



HAL
open science

Actes du congrès CIGI QUALITA 2021

Yannick Frein, Maurice Pillet

► **To cite this version:**

| Yannick Frein, Maurice Pillet (Dir.). Actes du congrès CIGI QUALITA 2021. 2021. hal-03236841

HAL Id: hal-03236841

<https://hal.science/hal-03236841v1>

Submitted on 26 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



CIGI QUALITA

Conférence internationale de Génie Industriel

5, 6 et 7 MAI 2021

Actes du congrès

CIGI QUALITA 2021

Grenoble Annecy

Président du comité d'organisation

- Michel Tollenaere

Présidents du comité scientifique

- Yannick Frein
- Maurice Pillet



Systems.Manufacturing.Academics.Resources.Technologies

CIGI QUALITA

Conférence internationale de Génie Industriel

5, 6 et 7 MAI 2021

THEMATIQUES

L'Humain au Cœur de l'Industrie du Futur sera le thème principal de cette conférence. Il sera centré autour des cinq transitions interreliées suivantes :

- Transition Numérique, d'une économie de la rareté des données aux opportunités liées à leur profusion.
- Transition Energétique, d'une économie de la profusion aux opportunités de la maîtrise et d'un meilleur rendement.
- Transition Ecologique vers une économie plus durable.
- Transition Sociétale vers une nouvelle place de l'homme dans l'industrie.
- Transition vers une industrie du futur à la lumière de la COVID 19.

PRESIDENTS DU COMITE SCIENTIFIQUE

YANNICK FREIN

G-SCOP Grenoble

MAURICE PILLET

SYMME Annecy

Conférences plénières

Vers les entreprises à mission dans l'industrie du futur,
- Blanche SEGRESTEIN, Mines Paristech

En route pour l'industrie 4.0 ! L'exemple de la fabrication de composants semiconducteurs
- Philippe VIALLETELLE, ST Microelectronics

L'industrie 4.0 au service de l'humain,
- Robert PELLERIN, Polytechnique Montréal

Transformation digitale en génie industriel : enjeux éthiques du monde de la recherche
- Damien TRENTESAUX, UPHF

Artificial intelligence in supply chain management: trends, challenges and opportunities,
- Gulgan ALPAN, G-SCOP



cigi-qualita21.sciencesconf.org



**Annecy
Grenoble**



Table des matières

Sessions Plénières	1
Artificial intelligence in supply chain management: trends, challenges and opportunities, Gülgün Alpan, G-SCOP	1
En route pour l'industrie 4.0 ! L'exemple de la fabrication de composants semi-conducteurs, Philippe VIALLETELLE, ST Microelectronics	15
L'industrie 4.0 au service de l'humain, Pr. Robert Pellerin, Polytechnique Montreal	32
Transformation digitale en génie industriel : enjeux éthiques du monde de la recherche, Pr. Damien Trentesaux, Université Polytechnique, Haut de Franc . . .	42
Session 1.1 : Design et Pilotage des systèmes de production de biens et de services	58
Métaheuristique pour le flow shop hybride avec machines dédiées, Ouled Bedhief Asma	58
Économie circulaire sous incertitude des délais : Revue de la littérature, Slama Ilhem [et al.]	67
PLANIFICATION DYNAMIQUE MULTILIGNES ET MULTIGAMMES AVEC MATIÈRES HÉTÉROGÈNES : UN MODÈLE DE BLENDING APPLIQUÉ À UNE CHAÎNE LOGISTIQUE MINIÈRE, Bamoumen Mouna [et al.]	75
A multi-product scheduling heuristic for a multiple-source slurry pipeline in the phosphate industry, Tchernev Nikolay [et al.]	84
Session 1.2 : Design et Pilotage des systèmes de production de biens et de services	92

Etude prospective d'un environnement d'aide à la décision rendu configurable par des modèles, Zhang Liwen [et al.]	92
Modélisation de CPPS : proposition d'un modèle générique, Arama Adama [et al.]	102
Exploring the role of Lean to cope with crisis: COVID-19., Saidi Hassani Alaoui Hajar [et al.]	109
Preserving resilience while managing changes, an assessment framework, Rigaud Eric	119

Session 1.3 : Design et Pilotage des systèmes de production de biens et de services **125**

Vers une approche globale pour soutenir le processus Adaptive Sales and Operations Planning, Vidal Jean-Baptiste [et al.]	125
Choix du niveau d'agrégation d'un processus S&OP dans un contexte de forte variété, Ghrab Yahya [et al.]	134
PERFORMANCE EVALUATION OF A CENTRALIZED MAINTENANCE WORKSHOP WITH REMANUFACTURING, Akbari Haghghi Sina [et al.]	145
Toward Sustainable Reconfigurable Manufacturing Systems (SRMS): Past, Present and Future, Dahmani Abdelhak [et al.]	152

Session 1.4 : Design et Pilotage des systèmes de production de biens et de services **153**

Redesign du mode de pilotage d'une chaîne logistique minière, Azzamouri Bassma [et al.]	153
Understanding Saturation of a dedicated Non-Elective Surgery Theater, Santamaria-Acevedo Gustavo [et al.]	162
Conception d'aménagement des pépinières forestières : Revue des pratiques et perspectives, Bélanger Olivier [et al.]	170
Factors influencing Drug Consumption and Prediction Methods, Koala Denis [et al.]	178
Conversational Bots with Reinforcement Learning for Multi-criteria Suppliers Selection Focused on post-COVID-19 Risk: Case of French Automobile Manufacturer, Chehbi Gamoura Samia [et al.]	186

Session 1.5 : Design et Pilotage des systèmes de production de biens et de services **194**

Coordination mechanisms in supply chains dealing with supply uncertainty: A state of the art, Lachance Emilie [et al.]	194
Planification et optimisation de la logistique de la construction dans les zones urbaines, Jarroudi Imen [et al.]	203
Conception d'un modèle de simulation pour l'évaluation de la pertinence de l'Internet Physique pour les chaînes d'approvisionnement humanitaires, Grest Manon [et al.]	211
Logistics Network Design in a Product-Service System case study, Metaireau Agathe [et al.]	219
Étude sur le potentiel de mutualisation logistique dans la grande région de Montréal, Tariq Sara [et al.]	228
Truck platooning transportation planning: A review with emphasis on Operations Research methods, Gazran Saba [et al.]	236

Session 2.1 - Session spéciale 3 : Surveillance intelligente des processus et maintenance prédictive pour l'industrie du futur **244**

Des modèles d'apprentissage d'ensemble pour la prédiction de l'état des systèmes de production, Gouiaa Amal [et al.]	244
Extraction et Exploitation de Règles Interprétables de Maintenance Cas d'un Four de Brasage, Hafdhallaoui Emna [et al.]	251
Processus de surveillance autonome des machines tournantes., Chimentin Xavier [et al.]	258
Dynamic grouping maintenance considering quality degradation, Vu Hai Canh [et al.]	264

Session 2.2 : Management de la Qualité & Amélioration continue **270**

Une démarche Lean dans la phase amont du cycle de vie produit – Application au processus opérationnel de l'Entreprise Stäubli Electrical Connectors SAS, Maranzana Sébastien [et al.]	270
Rôle de la conduite du changement dans la mise en place des systèmes de management, Alaoui Mdaghri Malika [et al.]	280

Proposition d'une méthodologie d'amélioration continue adaptée à une situation de crise. Application chez Safran Landing Systems pendant la crise du Covid-19, Minot-Lenfant Alicia [et al.]	287
Risk Analysis: Moroccan Drug Approval Process, Oullada Oumaima [et al.]	296
Continuous Improvement process model: A Knowledge Management approach, Cardenas-Cristancho Diana [et al.]	304

Session 2.3 - Session spéciale 3 et 7 : Surveillance intelligente des processus et maintenance prédictive pour l'industrie du futur **312**

Tolérancement GPS et Pilotage, Muller Thomas [et al.]	312
State Estimation considering the Sensor's Degradation, Hachem Hassan [et al.]	319
Combining First Principles and Machine Learning for optimal maintenance of electrical assets, Chevalier Marcel [et al.]	325
Incertitudes du pronostic de durée de vie de systèmes de production par IOHMM, Simon Christophe [et al.]	331

Session 2.4 - Session spéciale 12 : Ingénierie des exigences et sûreté de fonctionnement **339**

L'excellence opérationnelle au service de l'ingénierie système: Etat de l'art, Kokolo Joseph [et al.]	339
L'ingénierie dirigée par les modèles pour assurer la sécurité des logiciels embarqués en automobile, Sirgabsou Yandika [et al.]	347
Conception sûre des systèmes de production : intégration de la santé/sécurité des opérateurs à travers l'analyse des usages, Arias Corzo Sergio Orlando [et al.]	355
Application du MBSA avec AltaRica 3.0 dans un cadre système de systèmes pour le transport autonome, Batteux Michel [et al.]	363

Session 2.5 : Sûreté de Fonctionnement & Maintenance **371**

Optimisation de stratégies de maintenances par simulation stochastique AltaRica 3.0, Batteux Michel [et al.]	371
Stratégie intégrée de maintenance préventive périodique, de production et de contrôle qualité dans les systèmes manufacturiers, Bahria Nadia [et al.]	379

Estimating ignition probability of gas pipeline rupture using artificial neural networks modeling, Younsi Karim [et al.]	388
Production and maintenance policies for a deteriorated unreliable manufacturing system, Megoze Pomgha Patrick [et al.]	393
Stratégie intégrée de maintenance préventive conditionnelle, de production et de contrôle qualité dans les systèmes manufacturiers, Bahria Nadia [et al.]	394
Génération de modèles de fiabilité opérationnelle, Xu Changyi [et al.]	402
Session 3.1 : Conception, Innovation, Fabrication	408
Comment hybrider les avantages des approches agiles avec le phasage traditionnel pour les projets industriels ?, Gidel Thierry [et al.]	408
What are the origins of eco-innovations? Results from a case based reasoning approach on 300 eco-innovations, Pialot Olivier [et al.]	418
L'optimisation hybride multidisciplinaire pour concevoir les produits de demain, Cabarbaye Aurélien [et al.]	424
Improve idea generation process by using ideas graph representation compared to Linkography method, Dkhil Abdellatif [et al.]	431
Session 3.2 - Session spéciale 10 : Mesure de performance pour le développement durable et la responsabilité sociétale des organisations	439
BSC: a literature review and a roadmap towards its implementation in a healthcare network in the province of Québec, Ben Fradj Anes [et al.]	439
Toward an Innovative Risk- and Opportunity-Oriented System for SMEs' Decision-Makers, Ben Taleb Romain [et al.]	450
Vers un système d'évaluation de la performance globale et multidimensionnelle de l'usine du futur, Gentil Marie-Hélène [et al.]	458
Performance industrielle : une évolution intégrant l'éthique dans le contexte de l'Industrie 4.0, Berrah Lamia [et al.]	466
The key catalyst events for the evolution of the sustainable development concept, Belmabrouk Amal [et al.]	474
Session 3.3 - Session spéciale 14 : Supply chains of the future: Towards greener	

and more local supply chains	481
Simulation of a full containerization approach in road transportation: preliminary results in the waste management industry, Nozari Milad [et al.]	481
Insourcing versus outsourcing under economical and environmental considerations, and the role of cost sharing in case of outsourcing, Kandil Narjes [et al.]	489
Analysis of Price and Production Configuration in the Olive Oil Industry with organic and conventional products: The French Case, Nouira Imen [et al.]	490
On Retailers' Profit in A Price- and Greenness- Sensitive Market with Different Demand Functions, Asgari Erfan [et al.]	498
 Session 3.4 - Session spéciale 2 et 9 : Ingénierie de l'obsolescence et Ingénierie de la soutenabilité	 506
Sustainable urban factory: proposal of a viable scenario applied to Grenoble, Peigné Sarah [et al.]	506
Couplage de l'ingénierie des connaissances et de l'analyse du cycle de vie dans un contexte d'économie circulaire, Prioux Nancy [et al.]	514
Évaluation des pratiques de gestion de l'énergie dans l'industrie forestière, Hernández Thomas [et al.]	515
Prédiction du degré d'obsolescence en utilisant des techniques d'apprentissage automatique, Trabelsi Imen [et al.]	523
L'ingénierie des systèmes et les réseaux bayésiens pour la conception résiliente à l'obsolescence, Soltane Amel [et al.]	530
 Session 3.5 : Modélisation et gestion des entreprises, des connaissances et compétences	 537
Analysis of the forest industry market: A systematic literature review, Ferguene Mounia [et al.]	537
Vers un processus de changement plus efficient, Siebenborn Thierry [et al.]	546
Apports du marketing pour le management de projets innovants collaboratifs en lien avec une transition écologique et responsable, Galmiche Théo [et al.]	553
Use of Competency Management Methods and Tools in Project-based Organizations, Mlaouhi Kaouthar [et al.]	561

Supplying wood from private forests: Review of the barriers and means to increase mobilization, Le Minh [et al.] 569

Automation of competences extraction: HumaNeTrust, an experimental computer system used in an academic laboratory., Pouydebat Olivier [et al.] 577

Session 4.1 - Session spéciale 6 : Vérification et Validation en Ingénierie Système 585

Synthesis of software architecture for the control of embedded electrical generation and distribution system for aircraft under safety constraints: The case of simple failures, Zimmer Laurent [et al.] 585

Implementing Quality Function Deployment in Model Based Systems Engineering, Aouf Imen [et al.] 594

A Computer-Aided Verification Process for Engineered Systems, Romero Victor [et al.] 601

DGA – V&V anticipée pour les SDS, Havy Fabienne [et al.] 609

Session 4.2 - Session spéciale 8 : Transformation du secteur agroalimentaire à l'ère de l'industrie 4.0 615

Prédiction de la qualité de la farine après mouture, Parrenin Loic [et al.] 615

Implantation d'une architecture de CPS permettant l'interopérabilité des systèmes de production en contexte de PME, Piat Jean-Rémi [et al.] 625

Méthodologie d'étude de la rotation de cultures à partir des données, Dupuis Ambre [et al.] 634

C3PO : Une ontologie développée pour le pilotage de la production fromagère au Québec, Guilhen Lucas [et al.] 643

Estimation du poids et du volume des fruits et légumes via les technologies d'imagerie et de vision par ordinateur : état de l'art, Boudina Ahlam [et al.] . . . 651

Session 4.3 - Session spéciale 1 : Traitement du langage naturel dans l'industrie 4.0 et la chaîne logistique 659

Extraction générique de connaissances à partir de données textuelles et mesure de la performance des systèmes d'extraction de relations dans un contexte non supervisé., Chasseray Yohann [et al.] 659

L'analyse de sentiments pour la prévision de la demande en temps de crise : le cas de la pandémie de COVID-19, Nguyen Angie [et al.]	669
Etude de la Perte Focale pour le Traitement du Langage Naturel en Maintenance dans l'Industrie 4.0, Usuga Cadavid Juan Pablo [et al.]	675
Une approche pour la formalisation des connaissances en PHM : application aux PME, Omri Nabil [et al.]	682

Session 4.4 - Session spéciale 13 : L'excellence opérationnelle via la digitalisation des processus industriels **689**

Algorithme génétique pour l'optimisation intégrée du contrôle de la maintenance et de la qualité dans les systèmes manufacturiers, Harbaoui Imen [et al.]	689
Maturité du système physique de production : modèle de diagnostic 4.0, Damand David [et al.]	697
La transformation digitale de la chaîne d'approvisionnement, Mahraz Mohamed Iliasse [et al.]	705
Segmentation et caractérisation de la demande des magasins de détail : cas des produits cosmétiques, Laaroussi Mouhcine [et al.]	715

Session 4.5 : Industrie du futur - Transition numérique **723**

A Semantic Model for Cyber-physical production system inspired from Current developments, S A Puviyarasu [et al.]	723
Intégration verticale pour le pilotage des systèmes de production : un état de l'art, Cadoret Thomas [et al.]	730
Un cadre modulaire d'utilisation de Jumeaux Numériques dans un Système Manufacturier Reconfigurable, Haddou Benderbal Hichem [et al.]	738
Intégration d'objets connectés : Eléments d'aide au choix de plateforme IIoT, Santana Jussara [et al.]	746
Analyse de l'hybridation du QFD avec d'autres outils d'aide à la décision, Elomri Ameni [et al.]	754
Quality of sawmilling output predictions according to the size of the lot - The size matters!, Martineau Vincent [et al.]	762

Sessions Plénières



Plénière

Artificial intelligence in supply chain management:
trends, challenges and opportunities

Pr. Gulgun Alpan-Gaujaj, Grenoble INP

Artificial intelligence in supply chain management: trends, challenges and opportunities

Gülgün Alpan
UGA, Grenoble INP, Laboratoire G-SCOP

CIGI-Qualita21 : 5-7 May 2021
Grenoble FRANCE

1

Multidisciplinary Institute in Artificial Intelligence



- 2018 : Call in the context of priority research program (PIA3) as part of the national program for artificial intelligence
- One of the 4 Artificial Intelligence Institutes in France, accredited for an initial period of four years.
- Two main themes
 - Future AI systems: Machine learning (ML) and reasoning; Hardware and embedded architectures for AI; Perception and interaction.
 - AI for human beings and the environment (AI uses and applications): AI and society; Health; Environment and energy; Industry 4.0.
- 30 Chairs
- More than 55 companies
- Further information : <https://miai.univ-grenoble-alpes.fr/>

CIGI-Qualita21 : 5-7 May 2021
Grenoble FRANCE

2

AI for data-driven and reconfigurable supply chains Chair



Objectives

- Develop novel decision-making methods for highly reactive operations and supply chain management
 - Robust and (self)-adaptative solutions
 - With higher performance (better in time, quality and cost)

Perimeter

- Logistics, manufacturing, planning and control : at factory and supply chain level
- production and service systems

Team and Partners of the Chair



Senior scientific team

Gülgün Alpan, chair holder, Discrete Event Simulation, Operations Research
Hadrien Cambazard, Constraint Programming, AI
Iragaël Joly, Economics of transportation, Data Analytics
Pierre Lemaire, Optimization, Data Analytics
Bernard Penz, Operations Research
Abdourrahim Sylla, Knowledge based AI
Zakaria Yahouni, Scheduling, human centric industry 4.0



Laboratoire d'Économie Appliquée de Grenoble

Doctoral Students

Denis Koala, CIFRE @ KLS, Dec. 2019 – Dec. 2022
Suzanne Le Bihan, CIFRE @ Renault, Jan. 2020 – Jan. 2023
Yuzhen Wang, Financed by MIAI , Oct. 2020 – Oct 2023
Issa Bou Zeid, CIFRE@ Renault Jan. 2021-Jan 2024

International Partners

Fraunhofer Institutes IML (Dortmund) and IPA (Stuttgart)  **Fraunhofer**

Industrial Partners



Principles of AI

- Representing and processing information
 - Knowledge
 - Data
 - Often uncertain
- Reasoning and decision: formalizing and automation of human reasoning
- Learning: capacity to discover regularities from observation, self (re-)programming by experience
 - Machine learning (ML)
- Solving: Generic methods for problem solving
 - Constraint programming
 - Meta-heuristics
 - Multi-agent systems
 - ...

Intelligence artificielle: De quoi s'agit-il vraiment? Cepadues 2020, GDR IA

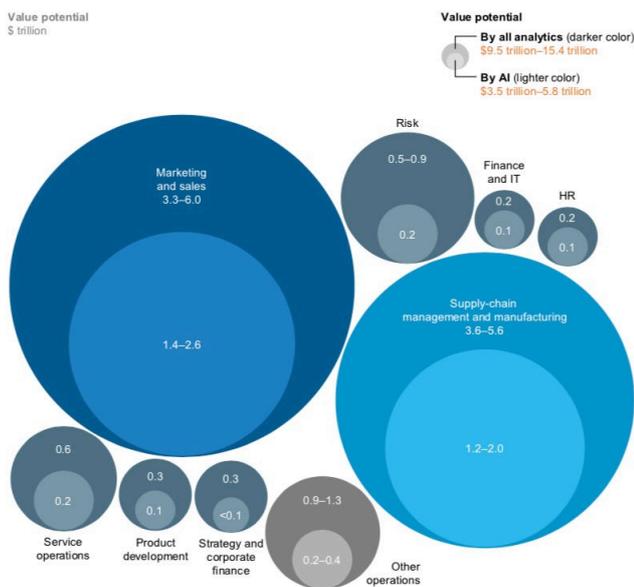
Main ML approaches

- **Supervised Learning** : training data X is labeled and known solutions Y to approximate a function $f(X)=Y$. Find Y for any new X
- **Unsupervised Learning**: Unlabeled dataset X. Explore X to find patterns and hidden structures.
- **Re-inforcement Learning**: by experimenting (trial-and-error) to maximize a reward

Trends: industrial point of view

Potential impact of AI in industry

AI's potential impact is greatest in marketing and sales and supply-chain management and manufacturing, based on our use cases



NOTE: Numbers may not sum due to rounding.
SOURCE: McKinsey Global Institute analysis

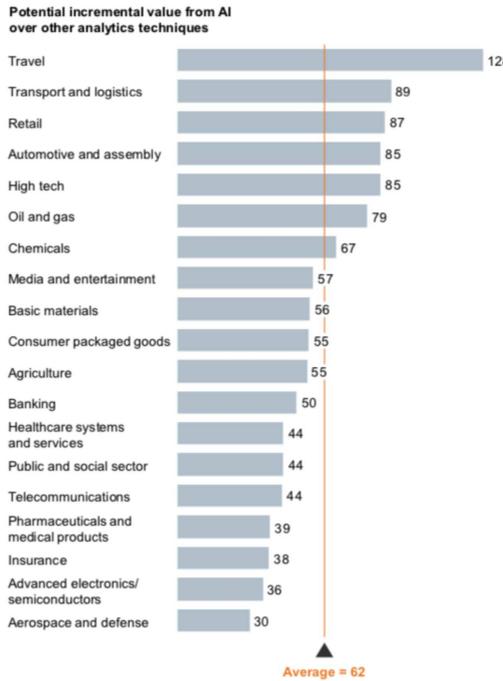
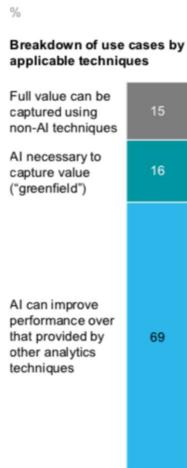
- >400 use cases across 19 industries
- Focus on deep learning
- Consumer facing industries (e.g. retail) high value in marketing and sales
- Advanced manufacturing industries (e.g. automotive) high value in supply chain, logistics and manufacturing

Source : Notes from the AI frontier: insight from hundreds of use cases. McKinsey Global Institute, Michael Chui, James Manyika, Mehdi Miremadi, Nicolaus Henke, Rita Chung, Pieter Nel, and Sankalp Malhotra, April 2018.

Impact by industrial sector

Exhibit 6

In more than two-thirds of our use cases, AI can improve performance beyond that provided by other analytics techniques

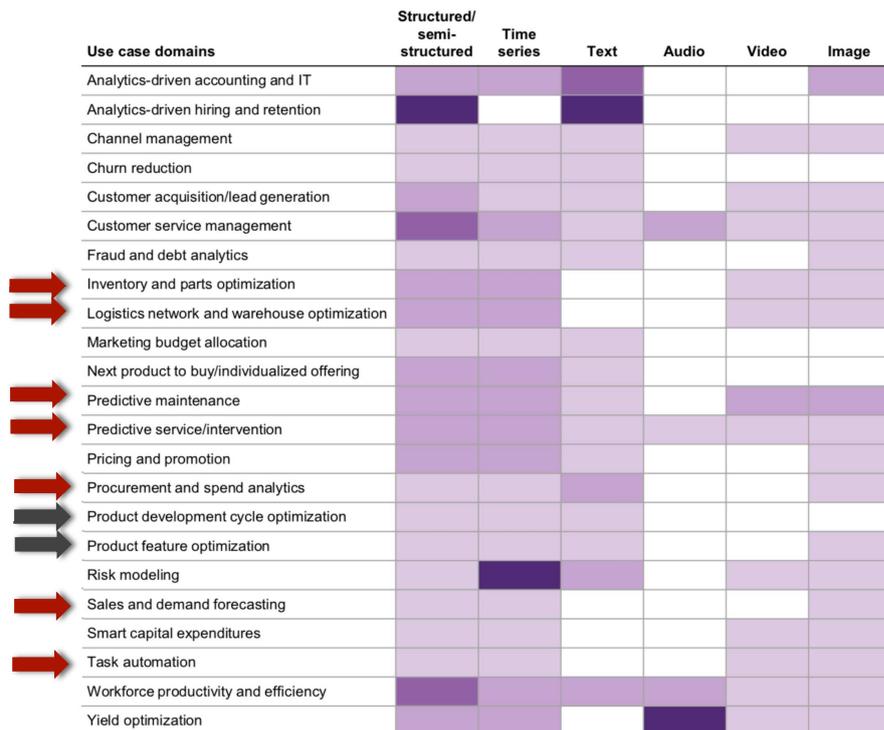


Source : Notes from the AI frontier: insight from hundreds of use cases. McKinsey Global Institute, Michael Chui, James Manyika, Mehdi Miremadi, Nicolaus Henke, Rita Chung, Pieter Nel, and Sankalp Malhotra, April 2018.

Problem types in McKinsey Case studies

Use case mapping to data types

Number of use cases Low High

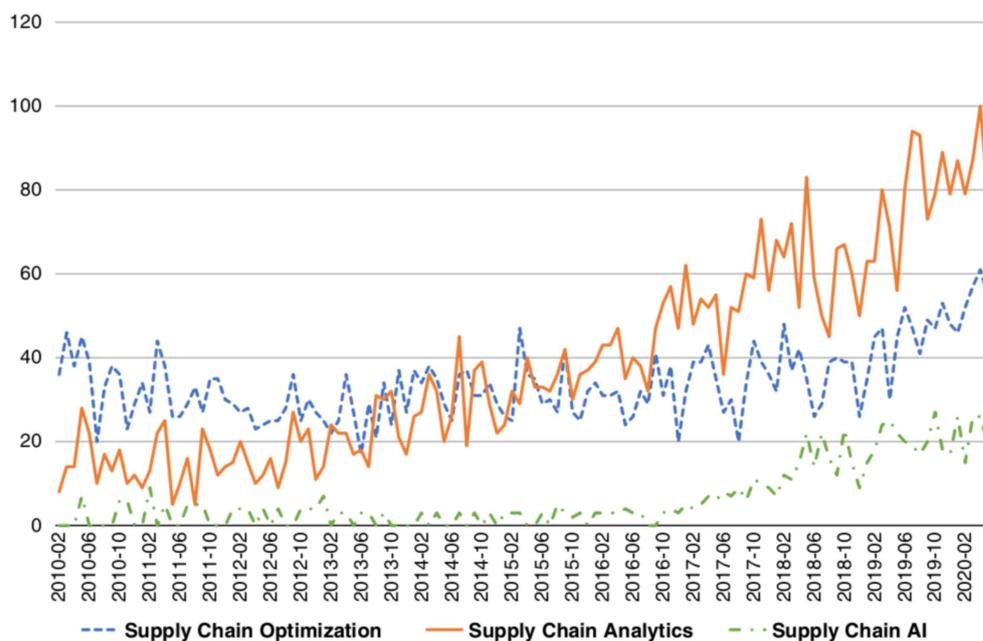


SOURCE: McKinsey Global Institute analysis

Source : Notes from the AI frontier: insight from hundreds of use cases. McKinsey Global Institute, Michael Chui, James Manyika, Mehdi Miremadi, Nicolaus Henke, Rita Chung, Pieter Nel, and Sankalp Malhotra, April 2018.

Trends: research point of view

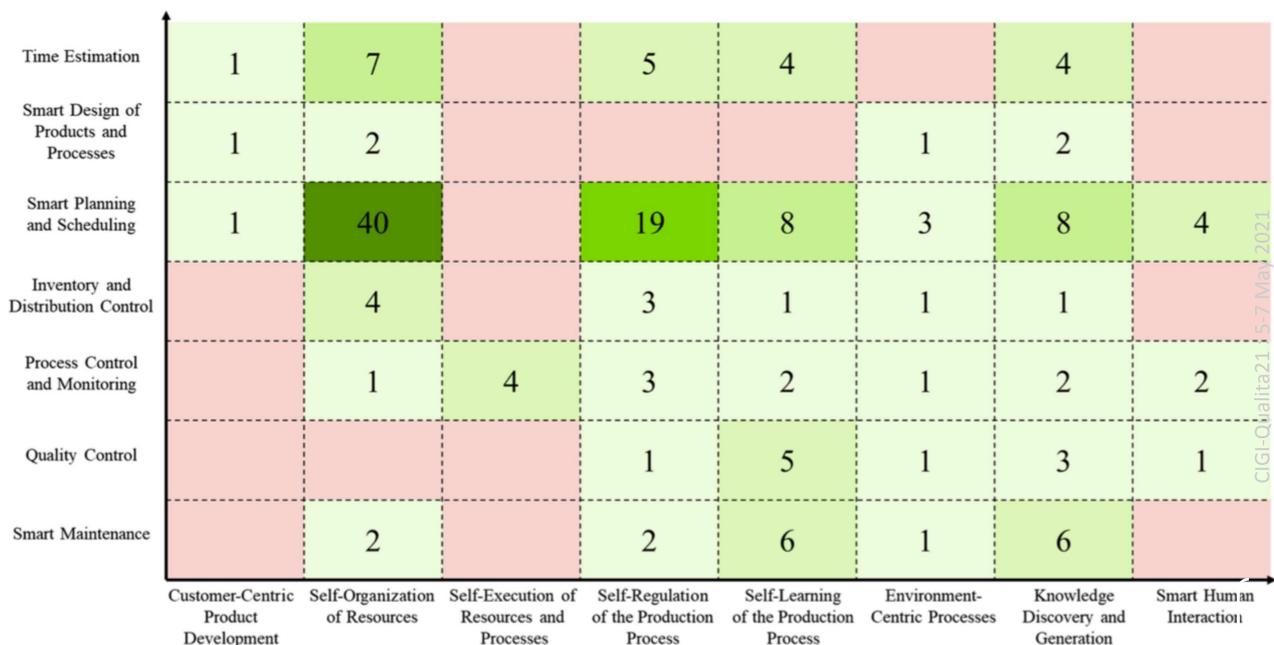
Relative Google search Frequencies last decade



Recent related review articles

- C. Gambelle, B. Ghaddar, J. Noam-Sawaya (2021), Optimization problems for machine learning: A survey, *EJOR* 290, 807-828
- D. Ivanov, C. Tang, A Dolgui, D. Battini, A. Das (2021) Researchers' perspectives on Industry 4.0: multi-disciplinary analysis and opportunities for operations management, *IJPR*, 59:7, 2055-2078.
- S. Gupta, S Modgil, S. Bhattacharya, I. Bose (2020), Artificial intelligence for decision support systems in the field of operations research: review and future scope of research, *Annals of OR*
- M. V. Velibor, G. Rerakis (2020), Data Analytics in Operations Management: A review, *M&SOM*, 22:1, 158-169.
- JPU Cadavid, S Lamouri, B Grabot, R Pellerin, A Fortin (2020), Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0, *JIM*, 31:6, 1531-1558.
- M. Qi, H-Y. Mak & Z-J M Shen (2020), Data-driven research in retail operations—A review, *NRL*, 67:8, 595- 616.
- G. Baryannis, S. Validi, S. Dani & G. Antoniou (2019) Supply chain risk management and artificial intelligence: state of the art and future research directions, *IJPR*, 57:7, 2179-2202.

ML and Production Planning and Control

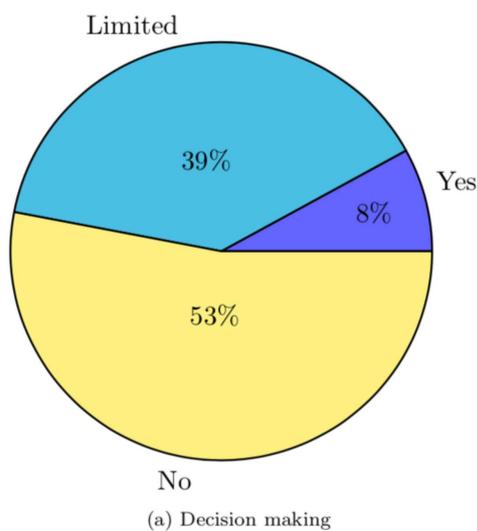


Successful use of ML is reported in Retail Operations and Revenue Management

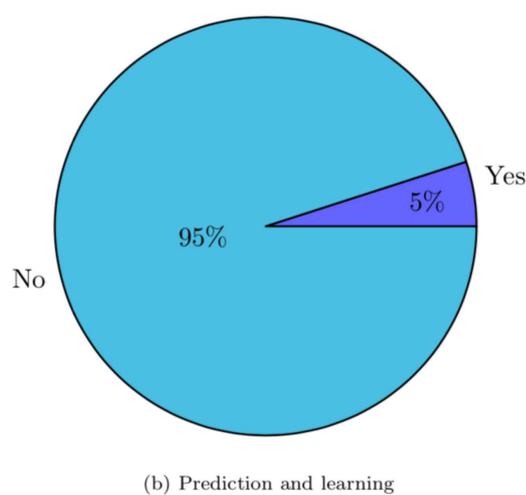
- Assortment optimization problem
 - Estimating choice model parameters from data
 - Forecasting customer behavior
- Fulfillment optimization models
 - Demande forecasting
 - Integrating prediction and optimization (e.g. in inventory management / fulfillment nodes)

M. Qi, et al. 2020, *Data-driven research in retail operations—A review*, *NRL*, 67:8, 595- 616.

AI is in its infancy in SCRM research

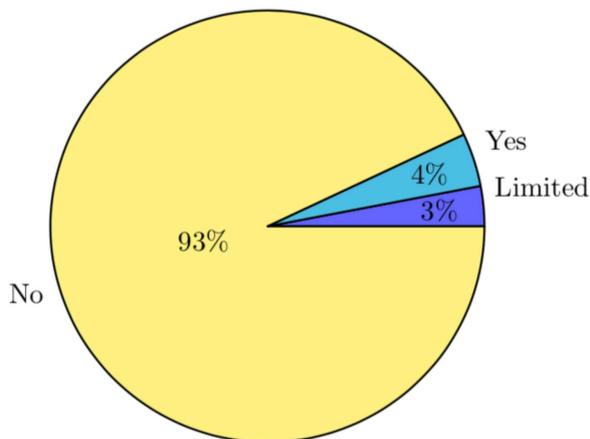


Very few propose applicable decision support tools



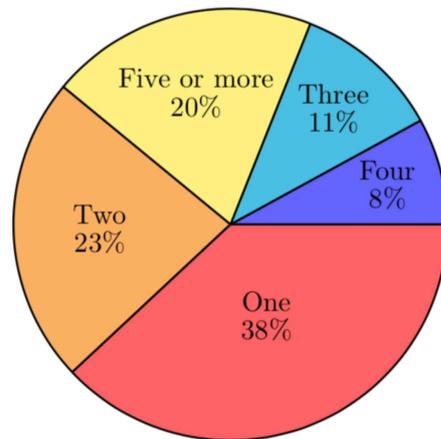
Little attention given to using knowledge gained to automate (even partially) decision making

Full potential of Data is not yet fully employed in SCRM reserach



(a) Large datasets

Very few studies exploit large data sets

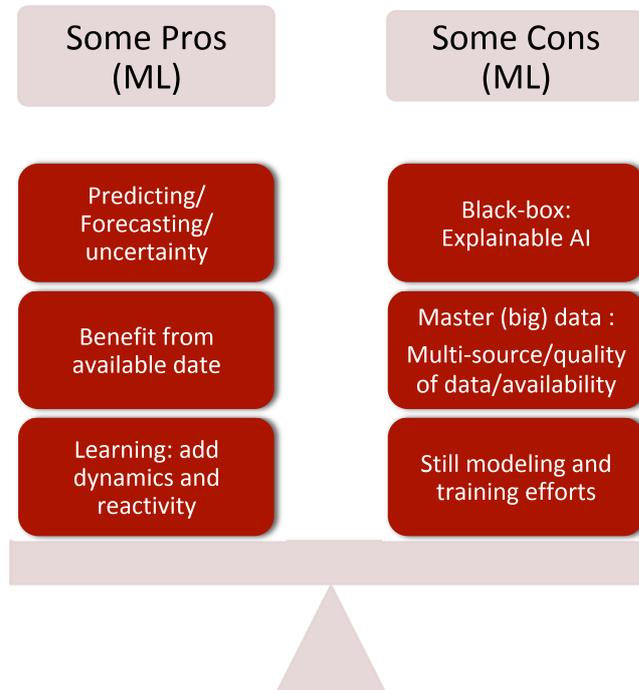


(b) Number of parameters

In terms of model complexity, most of the studies focus on a single uncertain parameter

G. Baryannis, et al(2019) Supply chain risk management and artificial intelligence: state of the art and future research directions, IJPR, 57:7, 2179-2202

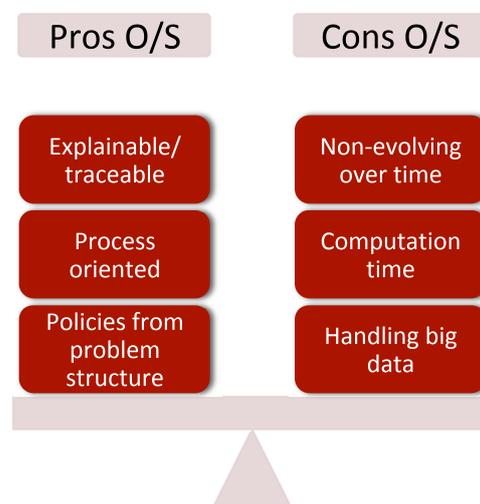
Opportunities and Challenges



Developing hybrid models



- MIAI Chair focuses on hybrid models
- Common consensus in literature review articles
- Complementarities of Optimization/Simulation and numerical AI



Some example hybridization ideas

- Use ML for parameter setting in optimization algorithms or target search space.
- Use optimal policies detected by optimization to enhance learning capabilities of ML algorithms
- ML to detect patterns in predictions provided by simulation
- Explain correlations detected by ML through simulation

Other Challenges

- New research opportunities in SCM and OM : 75% of areas related to PPC not explored. Very few works found in SCM and logistics (Cadavid et al 2020), areas to explore in SCRM (G. Bayanis et.al. 2019)), ...
- Many efforts on data is needed before (and after) applying ML
 - Collecting data
 - Need for adequate data : cleaning/ consolidation/conflicting data
 - Updating data (include new, remove obsolete data)
- Need for standard benchmark datasets for testing new models (Qi et al. 2020)
- Generic Yet modeling efforts are needed on application
- Benefit/cost analysis of using AI (e.g. Environmental impact of data intensive techniques)
- Human in the loop
- Need for interdisciplinary research
 - Privacy and security
 - Ethics
- ...

THANK YOU FOR YOUR ATTENTION
CONTACT : gulgun.alpan@grenoble-inp.fr

QUESTIONS?



Plénière

En route pour l'industrie 4.0 !

L'exemple de la fabrication de composants semiconducteurs

Philippe VIALLETELLE, ST Microelectronics



En route pour l'industrie 4.0 !

L'exemple de la fabrication de composants semiconducteurs.

Philippe Vialletelle
STMicroelectronics



Biography

Senior Member of Technical Staff at
STMicroelectronics France

Engineering degree in **Physics**

Topics of interest: production and process control, factory automation and **industry 4.0**

Currently busy in contributing to the development of a global **semantic model** within the Manufacturing Data & Analytics program which is today one of the key strategic priorities of ST





life.augmented

STMicroelectronics Company presentation

March 2021



We are creators and makers of technology



- One of the world's largest semiconductor companies
- 2020 revenues of **\$10.2 B**
- **46,000** employees of which **8,100** in R&D
- Over **80** Sales & marketing offices serving over **100,000** customers across the globe
- **11** Manufacturing sites
- Signatory of the United Nations Global Compact (UNGC), Member of the Responsible Business Alliance (RBA)

As of December 31, 2020

4

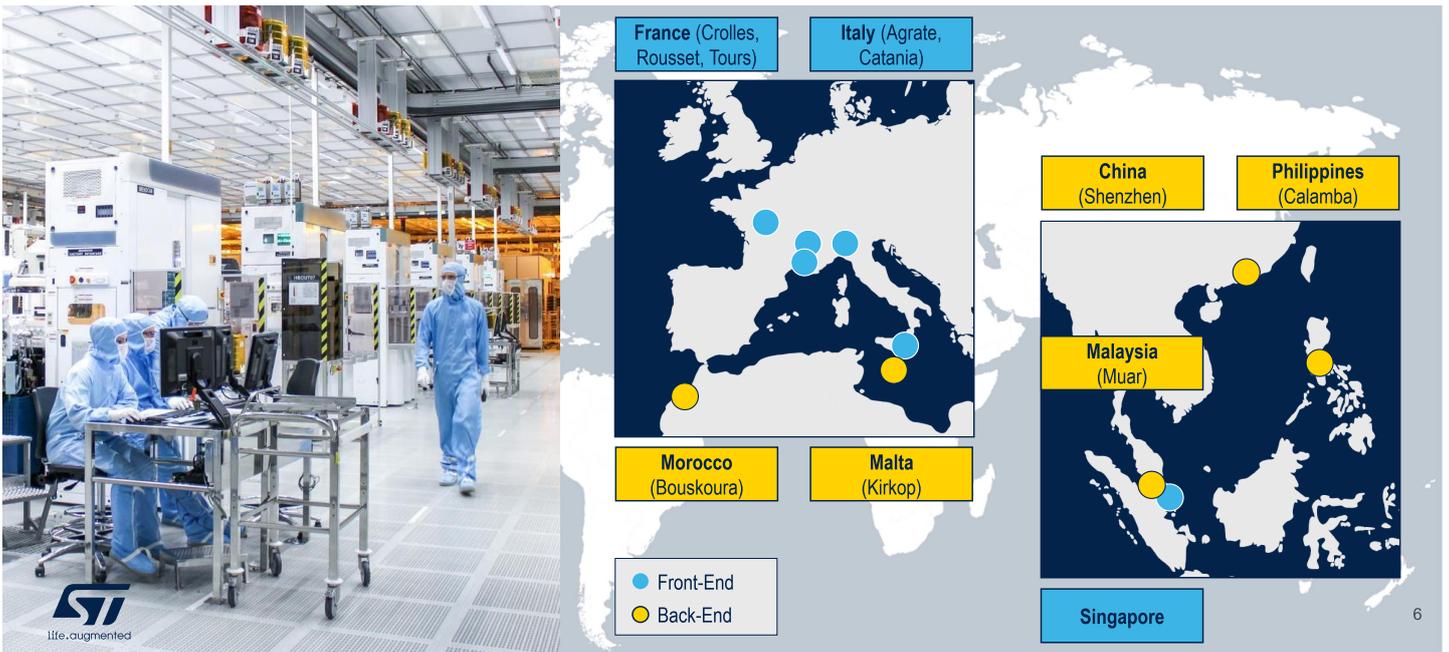


life.augmented

We address four end markets



We offer quality, flexibility and supply security



Bienvenue à ST Crolles!



L'univers de la salle blanche

5 équipes se relaient

 **24h/24**
7j/7
365j/365

Une qualité de l'air exceptionnelle

 L'air y est constamment filtré et renouvelé, **plusieurs fois par minute**
La température est contrôlée : **21,5°C (+/- 0,5°C)** toute l'année

Pourquoi autant se protéger ?

Parce qu'une poussière de **10 µm**
peut endommager jusqu'à **360** transistors !



(10 µm ~ 20% du diamètre d'un cheveu)

L'unité de production 300mm de Crolles fonctionne en mode automatique 24/7

- **100%** des démarrages équipement
- **100%** des transports de produits entre étapes élémentaires
- **94%** des décisions d'ordonnancement et de dispatching

L'APC (Advanced Process Control) est déployé sur 100% des équipements et des étapes de fabrication.

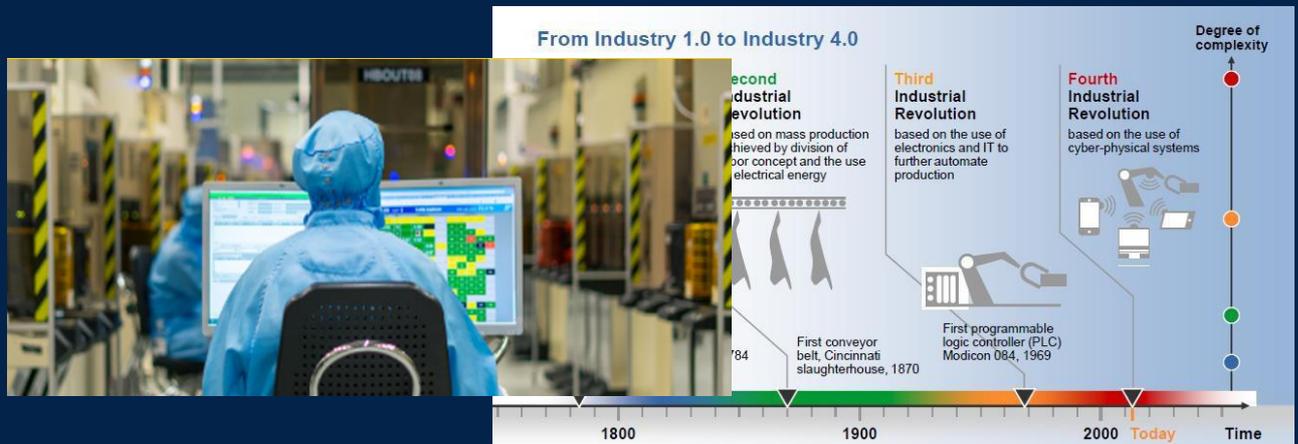
Système de transport automatisé au plafond.



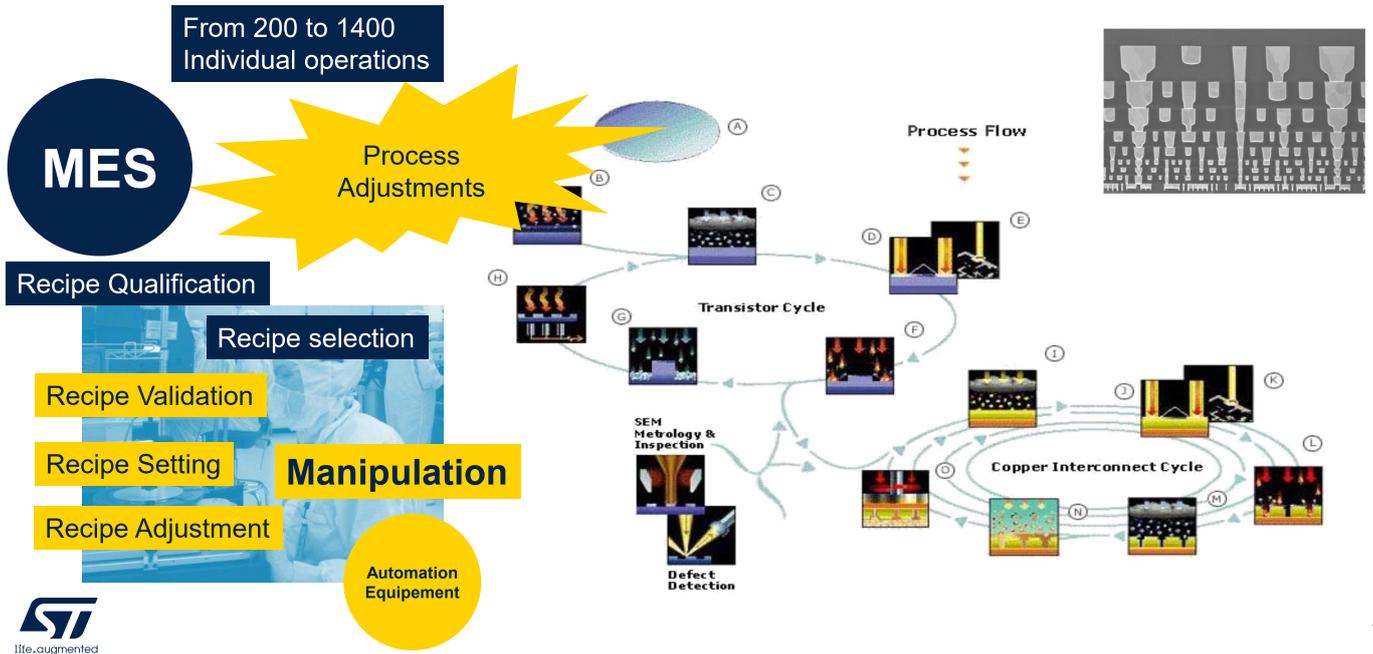
Les acteurs de la salle blanche vivent au quotidien l'Industrie 4.0!



Une longue route...

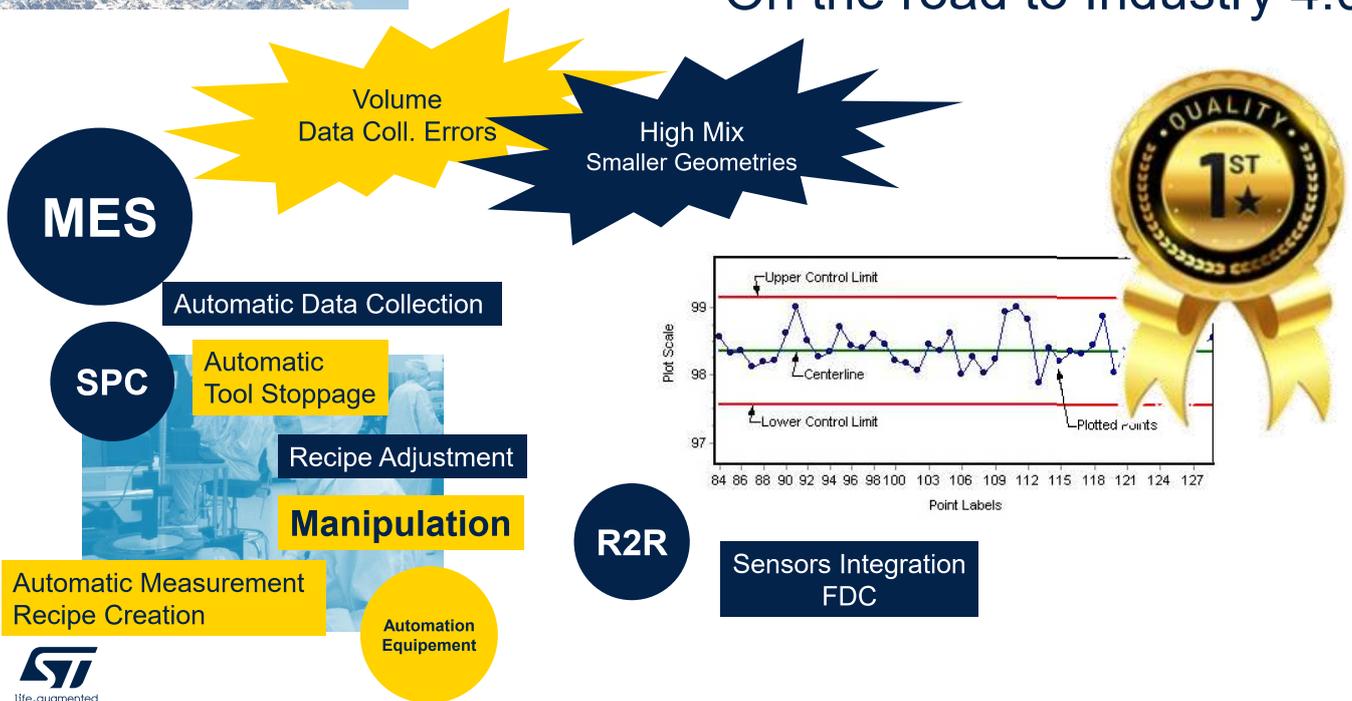


On the road to Industry 4.0



11

On the road to Industry 4.0

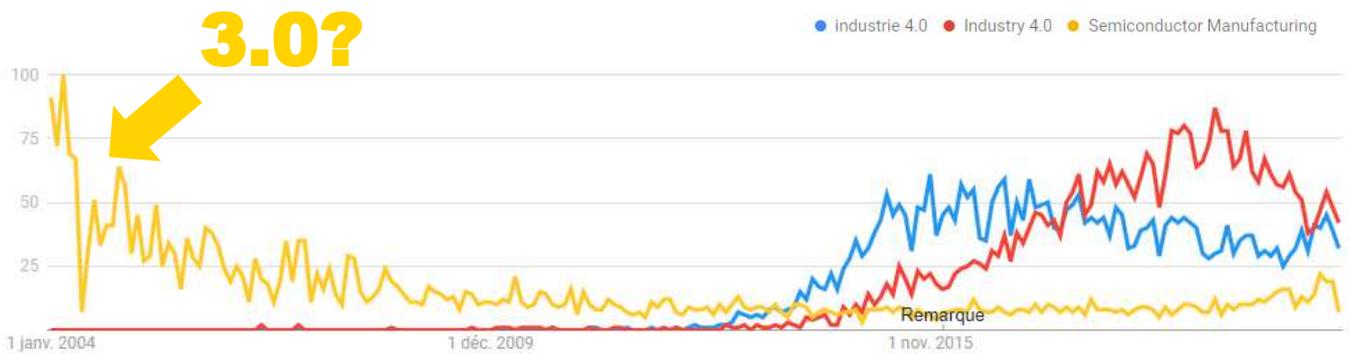


12

On the road to Industry 4.0



Intérêt pour la recherche Tous pays 2004 à aujourd'hui



Google Trends

On the road to Industry 4.0

300mm!

Clean Room sqm cost



AMHS



Logistics



Manipulation

Width	416 mm (16.4")
Depth	333 mm (13.1")
Height	335 mm (13.2")
Empty	4.2 kg (9.26 lb)
With wafers	7.3 kg (16.09 lb)



On the road to Industry 4.0

300mm!



AMHS



Decision

Real Time Scheduling



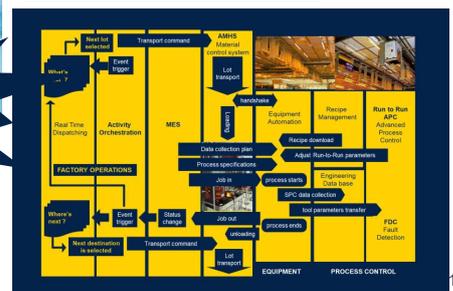
Logistics



Manipulation

Equipment complexity & heterogeneity

Real Time Dispatching



On the road to Industry 4.0



Prescription

Supervision

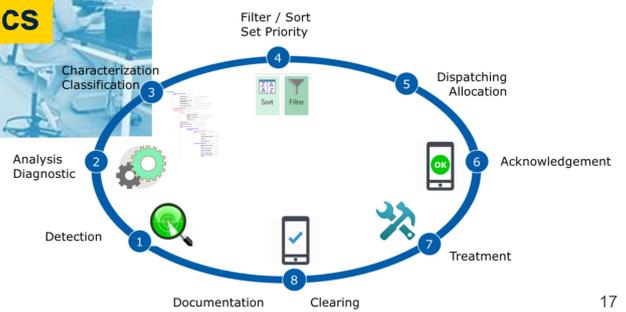
Decision

Diagnostic

Logistics



Manipulation



Et l'Humain 4.0 ?



Message 1

You're moving to a virtual world



200mm

- **Material** is visible
- **Problems are pushed** and disappear once fixed
- **Responsibility is implicit**

300mm

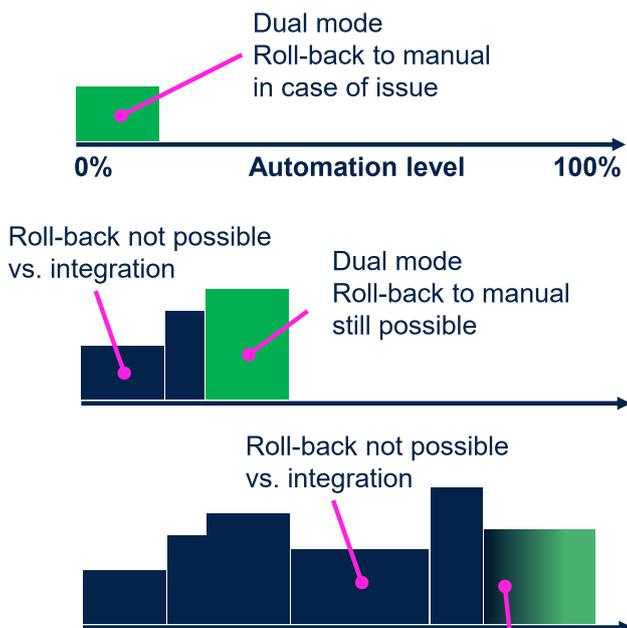
- **Virtual** objects
- **Information must be pulled** from the system
- **Responsibility is virtual**



19

Message 2

There is no turning back



- Partial manual modes (Semi-Auto) are transitory and less and less viable with time
- Global System's reliability and robustness become critical
- Manual operation becomes risky if not impossible and often detrimental to the system
- **Risk is moving from action to configuration / parametrization**

Roll-back impossible because no-one remembers the way it was done before!

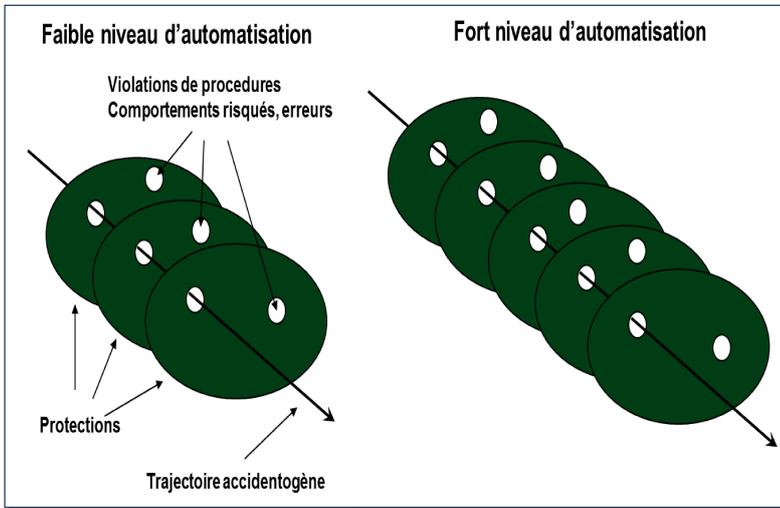


20

Message 3

Risk's perception is changing!

Risk is moving from action to configuration / parametrization



- Risk is still there
- But its perception decreases with the increase in automation
- Error becomes systematic and its consequences often heavier.



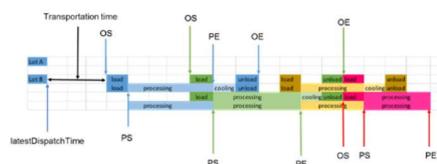
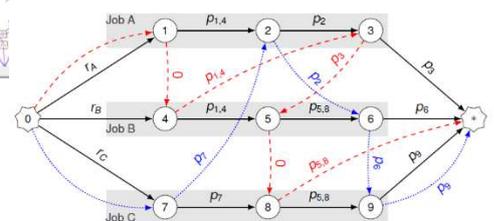
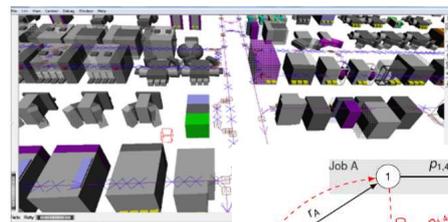
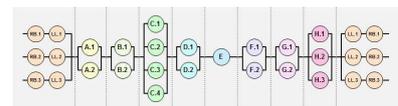
Important Warning:
Doesn't work in real world!



Message 4

There is no magic!

- Modelling activities are key (Digital Twin)
- Model Maintenance becomes critical with time
- Even the best system only knows... What's in the system!
- Understanding interactions at play in the system becomes a key competence



Le double effet boîte noire 1/2



- Tout est automatique !
- Tout a été pensé et optimisé !
- Tout est sécurisé !
- Le « système » va s'en occuper...
- Ca marche tout seul maintenant, non ?

Ce qui n'est pas connu du système
N'est pas pris en compte par le système.

Le double effet boîte noire 2/2

- Tout est verrouillé !
- On ne peut rien faire !
- Je tourne et il ne se passe rien !
- Le système fait n'importe quoi !
- Vous êtes trop compliqués !



Le système fait ce pour quoi il a été conçu
avec les règles qu'on lui a données.

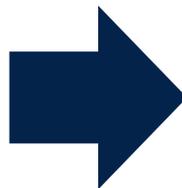
Message 5

Replacing the “human glue” is by far the most difficult!



Operator
 Drives the machines and moves the lots
 Physical presence on the production floor
 Uses his 5 senses to continuously analyze the situation.

Indoor patrols, Real-time detection of problems in the field and almost instant diagnosis

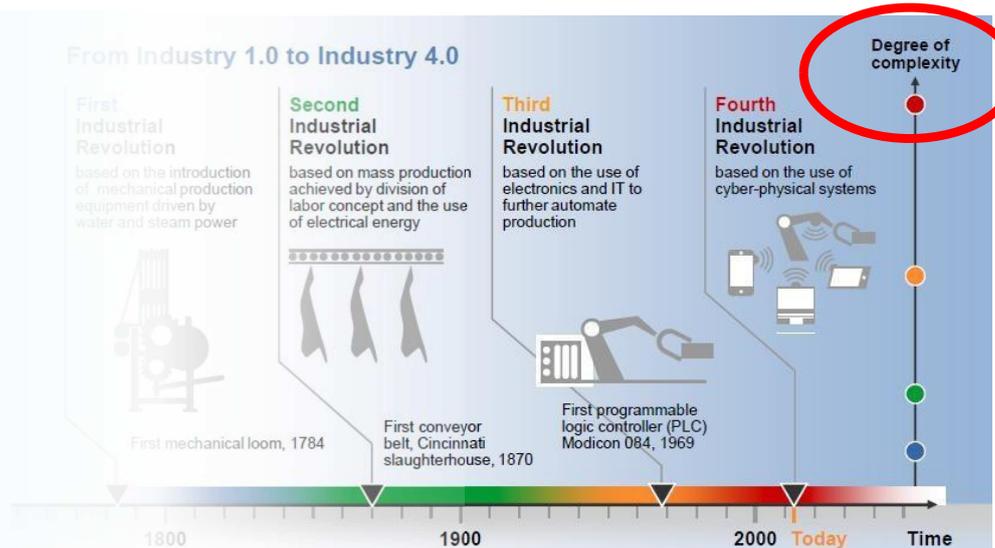


Technician
 Technical field expert
 Maintenance & Process diagnostic
 Resolving issues on his perimeter



life.augmented

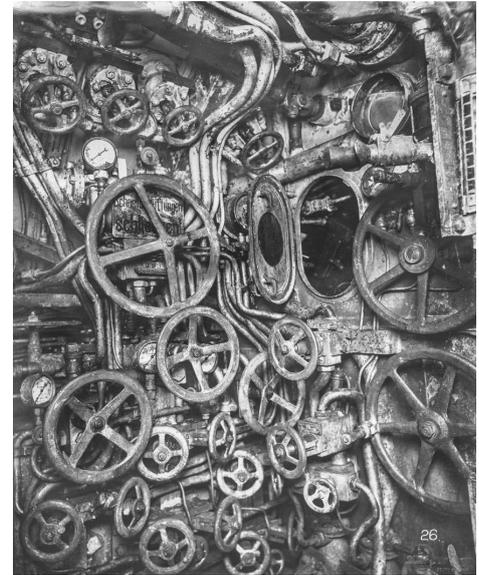
Une tâche (très) complexe



life.augmented

Pour éviter la « boîte noire »

- Choix judicieux des projets et des solutions
- Remise en question des « pratiques ancestrales »



Axes de collaboration

- Cartographie des besoins métiers par type d'industrie
 - Identifier les contextes « métier »
 - Identifier les solutions Industrie 4.0 adaptées
- Caractériser les domaines d'application des approches algorithmiques
 - Support académique aux industriels (algorithmes)
 - Support industriel à la recherche (data sets)
- Hybridation des techniques algorithmiques
 - Data Driven Automation
 - Remplacer ce qui est « systématique »



Pour éviter la « boîte noire »

- Choix judicieux des projets et des solutions
- Remise en question des « pratiques ancestrales »

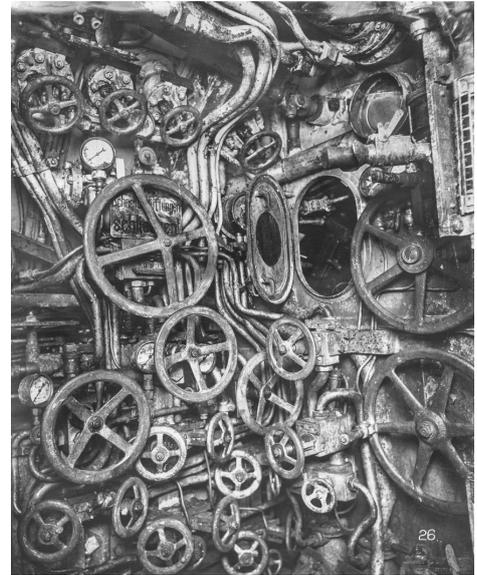
Gestion du risque

Responsabilisation

Circulation de l'information

Transfert des connaissances

Accompagnement du changement



life.augmented

29

**“ Face au monde qui change,
il vaut mieux
penser le changement
que changer le pansement. ”**

Francis Blanche



life.augmented

30

Merci.



© STMicroelectronics - All rights reserved.

ST logo is a trademark or a registered trademark of STMicroelectronics International NV or its affiliates in the EU and/or other countries.

For additional information about ST trademarks, please refer to www.st.com/trademarks.

All other product or service names are the property of their respective owners.





Plénière

L'industrie 4.0 au service de l'humain

Pr. Robert Pëllerin, Polytechnique Montreal

L'industrie 4.0 au service de l'humain

Robert PELLERIN, ing., Ph.D., CPIM, CD, PMP

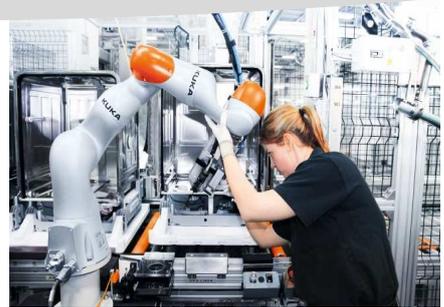
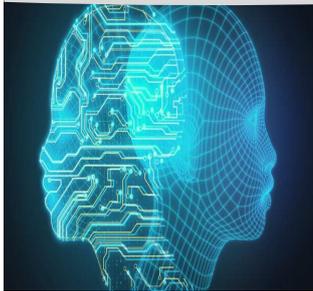
Professeur titulaire

Département de mathématiques et de génie industrie - Polytechnique Montréal

Titulaire de la Chaire de recherche Jarislowsky/SNC-Lavalin

Membre du CIRRELT, POLY-Industries 4.0 et IVADO

robert.pellerin@polymtl.ca



POLYTECHNIQUE MONTRÉAL - Département de mathématiques et de génie industriel

6 mai 2021

Plan de la présentation

Impacts du 4.0 sur l'humain



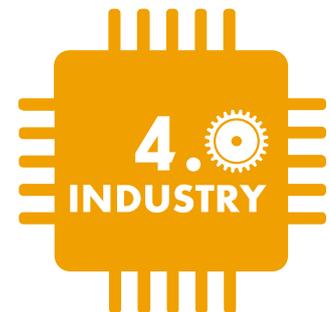
Recherche en GI



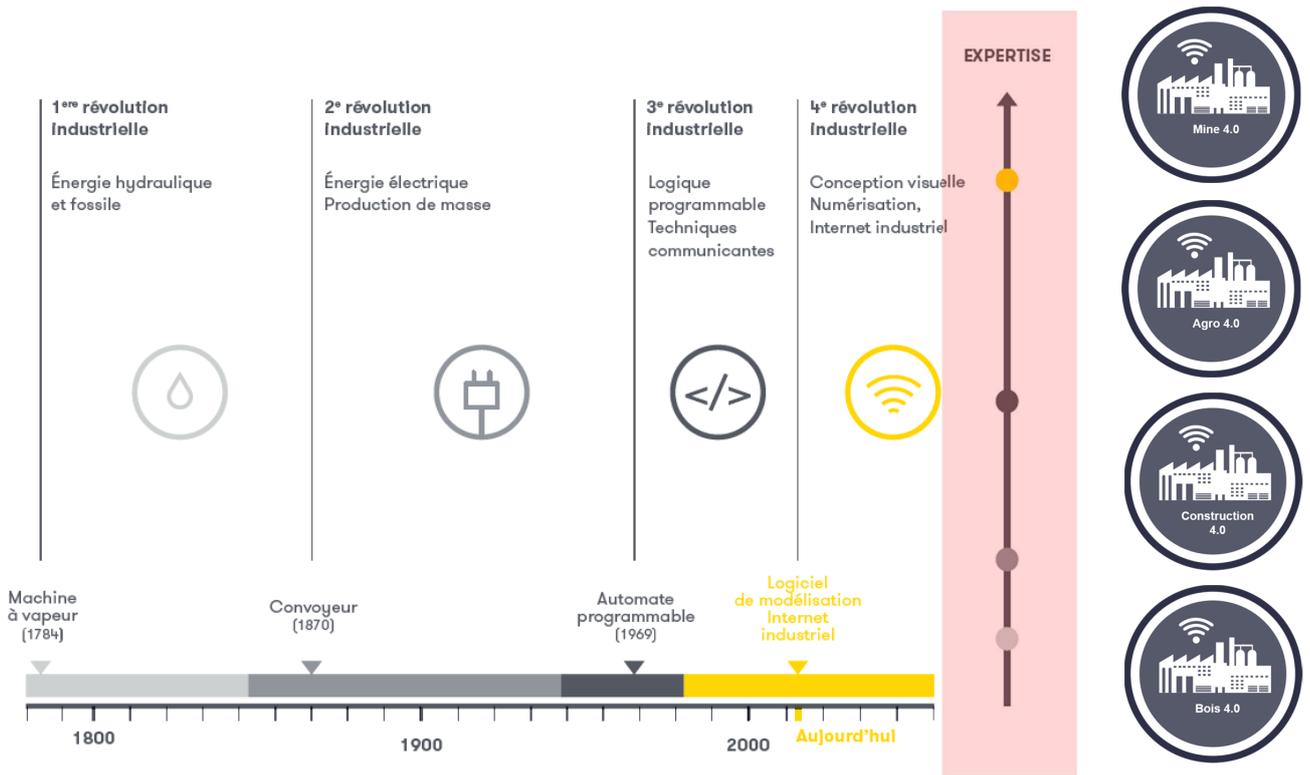
Constats



Perspectives



Révolution industrielle ou du savoir?

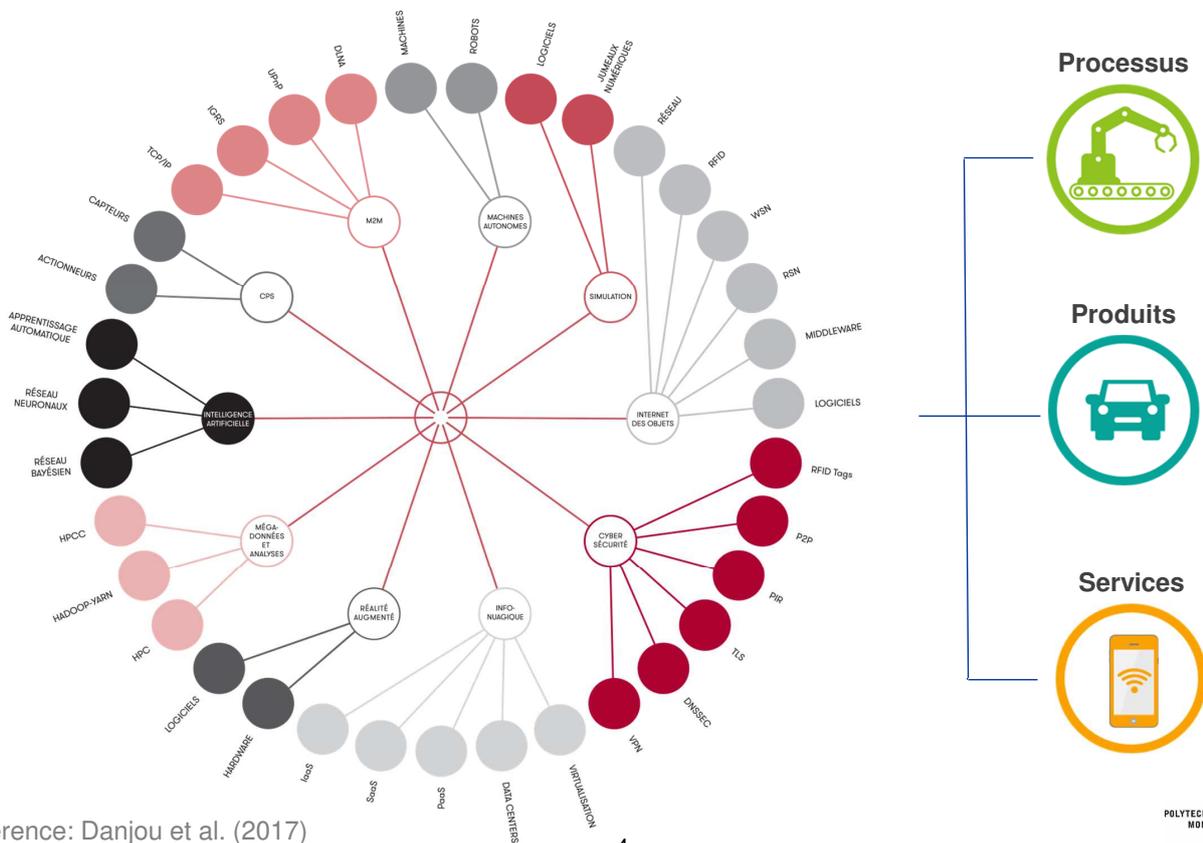


CIGI Qualita 2021

3



Leviers technologiques



Référence: Danjou et al. (2017)

CIGI Qualita 2021

4



Transformation des opérations et des opérateurs



Surveillance



Les capteurs et les sources de données externes permettent la surveillance de :

- La condition du produit
- L'environnement extérieur
- L'utilisation et l'opération du produit

La surveillance permet également de produire des alertes et des notifications pour informer des changements



Contrôle



Les logiciels embarqués dans les produits permettent :

- Le contrôle des fonctions produits
- La personnalisation de l'expérience de l'utilisateur



Optimisation



Les capacités de surveillance et de contrôle permettent aux algorithmes d'optimiser les utilisations du produit pour :

- Améliorer les performances
- Permettre des diagnostics préventifs, du service, et d'effectuer des réparations



Autonomie



La combinaison de la surveillance du contrôle et de l'optimisation permet d'avoir :

- Un produit autonome
- Une coordination des opérations avec les autres produits et systèmes
- Une personnalisation des produits en autonomie
- Des diagnostics de fonctionnement et des services

Référence: Danjou et al. (2017)

CIGI Qualita 2021

5



Impacts sur les employés

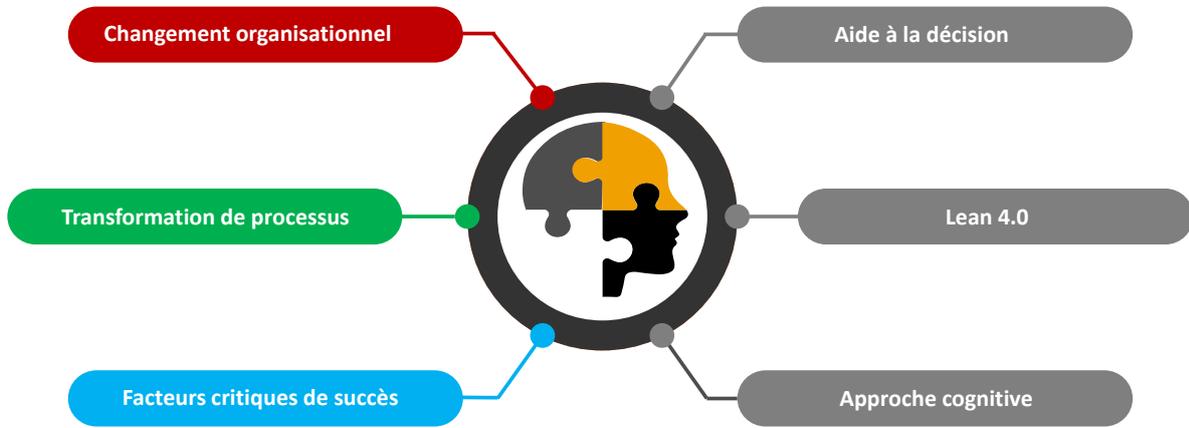
- Avec l'impulsion des technologies et des pratiques 4.0, la nature du travail se transforme et nous amène à :
 - Revoir le rôle des opérateurs;
 - Adapter les postes de travail;
 - Développer des nouvelles compétences auprès des employés;
 - Bonifier les solutions techniques de mécanismes d'aide cognitive pour faciliter et pérenniser leur utilisation par des opérateurs;
 - Revoir toutes les facettes de la gestion des ressources humaines :
 - Recrutement;
 - Formation;
 - Motivation;
 - Évaluation de la performance;
 - Rémunération;
 - ...

CIGI Qualita 2021

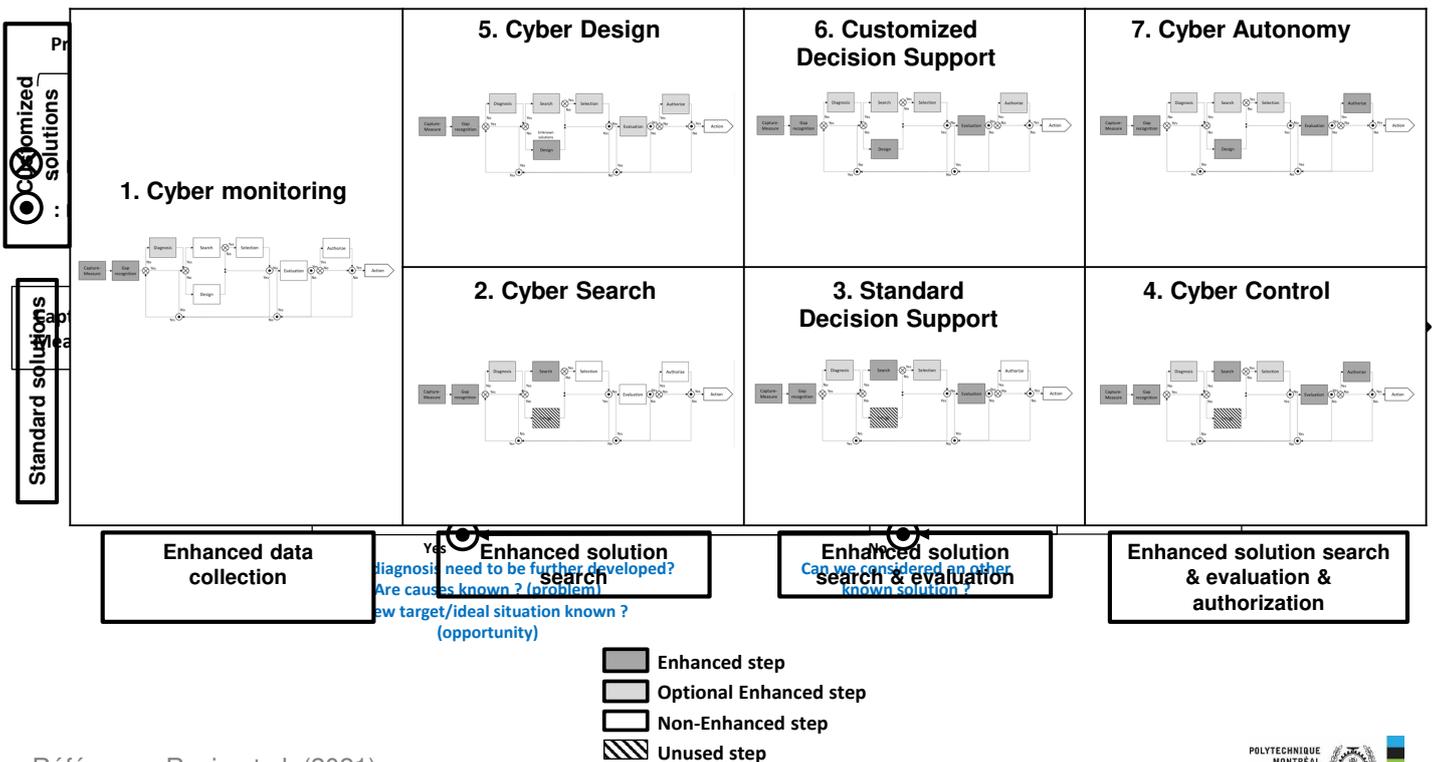
6



Recherche 4.0 en génie industriel axée sur l'humain



Recherche - Aide à la décision



Référence: Rosin et al. (2021)

Synergie Lean 4.0 et Industrie 4.0

		Autonomous robots	Simulation	System integration	Internet of Things	Cloud	Augmented reality	Big data and analytics	Cybersecurity		Monitoring	Control	Optimization	Autonomy				
Just-in-time	Takt time planning									0								
	Continuous flow	4	10	4	9	3		1		31	10		15	6				
	Pull system		2		7	1				10	6		2	2				
	Quick changeover		1							1			1					
	Integrated logistics	2		1	1					4	2			2				
Jidoka	Automatic stops				1					1		1						
	Andon				3					3	2	1						
	Person-machine separation	2								2				2				
	Error-proofing	2			3		2			7	4	1	2	2				
	In-station quality control	1			5	1		4		11	5		1	5				
Waste Reduction	Solve root cause of problem									0								
	Genshi Genbutsu									0								
	5 Why's									0								
	Eyes for waste		4	1	6			1		12	12							
	Problem solving									0								
People and Team work	Selection									0								
	Common goals									0								
	Ringi decision making									0								
Foundations	Cross-trained						4			5	4		1					
	Continuous improvement		2	1			2	1		6	3		3					
	Leveled production (heijunka)							1		1			1					
	Stable and Standardized processes	2					1	1		3	1			2				
	Visual Management		1		10	3	1	2		17	17							
Toyota Way Philosophy			2						2	2								
										13	21	9	45	8	10	10	0	116
			8	7	31	6	9	7			68							
					3						3							
			13	1	5	1	1	3					24					
			13	1	6	1								21				

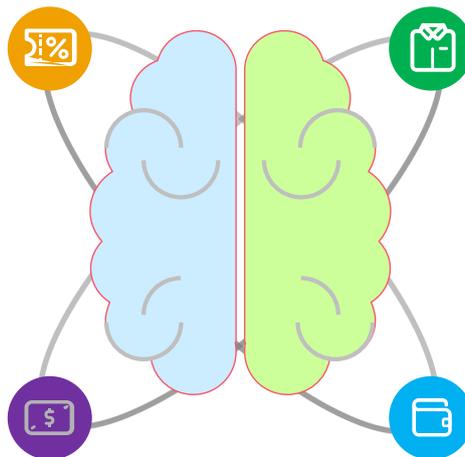


Référence: Rosin et al. (2017)

CIGI Qualita 2021

Approche cognitive

Impact de l'IA et de la transformation numérique sur l'employé

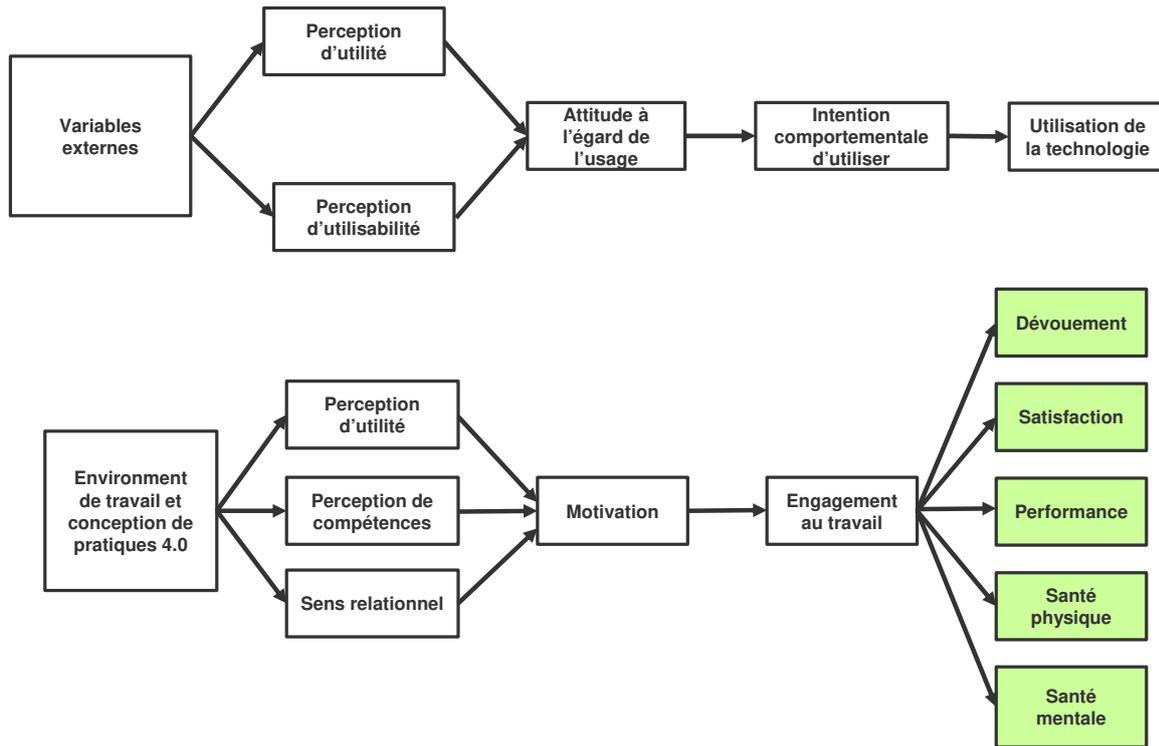


Adaptation des postes et procédures de travail

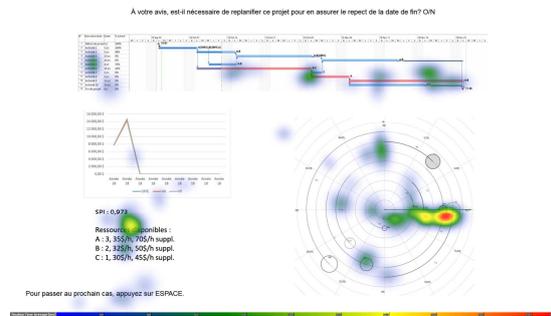
Ux – Expérience utilisateur

Symbiose humaine - technologie

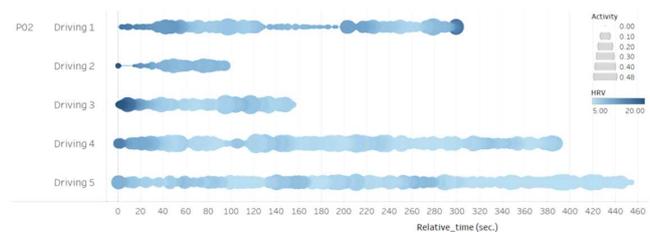
Une perspective différente de la conception 4.0



Des outils méthodologiques centrés sur l'humain



Visualisation des zones d'attention



Visualisation de la fatigue et de la réactivité émotionnelle

Références: Tech3Lab (site web) et Isac et al. (2020)

Constats

- L'amélioration de la productivité des systèmes demeure toujours la priorité d'une vaste majorité des chercheurs en génie industriel et **l'impact sur les employés est souvent étudié de façon secondaire et non structurée.**
- La collaboration en recherche en génie industriel s'est accentuée dernièrement avec un élargissement des équipes afin de combiner de multiples compétences et technologies. En revanche, les équipes de recherche demeurent très centrées sur des approches méthodologiques et des **techniques expérimentales peu variées.**
- Une grande part des travaux encourage une transformation de la nature du travail de l'opérateur vers un rôle de gestion d'alertes et d'exceptions **sans aborder les conséquences** en termes de :
 - Charge cognitive exceptionnellement variable;
 - Complexité liée à la formation des compétences sur des actions très rares;
 - Évaluation de la performance des employés.

Perspectives de recherche

- Modélisation des mécanismes de prise de décision et de gestion d'exceptions par les opérateurs au sein de processus d'affaires.
- Méthodes et mesures de comparaison (et instances de tests) permettant d'évaluer de nouvelles pratiques 4.0 sur des bases opérationnelles, mais aussi cognitives et émotionnelles.
- Migration d'une stratégie d'apprentissage automatique vers des modèles de co-apprentissage humain-machine à long terme avec prise en compte du niveau de motivation, d'utilité et d'engagement.

Conclusion

- L'industrie 4.0 demeurera adéquate pour connecter et améliorer les produits, processus et services des organisations.
- Les approches actuelles sont particulièrement utiles pour les opérateurs en atténuant le fardeau de tâches répétitives et leur procurant de nouveaux outils numériques.
- Au-delà des impératifs de productivité, les outils actuels du 4.0 peuvent aussi devenir source de génération de santé physique et mentale pour les employés.
- La prochaine révolution, l'Industrie 5.0, doit se distinguer par une préoccupation centrale sur l'humain visant à véritablement intégrer technologies et employés sur le long terme.
- Cette transformation passe nécessairement par l'emploi de techniques peu exploitées par l'ingénieur industriel et par le recours à de la véritable co-conception avec les travailleurs.

Session plénière – CIGI Qalita 21

L'industrie 4.0 au service de l'humain

Robert PELLERIN, ing., Ph.D., CPIM, CD, PMP

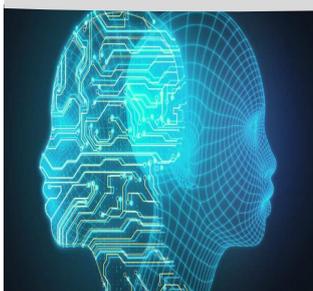
Professeur titulaire

Département de mathématiques et de génie industriel

Polytechnique Montréal

Membre du CIRRELT, POLY-Industries 4.0 et IVADO

robert.pellerin@polymtl.ca



- DANJOU, C., RIVEST, L. et PELLERIN, R. (2017), " Douze positionnements stratégiques pour l'Industrie 4.0 : entre processus, produit et service, de la surveillance à l'autonomie", CIGI 2017, 3-5 mai, Compiègne, France.
- ISAC, J., PELLERIN, R., and LÉGER, P.M. (2020), "Impact of a Visual Decision Support Tool in Project Control: A Comparative Study Using Eye Tracking", *Automation in Construction*, Vol. 110, pp. 1-19.
- ROSIN, F., FORGET, P., LAMOURI, S., et PELLERIN, R. (2020), "Impacts of Industry 4.0 technologies on Lean principles", *International Journal of Production Research*, Vol. 58, No. 6, pp. 1644-1661.
- ROSIN, F., FORGET, P., LAMOURI, S., and PELLERIN, R. (2020), "Industry4.0 and decision-making", JCM 2020, Aix-en-Provence, France, June 2-4, 2020.
- Tech3Lab, <https://tech3lab.hec.ca/>.



Plénière

Transformation digitale en génie industriel :
enjeux éthiques du monde de la recherche,

Pr. Damien Trentesaux, Université Polytechnique, Haut de France



Transformation digitale en génie industriel : enjeux éthiques du monde de la recherche

Damien Trentesaux

Damien.Trentesaux@uphf.fr



Partie I

Contextualisation et enjeux de la digitalisation des entreprises (GI)

Domaine civil: production, logistique, santé...

Une évolution des mentalités et des attentes sociétales

- Production/service solidaire
 - Impression 3D masques, les retraités à la rescousse...
 - Rétribution équitable des producteurs
 - Production locale («manufacturing as my neighbour »)
- Chaines logistiques humanitaires
- Des industriels en crise existentielle sur leur finalité
 - Des jeunes ingénieurs qui aspirent à donner du sens à leur vie
- Un modèle économique à bout de souffle en quête de renouveau
 - Empreinte environnementale, ressources limités
 - Une envie de ne pas trop emprunter à nos enfants
- Emergence de « l'impact sociétal » dans nos communautés
 - Bien-être des opérateurs...



TC 9.2. Systems and Control for Societal Impact

Welcome to the TC9.2. Systems and Control for Societal Impact. Please visit our pages to discover the main topics addressed by the TC9.2, the members and sponsored and co-sponsored events. If you wish to contact us, then please send an email to the Chair Dr. Mariana Netto, email: mariana.netto@univ-eiffel.fr. We thank you in advance for your interest in the IFAC TC9.2.



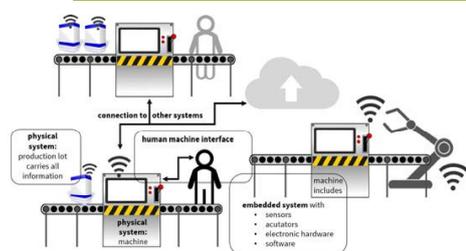
Damien Trentesaux

Transformation digitale des industries et services

La maturité de ces nouvelles technologies permet désormais des applications réelles

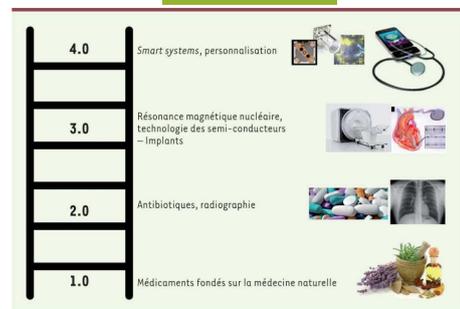
Des concepts et visions rapprochant matière, homme et information

Systèmes cyber-physiques/ Industrie 4.0



<http://dx.doi.org/10.15480/882.1458>

Médecine 4.0



<https://doi.org/10.1051/medsci/20183405019>

Damien Trentesaux

4



Transformation digitale des entreprises

Des technologies favorisant l'autonomie décisionnelle des systèmes industriels et produits

Cobotique , systèmes embarqués
Robotique mobile

Intelligence artificielle
Architectures IIoT, edge, fog, cloud



SurferLab

Damien Trentesaux

5

Enjeux des nouvelles technologies digitales

Mais les implications de la science sont trop sérieuses pour être laissées aux scientifiques, et encore moins aux industriels, aussi séduisants et «hype» soient-ils. Chacune des nouvelles thérapies doit être rendue publique, évaluée, scrutée, anticipée.

Hors série Libération, 6/12/18: l'IA, au cœur de l'humain



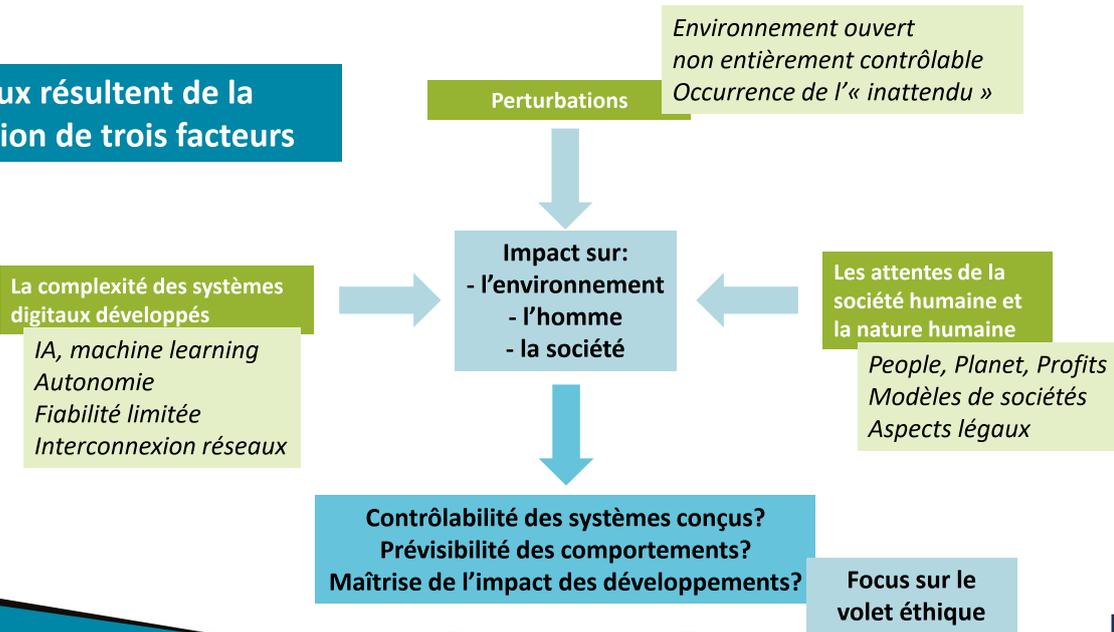
Chacune de nos contributions devrait subir une étude de son impact sur la société!

Damien Trentesaux

6

Contextualisation des enjeux « éthiques »

Ces enjeux résultent de la conjonction de trois facteurs



Damien Trentesaux

7

Trentesaux et Karnouskos, 2021



Partie II

De l'urgence de considérer l'éthique en recherche dans le domaine de la transformation digitale des entreprises

8

Quelques risques d'ordre éthiques

- Des grands classiques, la partie émergée de l'iceberg
 - les biais en IA: learning sur des données non représentatives
 - les données: usage, vente de données personnelles
 - Paradoxes et situations « extraordinaires » (MoralMachine du MIT)
- Et pour nous, les scientifiques
 - Détournement intentionnel de travaux scientifiques
 - Micro-travail (captchas), ubérisation du travail
 - Comportement non prouvé des systèmes autonomes
 - « Trop d'aide, tue l'aide! »
 - Béquille émotionnelle, disparition de compétences humaines, de métiers

« L'enfer est pavé de bonnes intentions »

Tout chercheur (surtout en GI) ne pense qu'à bien faire pour la communauté
 Pouvez vous garantir que vos développements sont éthiques? Vraiment?

Un collègue qui travaille sur l'IA en biométrie pense arrêter (« big brother »)

Il est facile d'imaginer ces risques éthiques en GI

La R&D scientifique et industrielle fourmille de bonnes idées... a priori...



Ordonnancement tenant compte de la fatigue des opérateurs

Paredes-Astudillo et al., 2020

Systèmes cobotiques

Traçabilité des produits et processus (RFID guns)

Operateur 4.0, symbioses, Transhumanisme...

Exosquelettes

Systèmes biométriques



Il est facile d'imaginer ces risques éthiques en GI

La R&D scientifique et industrielle fourmille de bonnes idées... a priori...



Ordonnancement tenant compte de la fatigue des opérateurs
Paredes-Astudillo et al., 2020

Une bonne intention:

La détection d'un courant galvanique permet d'estimer la fatigue des opérateurs 4.0: Sudation, température corporelle... Pourquoi ne pas l'utiliser pour le soulager de certaines tâches? (load balancing)

Qui conduit à des questions éthiques:

- Et si l'opérateur est réticent ?
- Et si ses statistiques sont utilisées pour évaluer ses performances, le surveiller ?
- Et si ce système est détourné pour l'inciter à travailler plus après un état de fatigue ?
- Et si une IA ordonnance les tâches en les affectant en priorité aux opérateurs les moins souvent fatigués pour les fatiguer autant que les autres ?
- Et si un état totalitaire le généralisait à tous les travailleurs ?

L'éthique: un nid de paradoxes et de dilemmes!



Freins à l'étude des aspects éthiques en GI

Réaction	Contre argument
« Cela ne me concerne pas! »	Projets Européens: Roboethics, ETHICBOTS, RoboLaw Montée en force de ces thématiques aux USA en science & technique (MIT, IEEE, GAFA, Navy)... <small>7th meeting of sub-experts group 14: Digital, Industry and Space, mars 21, HORIZON-CL4-2021-2022_03_16_2021.docx: "[...] six expected impacts in the Strategic Plan [...]: A human-centred and ethical development of digital and industrial technologies [...]" Call - A HUMAN-CENTRED AND ETHICAL DEVELOPMENT OF DIGITAL AND INDUSTRIAL TECHNOLOGIES 2021</small>
« Bah, on signe une charte et avec le RGPD c'est bon! »	Le comportement éthique des acteurs humains impliqués ne suffit pas pour garantir le comportement éthique des systèmes créés, surtout sur le long terme ou en dehors de son laboratoire. Impact sociétal, détournements... rarement étudiés.
« C'est facile! On met un gros bouton rouge »	Comment être sûr qu'il fallait bien le déclencher et que la situation ne s'empire pas? L'homme est lent... très lent... et appuyer sur un BRB aura forcément un coût: quid d'un système de transport en pleine vitesse ? D'un avion en vol? d'un robot en pleine opération sur un malade? d'une IA en pleine opération financière (μs) ? Quid d'une flotte d'une centaine d'AGV?
« Pas de soucis, l'homme restera au contrôle! »	L'humain superviseur doit réagir idéalement (temps, décision). Temps de réaction ? Connaissance du contexte ? Problèmes/délais de communication ? Gestion d'une flotte complète ? C'est donc un magicien omniscient ! Il est aussi source de catastrophe (volontaire ou non) !
« si un système est sûr, alors il est éthique! »	La sûreté est une notion fonctionnelle « interne », on ne peut prévoir tous les événements. Une réaction sûre n'est donc peut être pas éthique! (cf. les dilemmes).

Trentesaux D., Karnouskos S. (2020).

Avouez que votre réaction est dans ce tableau 😊





Partie III

OK, on ne peut s'affranchir de réfléchir à l'impact éthique de et dans nos recherches.
Mais au fait, c'est quoi l'éthique?

13

Ethique : définition(sss)



Définir l'éthique pour un être humain est déjà sujet à discussion depuis des siècles avec de nombreux paradoxes à la clé!

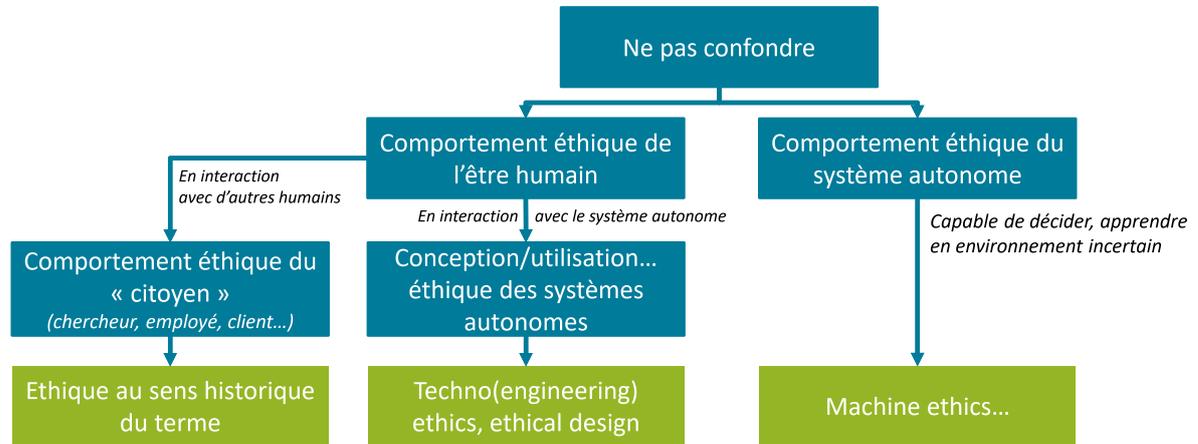
Socrate, Platon, Aristote, Descartes, Kant...

Un comportement en accord avec les attentes culturelles d'une société en relation avec la moralité et l'équité
(Morahan, 2015)

Une vie bonne, avec et pour autrui, dans des institutions justes
(Ricoeur, 1990)

Typologies

Typologie #1: a quoi/qui s'applique l'éthique (et en génie industriel) ?

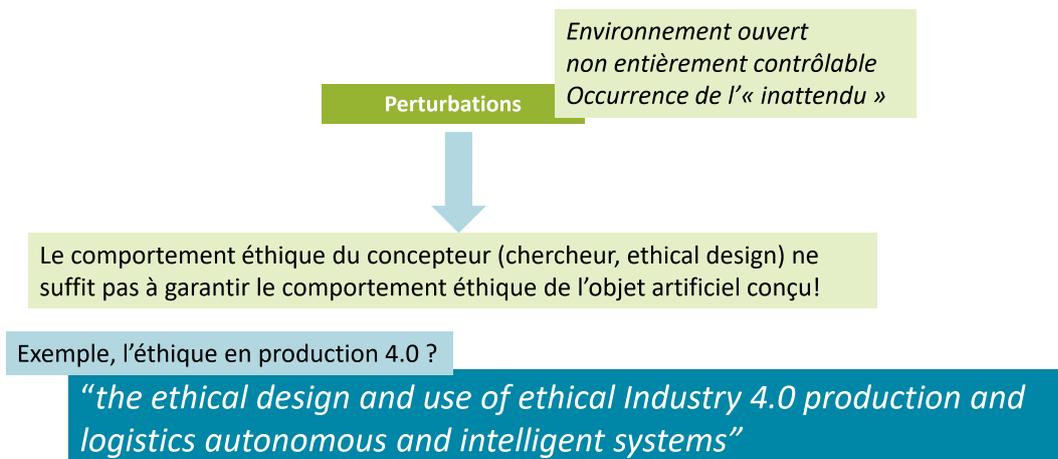


Adapté de Trentesaux et Karnouskos, 2021

Pourquoi aussi l'éthique du système autonome ?

Typologies

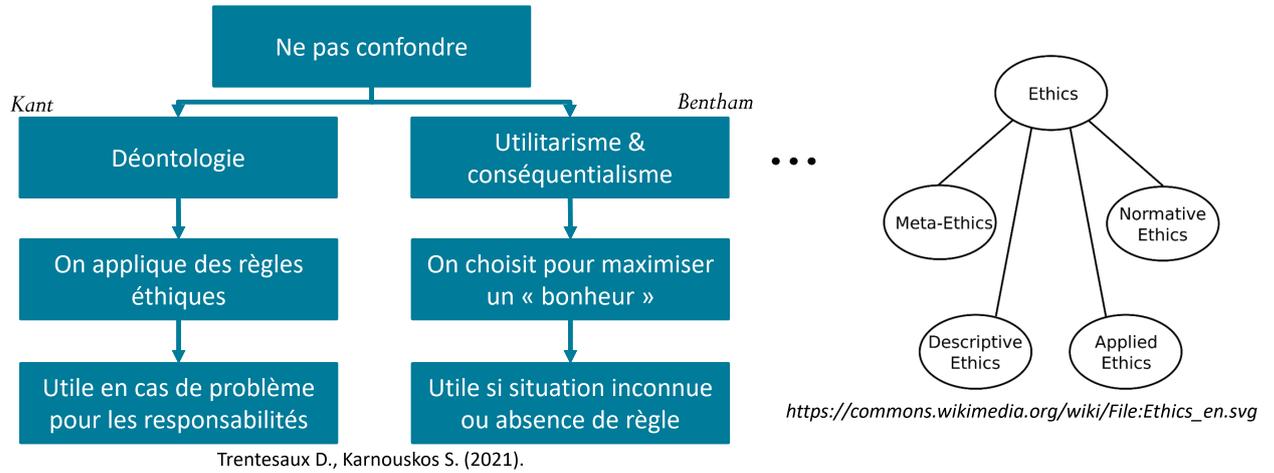
Typologie #1: a quoi/qui s'applique l'éthique (et en génie industriel) ?



Trentesaux et Caillaud, 2020

Typologies

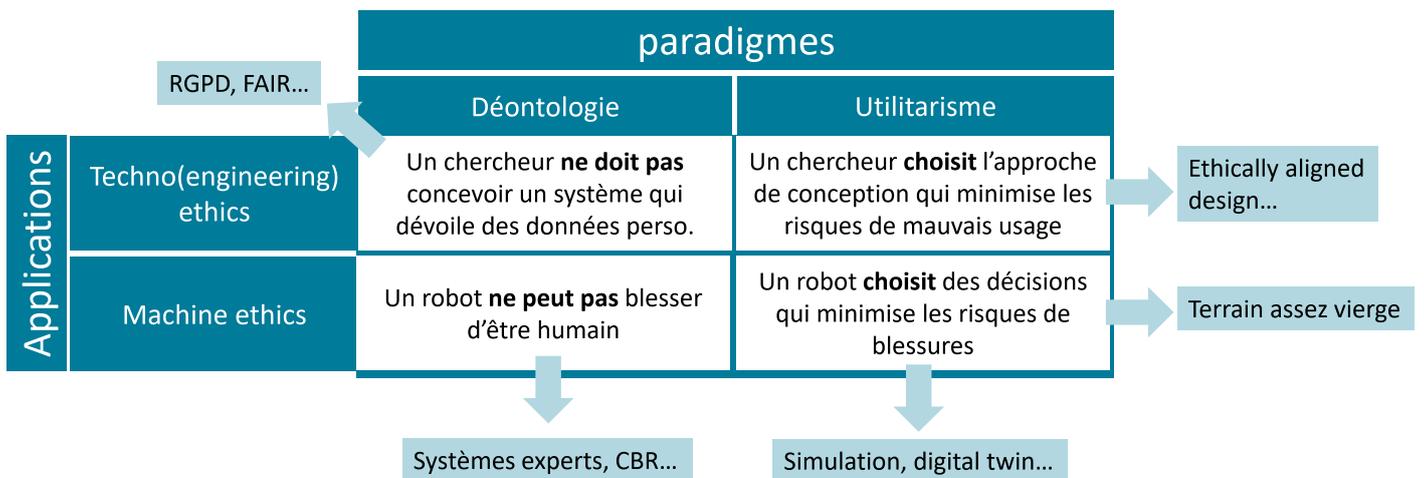
Typologie #2: quel type d'éthique ? Les paradigmes/doctrines « opérationnalisables »



Remarque: les règles déontologiques évoluent avec le temps, qu'est ce qui les fait évoluer?

Typologies

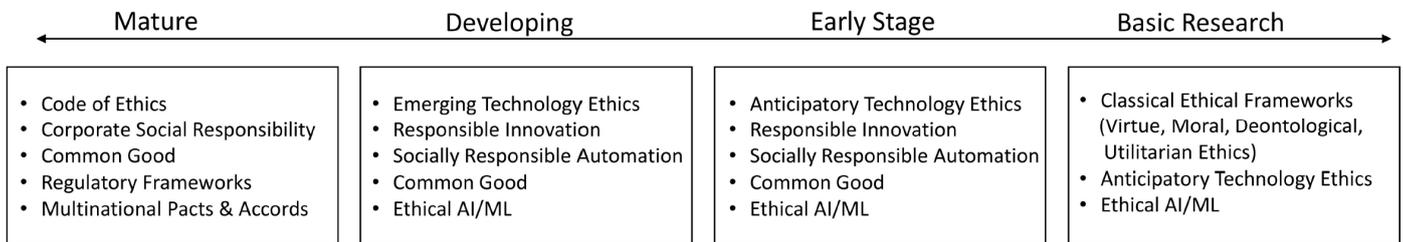
Croisement de ces deux typologies: exemples et positionnement



Trentesaux et Karnouskos, 2021

Typologies

Typologie #3: niveaux de maturité des développements en éthique



Khargonekar et Sampath, IFAC WG, 2020

Le déploiement d'une démarche éthique se fait en fonction du TRL (degré de maturité) des acteurs impliqués. Nous n'en sommes qu'au début!



Typologies

Typologie #4: niveaux de granularité auxquels s'applique l'éthique



Khargonekar et Sampath, IFAC WG, 2020

Importance de ce type de typologie : problème d'alignement de l'éthique par niveau (poupées russes): L'éthique d'une personne vis-à-vis de sa hiérarchie serait-elle la même vis-à-vis de la société?



Autour de l'éthique en génie industriel

Concept/principe	Finalité	Positionnement par rapport à l'éthique
RGPD: Règlement Général sur la Protection des Données GDPR: General Data Protection Regulation	Le RGPD encadre le traitement des données personnelles sur le territoire de l'Union européenne	<ul style="list-style-type: none"> • Ne couvre que la partie « Techno(engineering) ethics » déontologique. • Ne relève que des données, pas des comportements (on peut respecter le RGPD mais ne pas être éthique.) • Ne concerne que l'Europe, et donc, ailleurs ??
RSE (Norme ISO 26000): Responsabilité Sociétale des Entreprises CSR: Corporate social responsibility	Responsabilité d'une organisation vis-à-vis des impacts de ses décisions et de ses activités sur la société et sur l'environnement, se traduisant par un comportement transparent et <u>éthique</u> qui : - contribue au développement durable y compris à la santé et au bien-être de la société - prend en compte les attentes des parties prenantes - respecte les lois en vigueur et est compatible avec les normes internationales - est intégré dans l'ensemble de l'organisation et mis en œuvre dans ses relations	<ul style="list-style-type: none"> • Relève principalement de l'éthique financière et marketing: main d'œuvre à bas coût, corruption, respect des engagements financiers, des clients, etc. • Concerne l'éthique des employés vs. leur entreprise. • Ne concerne donc pas le bien-être, l'équité, le développement personnel, les systèmes autonomes, etc. ou les relations entre l'employé et la société dans sa globalité.
Développement durable: DD Sustainable development: SD	« un développement qui répond aux besoins du présent sans compromettre la capacité des générations futures de répondre aux leurs » (rapport Brundtland)	Le pilier « humain » intègre le volet éthique, mais préfère mettre l'accent sur la trajectoire temporelle (présent vs. Futur), alors que l'éthique est essentiellement du ressort du temps présent ou à très court terme

Des éléments nécessaires, mais non suffisants!

Merci à Lamia, Vincent, Frédéric, Claude, Chakib ☺...

Damien Trentesaux

21



Partie IV

Ethique et transformation digitale: un monde d'opportunités
 scientifiques à fort impact sociétal
 (que chercher de plus? ☺)

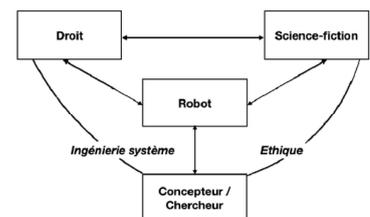
Domaines scientifiques proches et contributeurs

- Corpus scientifique naissant (tout le contraire du corpus philosophique 😊)
 - Rendre l'éthique « opérationnelle »
- Quelques grands domaines de recherche pour aborder l'éthique en GI (multifacette):
 - Sciences de l'ingénieur
 - Sûreté de fonctionnement et ingénierie système
 - Automatique, coopération homme-machine
 - Robotique et systèmes embarqués
 - Multi-agents et intelligence artificielle
 - Sciences humaines et de gestion
 - Psychologie, sociologie
 - Ergonomie cognitive
 - Sciences de gestion
 - Ne pas oublier: philosophie, littérature...

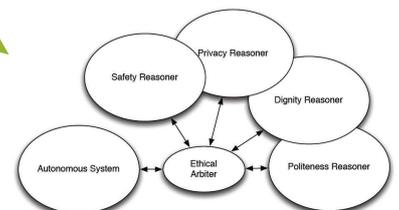
Trois exemples de contribution en éthique « pour l'ingénieur »

Human Value	Technology Shall Be Designed to ...
Self-actualization	... enable the workers to accomplish their own goals and obtain personal success, to demonstrate their own capabilities and competence in order to advance their career.
Accountability	... ensure full transparency of how the symbiotic CPPS operates and makes decision. It includes that the actions of a worker or group of workers is shared transparently to ensure that any problem can be traced back to the cause.
Trustworthiness	... be honest, trustworthy, reliable, and avoid systematic bias, thus performing its required function under given conditions for a stated time interval, as well as to improve interpersonal trust in a human-machine hybrid system.
Privacy	... ensure informational privacy, respect for the workers' private sphere and the right to determine what information can be communicated to others through informed consent.
Welfare	... ensure workers' health (physical well-being and peaceful psychological state) also thanks to a proper work-life balance, a balanced workload and a comfortable and pleasant work environment.
Autonomy	... allow workers to be independent, free and flexible while at work in order to express themselves, engage in autonomous thinking, make independent choices, utilize their creativity and fully use their own intelligence.
Altruism	... enhance teamwork and mutual care among the workers and to favour the wellbeing of all those one comes into contact with in their professional activities, thus demonstrating oneself as sincere, open to others and responsible.

Longo, F., Padovano, A., Umbrello, S., 2020.



Caillaud, E., 2020.



Dennis, L., Fisher, M., 2018.

Applications	paradigmes	
	Déontologie	Utilitarisme
Techno(engineering, ethics)	Un chercheur ne doit pas concevoir un système qui dévoile des données perso.	Un chercheur choisit l'approche de conception qui minimise les risques de mauvais usage
Machine ethics	Un robot ne peut pas blesser d'être humain	Un robot choisit des décisions qui minimise les risques de blessures

Quelques problématiques de recherche en GI

1. L'éthique est-elle un indicateur de performances comme les autres ?
2. Planification/ordonnancement/pilotage éthique des systèmes de production du futur ?
3. Les horizons informationnels et temporels en GI modifient-ils les décisions éthiques ?
4. Comment assurer l'éthique des décisions prises en environnement incertain ?
5. Comment aligner la stratégie d'entreprise avec l'éthique personnelle (« souffrance éthique ») ?
6. Comment embarquer des décisions éthiques dans les systèmes autonomes de production ?
7. Comment maintenir la compétence humaine face à un système d'aide trop efficace ?
8. Comment s'assurer que nos recherches seront toujours utilisées de manière éthique ?
9. Est-il possible de définir une évaluation éthique des systèmes et produits industriels ?
10. Comment s'assurer du bon impact sociétal des systèmes conçus dans le contexte de l'industrie 4.0 ?
11. Est-il possible de développer une méthode de conception générique (IS) garantissant l'éthique des parties prenantes ?
12. Comment s'assurer qu'une méthode « lean » est éthique ?
13. Quelles règles éthiques imposer aux systèmes industriels 4.0 ?
14. Les paradigmes éthiques sont-ils équivalents ? Quand utiliser l'un plutôt qu'un autre ?
15. Comment articuler RSE, qualité de vie... et éthique en entreprise ?

Damien Trentesaux

25

Impact sur nos formations

- Ecoles ingénieurs en GI
 - Former au RGPD, à la RSE, au développement durable, etc. c'est bien!
 - Ca ne suffit pas!
 - Quelques idées
 - Lean management: développer le volet éthique, par exemple: le lean est il la panacée?
 - Evolution des modèles de notre société: vers quoi voulons-nous aller ?
 - par exemple: robotisation et emploi
 - Cours de philosophie (transhumanisme, doctrines de l'éthique), littérature (science-fiction)



Evaluating ethical risks

- Etudes doctorales en GI
 - Former à l'éthique en recherche: auteurs, RGPD, plagiat, etc. c'est bien!
 - Ca ne suffit pas non plus!
 - Quelques idées
 - Pousser leurs développements aux limites en terme d'usage (systématiser l'analyse des risques éthiques)
 - Détournements, usages non imaginés, conséquence sur la société, etc. sans s'interdire d'aller voir!
 - Se questionner sur l'impact pour toutes les parties prenantes ('stakeholders') ou concernées par le travail
 - Ce qui est développé pour l'un, est-il bon pour l'autre?

Damien Trentesaux

26

Conclusion

Alors, l'éthique en GI :
un sujet de recherche ou pas ?
Juste une conviction personnelle ?

Peut-être une occasion pour contribuer au renouveau de notre domaine ?
Production, achats, logistique, cobotique, systèmes de soins...
Ne pas avoir peur de se lancer sur ce terrain vierge, tellement vaste!

Possibilité de monter un numéro spécial « éthique et GI », si vous êtes partant, contactez moi!

Peut-être une auto-critique de notre sphère scientifique:

Le rôle du chercheur en GI n'est-il pas d'éclairer la société sur les futurs industriels possibles?

Damien Trentesaux

27

Références

- Berrah L., Clivillé V., Trentesaux D., Chapel C. (2021). Performance industrielle : une évolution intégrant l'éthique dans le contexte de l'Industrie 4.0, CIGI-Qualita, Grenoble
- Berrah L., Trentesaux D. (2021). Decision-Making in Future Industrial Systems: Is Ethics a New Performance Indicator? Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future, 952, Studies in computational intelligence, Springer.
- Berrod F., Clermont P., Trentesaux D. (2020). Droit et robots: Droit science-fictionnel et fictions du droit, PUV, ISBN 978-2-36424-070-4
- Paredes-Astudillo Y., Moreno D., Vargas A.-M., Angel M.-A., Perez S., Jimenez gordillo J.-F., Saavedra-Robinson L., Trentesaux D. (2020). Human Fatigue Aware Cyber-Physical Production System. 1st IEEE International Conference on Human-Machine Systems, IEEE, Rome, september.
- Rault R., Trentesaux D. (2018). Artificial Intelligence, Autonomous Systems and Robotics: Legal Innovations. Service Orientation in Holonic and Multi-Agent Manufacturing, 762, Studies in computational intelligence, Springer.
- Trentesaux D., Karnouskos S. (2021). Engineering Ethical Behaviors in Autonomous Industrial Cyber-Physical Human Systems. Cognition, Technology & Work.
- Trentesaux D., Rault R. (2020). Ethique des robots intelligents dans la société humaine : Regards croisés issus du droit, de la science et de la littérature. In Berrod F., Clermont P., Trentesaux D. (Ed.), Droit et robots: Droit science-fictionnel et fictions du droit, PUV, pp. 89-128, ISBN 978-2-36424-070-4
- Trentesaux D., Rault R., Caillaud E., Huftier A. (2021). Ethics of autonomous intelligent systems in the human society: cross views from science, law and science-fiction. Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future, 952, Studies in computational intelligence, Springer.
- Trentesaux D., Caillaud E. (2020). Ethical stakes of Industry 4.0. 21st IFAC World Congress, Elsevier, Berlin.
- Trentesaux D., Karnouskos S. (2020). Ethical Behaviour Aspects of Autonomous Intelligent Cyber-Physical Systems. Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future, 853, Studies in computational intelligence, Springer.
- Trentesaux D., Rault R. (2017). Ethical behaviour of autonomous non-military cyberphysical systems. XIX International Conference on Complex systems: control and modeling problems, ofort, Samara, Russia, ISBN 978-5-473-01154-8
- Trentesaux D., Rault R. (2017). Designing Ethical Cyber-Physical Industrial Systems. 20th IFAC World Congress, 50(1), Elsevier, Toulouse.
- Trentesaux D. (2021). Ensuring ethics of cyber-physical and human systems: a guideline, 1st Sohoma Latin America workshop, Colombie.

Damien Trentesaux

28

Références

- Alsejgier, R.A., 2016. Roboethics: Sharing Our World with Humanlike Robots. *IEEE Potentials* 35, 24–28. <https://doi.org/10.1109/MPOT.2014.2364491>
- Anderson, S.L., 2008. Asimov's "Three Laws of Robotics" and Machine Metaethics. *AI Soc.* 22, 477–493. <https://doi.org/10.1007/s00146-007-0094-5>
- Caillaud, E., 2020. Droit des robots et science-fiction : quels enjeux pour l'ingénierie système ? In *Droit des robots: Droit science-fictionnel et fictions du droit*, PUV
- Coeckelbergh, M., 2012. Can we trust robots? *Ethics Inf Technol* 14, 53–60. <https://doi.org/10.1007/s10676-011-9279-1>
- Delvaux, M., 2016. Civil law rules on robotics, European Parliament Legislative initiative procedure 2015/2103.
- Dennis, L., Fisher, M., Slavkovik, M., Webster, M., 2016. Formal verification of ethical choices in autonomous systems. *Robotics and Autonomous Systems* 77, 1–14. <https://doi.org/10.1016/j.robot.2015.11.012>
- Fisher, M., List, C., Slavkovik, M., Winfield, A., 2016. Engineering moral machines. *Informatik-Spektrum*.
- Gurkaynak, G., Yilmaz, I., Haksever, G., 2016. Stifling artificial intelligence: Human perils. *Computer Law & Security Review* 32, 749–758. <https://doi.org/10.1016/j.clsr.2016.05.003>
- Hengstler, M., Enkel, E., Duelli, S., 2016. Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change* 105, 105–120. <https://doi.org/10.1016/j.techfore.2015.12.014>
- Karnouskos, S., 2018. Self-Driving Car Acceptance and the Role of Ethics. *IEEE Transactions on Engineering Management* 1–14. <https://doi.org/10.1109/TEM.2018.2877307>
- Khargonekar et Sampath, "A Framework for Ethics in Cyber-Physical-Human Systems", IFAC WG, 2020
- Longo, F., Padovano, A., Umbrello, S., 2020. Value-Oriented and Ethical Technology Engineering in Industry 5.0: A Human-Centric Perspective for the Design of the Factory of the Future. *Applied Sciences* 10, 4182.
- Palmerini, E., Bertolini, A., Battaglia, F., Koops, B.-J., Carnevale, A., Salvini, P., 2016. RoboLaw: Towards a European framework for robotics regulation. *Robotics and Autonomous Systems*. <https://doi.org/10.1016/j.robot.2016.08.026>
- Thekkilakattil, A., Dodig-Crnkovic, G., 2015. Ethics Aspects of Embedded and Cyber-Physical Systems, in: *Computer Software and Applications Conference (COMPSAC), 2015 IEEE 39th Annual*. pp. 39–44. <https://doi.org/10.1109/COMPSAC.2015.41>
- van Gorp, A., 2007. Ethical issues in engineering design processes; regulative frameworks for safety and sustainability. *Design Studies* 28, 117–131. <https://doi.org/10.1016/j.destud.2006.11.002>

Session 1.1 : Design et Pilotage des systèmes de production de biens et de services

Chairman : Christophe Danjou - Vincent Cliville

CIGI 2021

Métaheuristique pour le flow shop hybride avec machines dédiées

ASMA OULED BEDHIEF

Ecole Supérieure Privée d'ingénierie et de Technologie-Esprit, Tunis, Tunisie
OASIS - Ecole Nationale d'Ingénieurs de Tunis, Université Tunis el Manar B.P. 37, Le belvédère 1002, Tunis.

asma.ouledbedhief@esprit.tn

Résumé – Dans cet article, nous traitons un problème d'ordonnement de type flow shop hybride à trois étages avec machines dédiées. Nous cherchons un ordonnancement qui minimise la date de fin totale des jobs (makespan). Ce problème étant NP-difficile, nous proposons un algorithme génétique (GA) pour sa résolution. D'abord, un plan factoriel complet a permis d'optimiser le choix des paramètres de trois versions (variantes) de l'algorithme génétique proposé. Ces variantes diffèrent par l'opérateur de croisement utilisé. Ensuite, une étude comparative de ces trois variantes est effectuée afin d'apercevoir quelle variante de l'algorithme génétique se prête mieux à la résolution du problème. Les résultats de comparaison menée sur plusieurs classes d'instances montrent que notre algorithme génétique (GA) fonctionne relativement mieux lorsque l'opérateur de croisement LOX (Linear Order Crossover) est utilisé. Ils montrent également que l'opérateur de croisement que nous proposons, dans le cadre de ce travail, NOX (New Order Crossover) est bien meilleur que OX (Order Crossover) dans 94% des cas testés.

Abstract – In this paper, we deal with the three-stage hybrid flow shop with dedicated machines. The objective is to minimize the maximum completion time (makespan). To solve this NP-hard problem, we propose a genetic algorithm (GA). A preliminary computational test with factorial design is firstly conducted to tune the key parameters of three versions of the proposed genetic algorithm. These versions are resulted from three crossover operators. Then, the performance of the proposed genetic algorithms is evaluated.

The results with different problem configurations demonstrate that the genetic algorithm GA performs relatively better when the LOX (Linear Order Crossover) operator is used. The results also show that the crossover operator NOX (New Order Crossover) that we propose, in this study, performs better than OX (Order Crossover) for 94% of problem cases.

Mots clés – Flow shop hybride, machines dédiées, algorithme génétique, opérateur de croisement, plan factoriel complet

Keywords – Hybrid flow shop, dedicated machines, genetic algorithm, crossover operator, factorial design...

1. INTRODUCTION

Nous considérons un problème d'ordonnement de type flow shop hybride avec machines dédiées, composé de trois étages et fabriquant deux types de jobs. Le premier et le deuxième étage contiennent chacun une seule machine respectivement M_1 et M_2 et, deux machines dédiées D_1 et D_2 sont au troisième étage (voir figure 1).

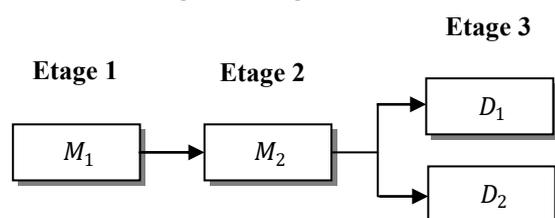


Figure 1. Flow shop hybride à trois étages avec deux machines dédiées

Chaque job nécessite exactement trois opérations. D'abord, il doit passer par la machine M_1 suivie de M_2 afin d'exécuter les deux premières opérations de sa gamme de fabrication. Ensuite, il suit un traitement particulier en fonction de sa destination finale. Il sera exécuté sur la machine dédiée correspondante à son type. L'objectif est la minimisation de la date de fin totale (makespan).

Selon la notation de Graham et al. (1979), nous notons ce problème $3FHD|1, 1, 2|C_{max}$. Il s'agit d'un problème de flow shop hybride à 3 étages avec une seule machine au premier et au deuxième étage et deux machines dédiées au troisième étage.

Les problèmes d'ordonnement avec machines dédiées présentent un grand intérêt pratique puisqu'ils ont d'importantes applications dans plusieurs industries telles que l'industrie pharmaceutique, l'industrie de meuble, d'étiquettes autocollantes et de verre... Prenons à titre d'exemple le cas d'une industrie pharmaceutique où un

même médicament dispose de différentes présentations (gélules, sirop ou poudre en sachets). En effet, les premières transformations sont identiques, par la suite, le médicament suit un cheminement spécifique à sa présentation, et il sera affecté à la ligne correspondante à la forme demandée.

Le reste de l'article est organisé comme suit :

Dans la section 2, nous développons une brève revue de littérature traitant les problèmes d'ateliers avec machines dédiées. La section 3 présente l'algorithme génétique (GA) que nous proposons. Dans la section 4, nous présentons l'étude expérimentale que nous effectuons afin d'évaluer la performance de l'algorithme génétique : Nous présentons d'abord le paramétrage des trois variantes de GA, puis nous exposons les résultats de l'étude comparative de ces variantes d'algorithme génétique. Nous terminons ce travail par une conclusion et quelques perspectives.

2. ETAT DE L'ART

Le problème d'ordonnement avec machines dédiées est introduit pour la première fois par Hermann et Lee (1992), en étudiant le cas à deux étages. Il s'agit d'une seule machine au premier étage et deux machines dédiées au deuxième. Ce problème est NP-difficile malgré son apparente simplicité [Hermann et Lee (1992), Lin (1999)].

Dans la suite, nous exposons les principaux travaux ayant étudié le problème du flow shop hybride avec machines dédiées en proposant des méthodes exactes et approchées.

Besbes et al. (2006) considèrent un atelier de type flow shop hybride avec machines dédiées à deux étages, composés respectivement de s et m machines dédiées. Les auteurs cherchent à minimiser le makespan et proposent un programme linéaire en variables mixtes. Mosheiov et Sarig (2010) étudient le cas où une seule machine est au premier étage et m machines dédiées sont au deuxième. Ils supposent que les jobs doivent être exécutés en blocs et partagent la même date d'échéance. Afin de minimiser le nombre pondéré des jobs en retard, les auteurs proposent un programme linéaire en nombres entiers. Des procédures par séparation et évaluation (PSE) ont été aussi proposées. Une PSE est développée par Hadda et al. (2014) pour le problème $Fm+1|T = m|Cmax$ où T désigne le nombre de types de jobs à réaliser. L'atelier se compose de deux étages avec m machines dédiées au deuxième étage, et une seule machine au premier. Cette procédure utilise une règle d'élimination développée par les mêmes auteurs dans [Hadda et al. (2012)]. Le même problème a été abordé dans [Liu et al. (2013)] où une borne inférieure, des règles d'élimination et une PSE ont été également présentées. Riane et al. (2002) développent un programme dynamique pour le cas à deux étages et deux machines dédiées. Là aussi, les auteurs cherchent à minimiser le makespan. L'analyse des résultats obtenus montrent que ce programme n'est pas adapté pour un nombre de jobs $n \geq 15$.

D'autres travaux ont eu recours à la résolution approchée du problème.

Oguz et al. (1997) traitent le problème à deux étages où deux machines dédiées sont au premier étage. Pour résoudre le problème, les auteurs développent une heuristique et étudient son comportement au pire des cas. L'étude expérimentale montre que l'heuristique a une performance garantie de $3/2$. Yang (2015) traite le même problème en se

donnant la somme des dates de fin comme critère d'optimisation. Il propose deux heuristiques de résolution. Yang (2013) considère le cas où m machines dédiées au premier étage, et une seule machine au deuxième. Il définit quatre heuristiques de résolution basées sur la règle de Johnson. Pour chaque heuristique, il fournit une analyse au pire des cas. Dans [Dridi et al. (2009)], les auteurs introduisent des règles d'élimination, des bornes inférieures et des cas polynomiaux pour le $Fm+1|T = m|Cmax$. Ils développent aussi une heuristique de complexité $O(n^2)$. Wang et Liu (2013) étudient le même problème. Ils proposent une heuristique de résolution basée sur une procédure de séparation et évaluation, et trois règles de dominance. L'étude expérimentale prouve que des problèmes dont les tailles vont jusqu'à $n = 100$ jobs avec 5 machines dédiées, sont résolus de façon optimale ou proche de l'optimum en un temps raisonnable. Les auteurs effectuent une comparaison avec deux métaheuristiques et prouvent la performance de l'heuristique proposée. Besbes et al. (2010) considèrent s et m machines dédiées respectivement dans les deux étages. Ils développent deux approches. La première est une heuristique qui se base sur la règle de Johnson, tandis que la deuxième est un algorithme génétique.

Très rares sont les travaux ayant considéré plus de deux étages. Riane et al. (1998) considèrent un cas d'atelier à trois étages. Le premier et le troisième étage contiennent une seule machine, et deux machines dédiées sont au deuxième étage. Les auteurs développent deux heuristiques. L'une se base sur la programmation dynamique et l'autre se base sur une PSE. Ouled Bedhief et al. (2019) considèrent le cas où deux machines dédiées sont au troisième étage. Les auteurs étudient un ensemble de cas particuliers et proposent une heuristique qui se base sur la programmation dynamique et la recherche locale.

L'examen de la littérature montre qu'une majeure partie des travaux réalisés sur le flow shop hybride avec machines dédiées portent sur des configurations à deux étages. Cette quasi-absence des travaux considérant trois étages représente à notre avis un pas à franchir pour se rapprocher de la réalité industrielle.

En plus, la plupart des travaux proposent des méthodes heuristiques pour la résolution du problème. Quant aux métaheuristiques, elles sont plutôt rares dans la littérature sur le problème d'ordonnement avec machines dédiées.

Dans ce papier, nous proposons un algorithme génétique. Nous étudions trois variantes de cet algorithme génétique qui diffèrent par l'opérateur de croisement utilisé.

La comparaison entre ces variantes d'algorithme génétique est l'objectif principal de ce papier. Nous cherchons ainsi à apercevoir quelle version de l'algorithme génétique se prête mieux à la résolution du problème étudié.

3. ALGORITHME GENETIQUE (GA)

Les algorithmes génétiques utilisent des opérateurs de croisement et de mutation, une méthode de sélection. Ils utilisent, aussi, des paramètres tels que la probabilité de croisement, la probabilité de mutation, la taille de la population et le nombre de générations. Le choix des paramètres et d'opérateurs les plus adéquats dépend du problème traité, et conditionne l'efficacité de l'algorithme génétique.

Dans ce qui suit, nous mettons l'accent sur les particularités

de notre algorithme génétique et nous exposons les choix des paramètres et d'opérateurs considérés dans ce travail.

3.1. Codage

Une solution du problème est une séquence des jobs à ordonnancer. Nous trouvons que la façon la plus simple de coder une solution est de la représenter par un chromosome dans lequel chaque gène correspond à un job, et où chaque job est défini par un numéro.

3.2. Génération de la population initiale

Dans cette étape, une population initiale d'individus doit être générée soit aléatoirement, soit par des heuristiques, soit à la fois des solutions aléatoires et heuristiques. Dans ce travail, nous choisissons de générer une population initiale de façon aléatoire afin de permettre un degré important d'hétérogénéité. Nous proposons, aussi, d'étudier deux valeurs pour la taille de la population: $P_{size} = 50$ et $P_{size} = 100$.

3.3. Fonction Fitness

Notre objectif étant la minimisation du makespan, nous choisissons une fonction fitness qui, pour chaque individu, représente la valeur du makespan.

3.4. Sélection

Cette étape consiste à identifier les individus qui permettent de produire la population de la génération suivante. Dans ce travail, nous utilisons la méthode de sélection par roulette. Elle associe chaque individu de la population à un secteur de la roulette proportionnel à sa performance. L'individu sélectionné correspond au secteur sur lequel la roulette s'est arrêtée.

Dans cette étape de l'algorithme génétique, nous sélectionnons $\frac{P_{size}}{2}$ individus de la population courante. Ces individus forment une population nommée Pop_{select} . Seuls les individus appartenant à la population Pop_{select} sont susceptibles d'être croisés dans la phase de croisement.

3.5. Croisement

Cette phase permet de générer de nouveaux individus (enfants) en se basant sur l'information génétique des individus sélectionnés (parents). Elle s'applique avec une probabilité P_{cross} appelée probabilité de croisement.

Dans ce travail, nous utilisons deux manières de croisement définies par les opérateurs : LOX (Linear Order Crossover) et OX (Order Crossover). En effet, ils sont les plus utilisés pour les problèmes d'ordonnancement. Nous développons, également, un nouvel opérateur de croisement en combinant les opérateurs LOX et OX. Nous nommons cet opérateur NOX (New Order Crossover).

Order Crossover (OX): Deux points de coupures sont choisis aléatoirement pour dissocier chacun des deux parents en trois fragments. La sous-séquence située entre les deux points de coupures du Parent1 (resp. Parent2) est transférée à Enfant1 (resp. Enfant2). En partant du deuxième point de croisement, les gènes manquants sont copiés du Parent2 (resp. Parent1) tout en maintenant le même ordre d'apparition.

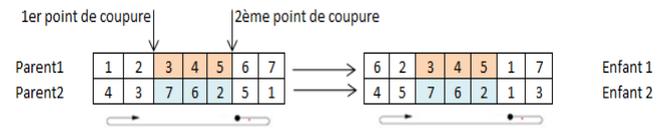


Figure 2. Fonctionnement de l'opérateur OX

Linear Order Crossover (LOX): Il a le même principe que l'opérateur OX: les deux opérateurs utilisent deux points de coupures et gardent inchangée la partie située entre ces deux points de coupures. Toutefois, le mode de la transmission des autres gènes est différent, dans la mesure où les gènes manquants de Enfant1 (resp. Enfant2) sont hérités du Parent2 (resp. Parent1) dans le même ordre d'apparition mais, en commençant par la première position de gauche à droite.

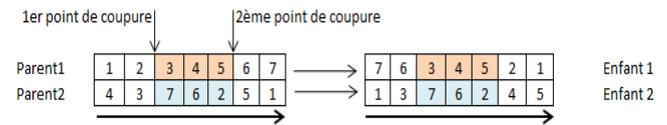


Figure 3. Fonctionnement de l'opérateur LOX

New Order Crossover (NOX): Les opérateurs OX et LOX restent des opérateurs qui donnent des résultats très satisfaisants sur un large spectre de problèmes de natures différentes. Afin d'en tirer avantage, nous proposons de combiner OX et LOX, en un nouvel opérateur NOX.

Pour l'opérateur LOX, nous partons toujours du début d'un chromosome et les gènes sont copiés dans leur ordre d'apparition dans l'autre parent (balayage des gènes linéaire). Toutefois, pour l'opérateur OX, nous partons toujours du deuxième point de coupure d'un chromosome et les gènes sont copiés dans leur ordre d'apparition dans l'autre parent (balayage des gènes circulaire). La différence entre ces deux opérateurs de croisement réside donc, dans la manière de balayage des gènes.

Nous pensons alors à utiliser les deux manières de balayage des gènes (balayage circulaire et linéaire) dans l'opérateur NOX. Ce dernier fonctionne comme suit:

Deux points de coupures sont choisis aléatoirement pour dissocier chacun des deux parents en trois fragments. La sous-séquence se trouvant entre les deux points du Parent1 (resp. Parent2) est transférée à Enfant1 (resp. Enfant2). Pour compléter les gènes manquants dans Enfant1 (resp. Enfant2), nous partons, tout comme dans OX, de son deuxième point de coupure mais, les gènes sont copiés du Parent2 (resp. Parent1), dans leur ordre d'apparition en balayant de gauche à droite, tout comme dans LOX.

Une fois la fin du chromosome est atteinte, nous continuons à partir de la première position. La figure 4 permet de mieux comprendre le fonctionnement de l'opérateur NOX.

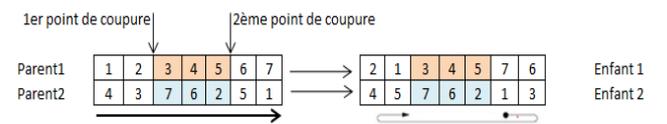


Figure 4. Fonctionnement de l'opérateur NOX

Dans cette phase de l'algorithme génétique, deux parents Parent1 et Parent2 sont choisis parmi la population sélectionnée Pop_{select} . Ensuite, nous générons un nombre

aléatoire $\alpha \in [0,1]$. Si $\alpha \leq P_{cross}$, nous appliquons un opérateur de croisement sur les deux parents. Nous notons Pop_{enfant} l'ensemble des enfants générés à la phase de croisement, qui est de taille $\frac{P_{size}}{2}$.

Par ailleurs, Grefenstette (1986) recommande une forte probabilité de croisement pour obtenir une bonne exploration de l'espace des solutions. Comme elle est typiquement comprise entre 0.6 et 1, nous utilisons, dans ce travail $P_{cross} = 0.7$ et $P_{cross} = 0.9$. Elles sont les plus utilisées pour les ateliers de flow shops hybride dans la littérature.

3.6. Mutation

Elle apporte l'aléa nécessaire à une exploration efficace de l'espace des solutions en permettant de surpasser les minimas locaux. Elle s'applique aux enfants générés selon une probabilité de mutation P_{mut} . Plusieurs opérateurs de mutation ont été présentés dans la littérature. Dans ce travail, nous choisissons de travailler avec l'opérateur de mutation : la mutation-échange. Elle permet de sélectionner, aléatoirement, deux gènes et de les permuter au niveau de l'enfant à muter.

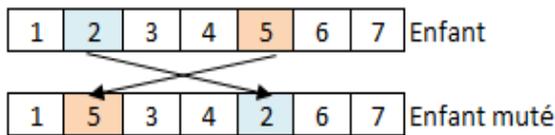


Figure 5. Fonctionnement de la mutation-échange

Dans cette étape de l'algorithme génétique (la mutation), nous générons un nombre aléatoire $\beta \in [0,1]$, pour chaque enfant de la population Pop_{enfant} . Dans le cas où $\beta \leq P_{mut}$, nous appliquons l'opérateur de mutation sur l'enfant à muter.

Comme la probabilité de mutation est typiquement très faible [Goldberg (1989)], nous comptons utiliser, dans ce travail, deux valeurs: $P_{mut} = 0.05$ et $P_{mut} = 0.1$. Ainsi, une forte probabilité de mutation risque de détruire les bons fragments que nous essayons de conserver.

3.7. Insertion

Elle est utilisée afin d'intégrer la population des enfants générés et mutés dans la population courante, tout en respectant que la taille de la population reste constante.

Dans ce travail, nous utilisons une méthode d'insertion qui consiste à éliminer les individus les plus faibles, et à conserver les individus les plus intéressants en permettant leur passage d'une génération à une autre sans les altérer. Cette méthode permet de trier les individus des populations P et Pop_{enfant} dans un ordre croissant de leur valeur de fitness (makespan), et de sélectionner les P_{size} meilleurs individus (ayant les plus faibles valeurs de fitness).

3.8. Critère d'arrêt

Les phases de sélection, croisement et mutation sont itérées jusqu'à ce qu'un critère d'arrêt soit atteint.

Dans ce travail, nous faisons le choix d'arrêter l'exploration de l'espace de recherche après un certain nombre d'itérations, noté $Iter_{max}$. Notre choix du nombre maximum d'itérations s'est porté sur deux valeurs: $Iter_{max} = 200$ et $Iter_{max} = 800$.

En résumé, nous considérons, dans ce travail, différents choix pour chacun des paramètres et opérateurs de notre algorithme génétique. Les valeurs de ces paramètres et opérateurs sont résumés dans le tableau 1.

Tableau 1 : Paramètres et Opérateurs à considérer

Paramètres	Taille de la population: P_{size}	50; 100
	Probabilité de croisement: P_{cross}	0.7; 0.9
	Probabilité de mutation: P_{mut}	0.05; 0.1
	Nombre maximum d'itérations: $Iter_{max}$	200; 800
Opérateurs	Opérateur de sélection	sélection par roulette;
	Opérateur de croisement	LOX; OX; NOX
	Opérateur de mutation	Mutation-échange

Afin d'étudier l'effet de trois opérateurs de croisement et d'un opérateur de mutation, nous allons exécuter trois versions de l'algorithme génétique proposé. Nous appelons ces algorithmes : LOX/Echange, OX/Echange, NOX/Echange. Ils correspondent aux trois paires possibles d'opérateurs croisement/mutation:

Dans ce qui suit, nous décrivons les résultats expérimentaux obtenus dans le but de tester la performance de ces trois versions d'algorithme génétique.

4. RÉSULTATS EXPÉRIMENTAUX

D'abord, il est à signaler que nous avons programmé notre algorithme génétique avec le langage C++, en utilisant l'environnement de développement DEV C++. Les simulations ont été effectuées sur un ordinateur portable équipé d'un processeur Intel (R) Core TM i3 de 2.50 GHz. L'objectif de cette étude expérimentale est de comparer les performances respectives des trois variantes d'algorithme génétique.

4.1. Génération des problèmes tests

Avant de passer à évaluer la performance de chacune des trois variantes d'algorithme génétique, nous décrivons les différentes familles d'instances utilisées.

Cinq familles d'instances ont été générées aléatoirement de F1 à F5. Chaque famille est testée pour plusieurs tailles.

Pour chaque taille, 20 instances sont générées. Les types de jobs sont choisis de façon aléatoire, et quelque soit la famille d'instances, les temps opératoires des jobs suivent une distribution uniforme.

Dans les cinq familles d'instances, nous supposons que les temps opératoires des jobs sur les machines M_1 et M_2 , respectivement p_{i1} et p_{i2} soient générés dans $[1,20]$. Tandis que les temps opératoires des jobs sur les machines dédiées (dp_{ik}) sont générés aléatoirement dans les distributions uniformes suivantes :

- **F1:** $dp_{ik} \sim U[1,20] \forall k \in \{1,2\}$
- **F2:** $dp_{ik} \sim U[1,40] \forall k \in \{1,2\}$
- **F3:** $dp_{ik} \sim U[20,40] \forall k \in \{1,2\}$

- **F4:** $dp_{ik} = p_{i2} + 5 \forall k \in \{1,2\}$
- **F5:** $dp_{ik} = p_{i2} + 10 \forall k \in \{1,2\}$

4.2. Plan Factoriel complet

Chaque variante de l'algorithme génétique est initialisée par une taille de la population (P_{size}), une probabilité de croisement (P_{cross}), une probabilité de mutation (P_{mut}) et par un nombre maximum d'itérations ($Iter_{max}$). Afin de déterminer les meilleurs choix de ces paramètres, nous utilisons un plan factoriel complet. En général, un tel plan consiste à déterminer les facteurs clés qui influencent une réponse donnée, et l'évolution de la réponse en fonction de ces facteurs. Il est le plus utilisé grâce à sa simplicité et à sa mise en œuvre rapide.

Dans ce travail, nous effectuons un plan factoriel complet à deux niveaux, puisque chaque paramètre peut prendre deux valeurs possibles.

Le tableau 2 rappelle les deux niveaux de chacun des facteurs, utilisés dans cette étude. Les facteurs sont les quatre paramètres que nous cherchons à ajuster.

Tableau 2: Les niveaux des paramètres

	Niveau bas (-)	Niveau haut (+)
La taille de population (P_{size})	50	100
La probabilité de croisement (P_{cross})	0.7	0.9
La probabilité de mutation (P_{mut})	0.05	0.1
Le nombre de générations maximum ($Iter_{max}$)	200	800

Un plan factoriel complet à deux niveaux et quatre facteurs s'effectue en 2^4 combinaisons. Nous menons, alors, seize expériences pour chaque variante. Chaque expérience correspond à une combinaison possible des quatre paramètres.

Tableau 3 : Matrice d'expériences du plan factoriel complet

Numéro d'expérience	P_{size}	P_{cross}	P_{mut}	$Iter_{max}$
1	-	-	-	-
2	-	-	-	+
3	-	-	+	-
4	-	-	+	+
5	-	+	-	-
6	-	+	-	+
7	-	+	+	-
8	-	+	+	+
9	+	-	-	-
10	+	-	-	+
11	+	-	+	-
12	+	-	+	+
13	+	+	-	-
14	+	+	-	+
15	+	+	+	-
16	+	+	+	+

Pour mener toutes ces expériences, nous considérons la famille d'instances F5. Nous générons aléatoirement 10 instances avec un nombre de jobs égal à 80, et nous retenons la déviation moyenne par rapport à la borne inférieure (en pourcentage). Nous notons ici que la borne inférieure utilisée est celle proposée dans Ouled Bedhief et al. (2019).

Les résultats des simulations obtenus sont indiqués dans le tableau 4. Il suffit de regarder pour quelle expérience la valeur de la déviation moyenne est la plus petite pour connaître le meilleur réglage pour chaque variante de l'algorithme génétique.

Tableau 4 : Résultats d'ajustement des paramètres

Combinaison ($P_{size}; P_{cross}; P_{mut}; Iter_{max}$)	NOX/Échange	LOX/Échange	OX/Échange
	50;0.7; 0.05; 200	1,28	0,96
50; 0.7; 0.05; 800	1,16	0,72	0,77
50; 0.7; 0.1; 200	1,20	0,99	1,28
50; 0.7; 0.1; 800	0,89	0,72	0,69
50; 0.9; 0.05; 200	1,25	1,04	1,40
50; 0.9; 0.05; 800	0,92	0,64	0,99
50; 0.9; 0.1; 200	1,32	1,04	1,27
50; 0.9; 0.1; 800	0,92	0,64	0,79
100; 0.7; 0.05; 200	1,19	0,99	1,14
100; 0.7; 0.05; 800	0,83	0,77	0,92
100; 0.7; 0.1; 200	1,17	1,02	0,93
100; 0.7; 0.1; 800	0,84	0,72	0,48
100; 0.9; 0.05; 200	1,22	0,86	0,90
100; 0.9; 0.05; 800	1,00	0,72	0,48
100; 0.9; 0.1; 200	1,08	0,88	0,92
100; 0.9; 0.1; 800	0,83	0,63	0,46

A partir des résultats obtenus, nous présentons, dans le tableau 5, le meilleur réglage des paramètres des trois variantes de l'algorithme génétique.

Tableau 5: Choix des paramètres des variantes génétiques

	P_{size}	$Iter_{max}$	P_{cross}	P_{mut}
NOX/Échange	100	800	0.9	0.1
LOX/Échange	100	800	0.9	0.1
OX/Échange	100	800	0.9	0.1

Pour chaque classe d'instances de F1 à F5, nous testons plusieurs tailles différentes (5, 10, 40, 80, 120, 160, 200). Pour chaque taille du problème, nous générons 20 instances. Pour chaque instance, nous calculons le pourcentage de déviation de la solution obtenue par toute variante, par rapport à la borne inférieure proposée dans Ouled Bedhief et al. (2019).

Nous notons que pour les petites tailles (5 et 10 jobs), nous calculons la déviation de la solution obtenue par la variante génétique par rapport à la solution optimale donnée par le programme mathématique proposé dans [Ouled Bedhief et al. (2015)] pour le même problème.

Les tableaux 6 à 10 présentent les résultats obtenus en termes de déviation moyenne (en pourcentage) pour les cinq familles d'instances.

Tableau 6: Comparaison des variantes d'algorithme génétique pour F1

	<i>OX/ Échange</i>	<i>LOX/ Échange</i>	<i>NOX/ Échange</i>
$n = 5$	0.50	0.03	0.01
$n = 10$	0.52	0.05	0.00
$n = 40$	0.48	0.09	0.09
$n = 80$	0.12	0.01	0.00
$n = 120$	0.05	0.01	0.01
$n = 160$	0.04	0.04	0.03
$n = 200$	0.04	0.04	0.02

Tableau 7: Comparaison des variantes d'algorithme génétique pour F2

	<i>OX/ Échange</i>	<i>LOX/ Échange</i>	<i>NOX/ Échange</i>
$n = 5$	0.61	0.02	0.00
$n = 10$	0.49	0.06	0.01
$n = 40$	0.52	0.05	0.02
$n = 80$	0.21	0.02	0.02
$n = 120$	0.14	0.01	0.01
$n = 160$	0.09	0.004	0.004
$n = 200$	0.09	0.01	0.01

Tableau 8: Comparaison des variantes d'algorithme génétique pour F3

	<i>OX/ Échange</i>	<i>LOX/ Échange</i>	<i>NOX/ Échange</i>
$n = 5$	0.27	0.00	0.00
$n = 10$	0.17	0.00	0.00
$n = 40$	0.13	0.00	0.00
$n = 80$	0.06	0.00	0.00
$n = 120$	0.04	0.00	0.00
$n = 160$	0.07	0.01	0.01
$n = 200$	0.12	0.007	0.005

Tableau 9: Comparaison des variantes d'algorithme génétique pour F4

	<i>OX/ Échange</i>	<i>LOX/ Échange</i>	<i>NOX/ Échange</i>
$n = 5$	1.89	0.76	0.96
$n = 10$	1.46	0.63	0.63
$n = 40$	1.16	0.68	0.81
$n = 80$	0.95	0.32	0.36
$n = 120$	0.35	0.17	0.18
$n = 160$	0.42	0.16	0.19
$n = 200$	0.17	0.07	0.08

Tableau 10: Comparaison des variantes d'algorithme génétique pour F5

	<i>OX/ Échange</i>	<i>LOX/ Échange</i>	<i>NOX/ Échange</i>
$n = 5$	2.63	1.62	1.84
$n = 10$	2.45	1.50	1.59
$n = 40$	2.23	1.45	1.67
$n = 80$	1.16	0.64	0.85
$n = 120$	0.73	0.21	0.50
$n = 160$	0.32	0.27	0.46
$n = 200$	0.26	0.20	0.27

Nous constatons d'abord que les familles d'instances F1, F2 et F3 sont faciles à résoudre par l'une ou l'autre des variantes NOX/Échange, LOX/Échange et OX/Échange avec une légère préférence pour NOX/Échange, et qu'une solution optimale ou très proche de la borne inférieure est vite trouvée.

Par contre, les résultats des tableaux 9 et 10 montrent que la résolution des familles F4 et F5 est plus difficile, et que la borne inférieure est rarement atteinte. Nous trouvons ainsi que LOX/Échange s'empare de la première place pour la résolution de ces deux classes. En effet, bien que NOX/Échange ait donné de bons résultats en ne dépassant pas une déviation moyenne de 1.67%, elle est souvent dominée par la variante LOX/Échange.

Nous constatons, aussi, que la variante NOX/Échange utilisant l'opérateur de croisement NOX domine OX/Échange dans 94% des cas testés. En effet, cette dernière (OX/Échange) n'a permis d'obtenir les meilleures déviations moyennes par rapport à NOX/Échange que 2 fois sur 35.

Quant au temps de calcul, les algorithmes génétiques consomment, en général, énormément de temps. La figure 6 présente, par exemple, l'évolution du temps de résolution moyen mis par les trois variantes en fonction de la taille du problème, pour la famille F5. Ainsi, nous constatons que l'allure de la courbe de l'évolution du temps de calcul est similaire pour les trois variantes où il augmente proportionnellement avec la taille du problème. En plus, toutes les variantes de l'algorithme génétique nécessitent pratiquement une même durée moyenne qui ne dépasse pas une heure pour un nombre de jobs égal à 200.

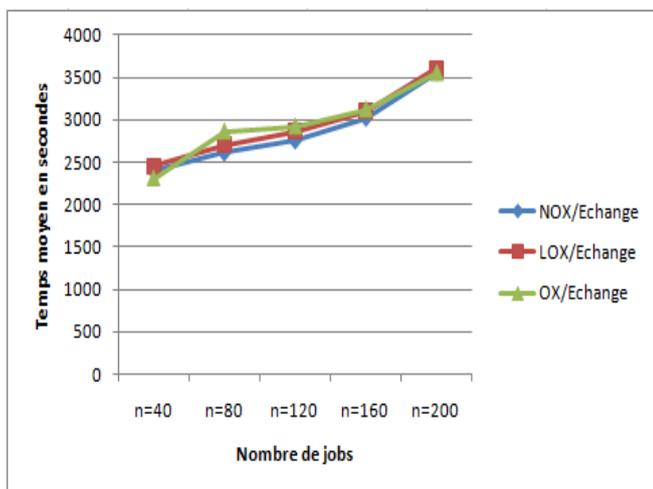


Figure 6. Temps moyen de résolution en secondes pour F5

CONCLUSION

Dans ce travail de recherche, nous avons traité une configuration de flow shop hybride à trois étages avec deux machines dédiées. L'objectif retenu dans cette étude est la minimisation du makespan. Pour la résolution de ce problème NP-difficile, nous avons proposé une métaheuristique. Elle consiste en un algorithme génétique. Nous avons ainsi comparé trois variantes de cet algorithme génétique qui diffèrent par l'opérateur de croisement utilisé.

Mais d'abord, nous avons effectué un plan factoriel complet afin de fixer le meilleur choix de chaque paramètre pour chaque variante de l'algorithme génétique.

Les résultats de simulation montrent que toutes les trois variantes sont performantes pour la résolution du problème. En effet, la déviation moyenne par rapport à la borne inférieure ne dépasse pas 2.23% pour les trois variantes.

Les résultats obtenus montrent, également, que l'algorithme génétique proposé fonctionne mieux lorsque l'opérateur de croisement LOX (Linear Order Crossover) est utilisé, et que l'opérateur NOX que nous proposons donne des résultats meilleurs que OX dans 94% des cas.

REFERENCES

- Besbes W, Loukil T, Teghem J, (2006). Using genetic algorithm in the multiprocessor flow shop to minimize the makespan, in: *Proceedings of the International Conference on Service Systems and Service Management*, France, pp. 1228–1233
- Dridi, N., Hadda, H., Hajri-Gabouj, S. (2009). Méthode heuristique pour le problème de flow shop hybride avec machines dédiées. *RAIRO Operations Research*, 43(4): 421-436. <http://dx.doi.org/10.1051/ro/2009024>
- Goldberg, D. (1989). *Genetic Algorithm in Search, Optimization, and Machine Learning*. The university of Alabama Addison Wesley publishing 1989.
- Graham, R.L., Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G. (1979). Optimization and approximation in deterministic sequencing and scheduling: A survey. *Annals of Discrete Mathematics*, 5: 287–326. [http://dx.doi.org/10.1016/S0167-5060\(08\)70356-X](http://dx.doi.org/10.1016/S0167-5060(08)70356-X)
- Grefenstette J.J, (1986). Optimization of control parameters for genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-16, No. 1: 122-128
- Hadda H, Dridi N, Hajri-Gabouj S, (2012). A note on the two-stage hybrid flow shop problem with dedicated machines. *Optim. Lett.* 6(8), 1731–1736
- Hadda, H., Dridi, N., Hajri-Gabouj, S. (2014). Exact resolution of the two-stage hybrid flow shop with dedicated machines. *Optimization Letters*, 8: 2329-2339. <http://dx.doi.org/10.1007/s11590-014-0741-y>
- Herrmann, J.W., Lee, C.Y. (1992). Three-machine look-ahead scheduling problems. Research Report No. 92–93, Department of Industrial Engineering, University of Florida, FL.
- Lin, B.M.T. (1999). The strong NP-hardness of two-stage flow shop scheduling with a common second-stage machine. *Computers & Operations Research*, 26(7): 695-698. [https://doi.org/10.1016/S0305-0548\(98\)00080-X](https://doi.org/10.1016/S0305-0548(98)00080-X)
- Liu Y, Kuei-Tang F, Lin B.M.T, (2013). A branch-and-bound algorithm for makespan minimization in differentiation flow shops. *Engineering Optimization*.

Mosheiov, G., Sarig, A. (2010). Minimum weighted number of tardy jobs on an m-machine flow shop with a critical machine. *European Journal of Operational Research*, 201(2): 404–408.
<http://dx.doi.org/10.1016/j.ejor.2009.03.018>

Oguz, C., Lin, B.M.T., Cheng, T.C.E. (1997). Two-stage flow shop scheduling problem with a common second-stage machine. *Computers & Operations Research*, 24(12): 1169–1174. [https://doi.org/10.1016/S0305-0548\(97\)00023-3](https://doi.org/10.1016/S0305-0548(97)00023-3)

Ouled Bedhief, A., Dridi, N. (2015). Méthode heuristique pour le flow shop hybride à trois étages avec machines dédiées. 11^{ème} congrès International de Génie industriel, CIGI_2015, Québec- Canada.

Ouled Bedhief, A., Dridi, N. (2019). Minimizing makespan in a three-stage hybrid flow shop with dedicated machines. *International Journal of Industrial Engineering Computations*, 10(2): 161-176.
<http://dx.doi.org/10.5267/j.ijiec.2018.10.001>

Riane, F., Artiba, A., Elmaghraby, S.E. (1998). A hybrid three-stage flow shop problem: Efficient heuristics to minimize makespan. *European Journal of Operational Research*, 109(2): 321–329. [https://doi.org/10.1016/S0377-2217\(98\)00060-5](https://doi.org/10.1016/S0377-2217(98)00060-5)

Riane, F., Artiba, A., Elmaghraby, S.E. (2002). Sequencing a hybrid two-stage flow shop with dedicated machines. *International Journal of Production Research*, 40(17): 4353–4380. <https://doi.org/10.1080/00207540210159536>

Wang, S., Liu, M. (2013). A heuristic method for two-stage hybrid flow shop with dedicated machines. *Computer & Operations Research*, 40: 438-450.
<http://dx.doi.org/10.1016/j.cor.2012.07.015>

Yang, J. (2011). Minimizing total completion time in two-stage hybrid flow shop with dedicated machines. *Computers and Operations Research*, 38(7): 1045–1053.
<http://dx.doi.org/10.1016/j.cor.2010.10.009>

Yang, J. (2013). A two-stage hybrid flow shop with dedicated machines at the first stage. *Computers & Operation Research*, 40(12): 2836–2843.
<http://dx.doi.org/10.1016/j.cor.2013.05.020>

Yang, J. (2015). Minimizing total completion time in two-stage hybrid flow shop with dedicated machines at the first stage. *Computers & Operations Research*, 58: 1–8.
<http://dx.doi.org/10.1016/j.cor.2014.11.012>

CIGI 2021

Économie circulaire sous incertitude des délais : Revue de la littérature

ILHEM SLAMA¹, OUSSAMA BEN-AMMAR², SADOK TURKI³, ALEXANDRE DOLGUI¹

¹ IMT Atlantique, LS2N-CNRS, La Chantrerie, 4 rue Alfred Kastler - B.P. 20722, 44307 Nantes, France
{ilhem.slama, alexandre.dolgui}@imt-atlantique.fr

² Mines Saint-Étienne, Université Clermont Auvergne, UMR-CNRS 6158 LIMOS, CMP Sciences de la Fabrication et Logistique, 880 route de Mimet, F-13541 Gardanne, France
oussama.ben-ammar@emse.fr

³ Laboratoire de Génie Informatique, de Production et de Maintenance (LGIPM), Université de Lorraine, UFR MIM, 3 Rue Augustin Fresnel, F-57070 Metz, France
sadok.turki@univ-lorraine.fr

Résumé – L'économie circulaire qui découle de la logistique inverse est actuellement un concept populaire promu par l'union européenne, par plusieurs gouvernements nationaux et par de nombreuses entreprises dans le monde entier. Dans ce papier préliminaire nous proposons une étude bibliographique non exhaustive sur la planification des opérations de désassemblage et d'assemblage, en mettant l'accent sur l'incertitude des délais. Ce papier est classé par catégories et discuté en fonction des principales opérations, à savoir le désassemblage et l'assemblage. Pour chacune des options de susmentionnées, ce papier élucide les définitions correspondantes, avec un aperçu structuré des méthodes de résolution. Une analyse transversale de la documentation examinée a mis en évidence un certain nombre de lacunes en matière de recherche et a révélé de multiples possibilités de recherche pour soutenir le développement de l'économie circulaire. Les principales conclusions montrent un intérêt croissant pour la durabilité des processus industriels linéaires traditionnels dans le contexte de l'économie circulaire.

Abstract – The circular economy that results from reverse logistics is currently a popular concept promoted by the European Union, several national governments and many companies around the world. In this paper, we propose a non-exhaustive bibliographical study on the planning of disassembly and assembly operations, focusing on the uncertainty of lead times. This paper is categorized and discussed according to the main operations, namely disassembly and assembly. For each of the above operation, this paper elucidates the corresponding definitions, with a structured overview of the methods of resolution. A cross-sectional analysis of the literature reviewed highlighted a number of research gaps and revealed multiple research opportunities to support the development of the circular economy. The main findings show a growing interest in the sustainability of traditional linear industrial processes in the context of the circular economy.

Mots clés : État de l'art, économie circulaire, système d'assemblage, système de désassemblage, délais aléatoires.

Keywords: Literature review, circular economy, assembly systems, disassembly systems, random lead-times.

1 INTRODUCTION

La logistique inverse et la gestion des déchets désignent toutes les opérations respectueuses de l'environnement liées à la réutilisation des produits et des matières premières. La chaîne mondiale d'approvisionnements, y compris pour les systèmes d'assemblage, de désassemblage et de recyclage, est perturbée par la crise liée au COVID19. Selon plusieurs analystes, Cette perturbation risque de durer et on va même jusqu'à la comparer à un grand coup de fouet dont la disparition de ces conséquences va prendre un temps très long. Une autre raison afin d'expliquer que c'est difficile de faire stabiliser la chaîne logistique, est qu'en ce temps de crise, les entreprises ainsi que leurs clients s'organisent souvent au jour le jour, alors d'habitude ils fonctionnent en flux tendus. Pour résumer, la stabilité des chaînes logistiques n'est pas de l'ordre du jour.

L'économie circulaire est un concept vaste. C'est tout un processus long qui inclut les opérations d'assemblage et de

désassemblage. Dans cette étude préliminaire, nous élaborons une exploration non exhaustive de la littérature concernant ces deux systèmes tout en tenant en considération les perturbations engendrées par les incertitudes des délais. Dans l'économie circulaire, les produits usagés sont collectés auprès des clients avant qu'ils soient recyclés, reconditionnés, réparés, ou même éliminés [Galantucci et al. 2004]. Dans cette étude bibliographique, nous nous focalisons sur le processus de re-fabrication, c'est-à-dire la réutilisation de composants obtenus à partir de produits en fin de vie afin de les réintégrer dans le processus d'assemblage.

Le processus de re-fabrication préserve l'identité du produit (ou de la pièce) et effectue les opérations de désassemblage, de tri, de remise à neuf et d'assemblage nécessaires pour amener le produit à un niveau de qualité souhaité. L'étude du processus de re-fabrication de produits usagés est largement abordée dans la

littérature scientifique sur la logistique inverse [Tighazoui et al., 2019, Suzanne et al. 2020].

L'assemblage et le désassemblage sont deux processus qui bénéficient de nombreux avantages de processus de re-fabrication [Brennan et al. 1994]. Le processus d'assemblage est l'une des étapes de fabrication les plus longues et les plus coûteuses. Les aspects de désassemblage doivent être pris en compte dans le processus de re-fabrication sans perturber le processus d'assemblage. Autrement dit, les chaînes de production subissent déjà plusieurs sources et facteurs d'incertitudes. La réinjection de pièces ou de composants remis à neuf ne doit se faire sans perturber les plans d'approvisionnement et de production existants. C'est pour cette raison, que dans cette étude nous nous intéressons à ces deux systèmes primordiaux dans l'économie circulaire. L'idée est de chercher où encore faire des économies financières afin que les produits issus de l'économie circulaire soient plus compétitifs par rapport aux produits tout neufs.

Les principaux objectifs du désassemblage sont la maintenance, la remise à neuf, le recyclage ou l'élimination des produits en fin de vie. À mesure que la complexité des produits et des systèmes de production augmente, il est nécessaire de prévoir une planification et un contrôle appropriés qui aident les planificateurs à traiter les aspects d'assemblage et de désassemblage.

Dans la pratique, la planification de la production et le contrôle des stocks, tous deux gérés dans le cadre des systèmes d'assemblage et de désassemblage, sont souvent soumis à diverses sources d'incertitude exogène telles que la variabilité de la demande, les pannes de machines, les retards de transport, taux de récupération, problèmes de qualité, une mauvaise estimation de délais d'obtention des composants, etc. [Daniel et al. 1997].

Le délai de désassemblage et/ou délai d'obtention des composants est un paramètre crucial dans la planification des opérations de désassemblage. Concernant les systèmes d'assemblage, les délais d'approvisionnement en composants, représentent le temps nécessaire pour qu'une commande en composants arrive. Cependant, la variabilité de délais affecte fortement l'efficacité de la chaîne logistique et peut affecter les commandes des clients avec une disponibilité aléatoire des articles. À cet égard, le présent document se concentre sur le processus de re-fabrication de produits usagés avec des activités de désassemblage et d'assemblage sous incertitude de délais. L'objectif de ce travail préliminaire est d'examiner, de classer et d'évaluer les études existantes afin de fournir une vue systématique des travaux passés et une vision appropriée pour les études futures.

Le reste de cet article est organisé comme suit. La section 2 présente une revue de la littérature des travaux traitant les systèmes d'assemblage et/ou de désassemblage sous incertitude des délais. Dans la section 3, plusieurs pistes de recherche futures sont suggérées. Enfin, la section 4 résume et conclut cette recherche.

2 PLANIFICATION D'ASSEMBLAGE/DESASSEMBLAGE SOUS INCERTITUDE DE DELAIS

Cette recherche est basée sur la planification des besoins en matériaux (*Material Requirements Planning* : MRP) dans le cadre des systèmes d'assemblage et de désassemblage sous

incertitude de délais. Les systèmes MRP ont été parmi les méthodes de production et de planification des stocks les plus populaires auprès des entreprises. Comme les systèmes MRP, les R-MRP font partie des systèmes ERP. Cependant la quasi majorité des systèmes existants ne prennent pas en considération les incertitudes. Ceci est expliqué par le fait que l'optimisation de la production sous incertitude est très complexe. Dans la littérature scientifique, nous observons un intérêt croissant pour la planification sous incertitude. Les études existantes peuvent être scindées en plusieurs catégories. Une première catégorie qui s'intéresse aux systèmes de désassemblage et une deuxième qui se focalise sur les systèmes d'assemblage. A notre connaissance, la majorité des études existantes se concentre sur la deuxième catégorie (voir Tableau 1. De plus, nous pouvons également les classer en problèmes mono et multi-produit, avec ou sans pièces communes ainsi avec ou sans des contraintes de capacité. Dans les sous-sections suivantes, nous allons classer les travaux existants en fonction de type de systèmes.

2.1 Système d'assemblage

La planification des besoins en matériaux, qui est un élément essentiel dans la planification des ressources, est nécessaire pour déterminer la taille des lots de production et d'approvisionnement en composants. La tailles des lots est généralement calculée en supposant que l'ensemble des paramètres, à savoir la demande, les délais, la capacité, le rendement, etc., sont déterministes.

Compte tenu des demandes en produits finis, le système MRP calcule les quantités à produire (ou à commander) pour chaque composant à chaque période. Ces calculs sont basés sur les nomenclatures (voir Figure 1), qui indiquent la hiérarchie des composants (c'est-à-dire la structure du produit).

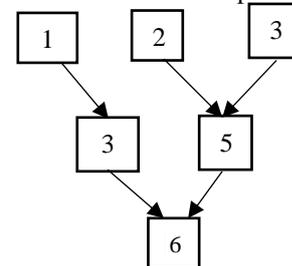


Figure 1. Système d'assemblage à deux niveaux.

Au cours des dernières décennies, l'industrie et les chercheurs ont accordé une plus grande attention à la gestion de l'incertitude et à ses effets sur la performance de la chaîne d'approvisionnement. La vulnérabilité croissante des chaînes d'approvisionnement face à ces incertitudes et la nécessité de définir des stratégies appropriées pour en faire face ont été fréquemment signalées [Snyder et al., 2016 ; Speier et al., 2011 ; Kleindorfer et Saad, 2005]. Ainsi, il existe un intérêt croissant pour étudier la planification des approvisionnements en tenant compte de l'incertitude d'un ou plusieurs paramètres.

Concernant l'incertitude des délais, les approches de modélisations proposées dans la littérature se limitent à :

- Un seul niveau dans la nomenclature, avec une demande constante sur l'horizon de planification ;
- Deux niveaux dans la nomenclature, avec une planification mono-période. C'est-à-dire faire la planification pour satisfaire une seule demande.

Tableau 1. Systèmes d'assemblage et de désassemblage : classification des travaux existants.

A. Système d'assemblage

Auteurs	Critères	Type de système					Commentaires	
		Mono-niveau	Deux-niveau	Multi-niveau	Multi-période	Une-période		
[Kumar, 1989]	Sc, Rc	✓				✓	Étude générique	
[Chu et al., 1993]		✓				✓	Algorithme itératif	
[Dolgui, 2002]		✓			✓		Modèle PLNE, simulation, heuristiques, point de commande optimale, lot d'assemblage optimal	
[Proth et al., 1997]		✓			✓		Optimisation stochastique, heuristiques	
[Dolgui et O. Louly, 2002]		✓			✓		Markov, Newsboy, politique L4L, les délais d'approvisionnement suivent la même distribution de probabilité	
[O. Louly et Dolgui, 2002a]		✓			✓		Markov, Newsboy, tous les composants ont les mêmes propriétés	
[O. Louly et Dolgui, 2004]		✓			✓		Markov, Newsboy, politique POQ, les composants suivent la même distribution et ont le même coût de stockage	
[O. Louly et Dolgui, 2009]		✓			✓		Politique EOQ, Branch et Bound (B&B)	
[O. Louly et al., 2008a]		Sc, Ts	✓			✓		B&B
[O. Louly et al., 2008b]			✓			✓		Généralisation de [O. Louly et Dolgui, 2002a]
[Chauhan et al., 2009]	Sc, Rc	✓				✓	Recuit simulé, solution exacte, modèle continue	
[O. Louly et Dolgui, 2011]		✓			✓		Politique POQ, une généralisation du modèle Newsboy	
[O. Louly et Dolgui, 2013]	Oc, Sc, Ts	✓			✓			
[Shojae et al., 2015]	Sc, Ts	✓			✓		Politique POQ, optimisation	
[Borodin et al., 2016]	Sc, Rc	✓				✓	Chance-constrained programming	
[Hnaïen et Afsar, 2017]	Oc, Sc, Ts	✓				✓	Modèles PLMNE	
[Yano, 1987 c]	Sc, Rc		✓			✓	Programmation non-linéaire	
[Hegedus et Hopp, 2001]	Sc, Ts		✓			✓	Méthode d'optimisation combinatoire	
[Tang et Grubbström, 2003]	Sc, Rc		✓			✓	Transformation de Laplace	
[Hnaïen et al., 2009]			✓			✓	Algorithme Génétique (AG), B&B	
[Hnaïen et al., 2010]	Sc, Ts		✓			✓	AG, multi-objectif, sélection élitiste	
[Fallah-Jamshidi et al., 2011]	Sc, Rc		✓			✓	Approche hybride, AG spécifique	
[Sakiani et al., 2012]	Sc, Ts		✓			✓	AG, multi-objectif, sélection par tournoi	
[Guiras et al., 2016]	Mc, Sc, Rc		✓			✓	AG, coût de maintenance	
[Ben-Ammar et Dolgui, 2018]	Sc, Rc		✓			✓	B&B et AG	
[Axsäter, 2005]	Sc, Rc		✓*			✓	Techniques d'approximation et de décomposition, distributions continues	
[Ben-Ammar et al., 2018]	Sc, Rc			✓		✓	Formulation récursive, B&B	
[Ben-Ammar et al., 2020]	Sc, Rc			✓		✓	Modélisation générique, processus itératif, AG hybride	

B. Système de désassemblage

Auteur	Niveau		Produit		Capacité	Objectif	Approche de résolution
	Deux-	Multi-	Mono-	Multi-			
[Slama et al., 2019b]	✓		✓			Rc, Sc, Oc	PLMNE
[Slama et al., 2020a]	✓		✓			Rc, Sc	Newsboy
[Slama et al., 2020b]	✓		✓		✓	Rc, Sc, Oc	SAA

(Rc: coût de retard, Sc: coût de stockage, Ts: Taux de service, Oc: coût de lancement, Mc: coût de maintenance, ✓*: Trois-niveau)

Pour le cas des nomenclatures multi-niveau, [Ben-Ammar et al. 2018, 2020] ont été les premiers à proposer une approche analytique pour optimiser l'approvisionnement en composants. Cependant, malgré la contribution remarquable de ces travaux, Elles se limitent à une seule demande à satisfaire. Ceci est expliqué par la difficulté de modéliser un système de production multi-niveau et multi-période où il existe une interdépendance entre les périodes et les niveaux.

Malgré l'importance de ce problème, les études existantes restent insuffisantes comparées à celles développées pour faire face à l'incertitude de la demande. Dans [Dolgui et al., 2013], les auteurs présente une revue de la littérature classant les techniques et les approches existantes. Les auteurs confirment ce qui a été mentionné précédemment : la plupart des modèles existants s'intéressent à (i) des systèmes d'assemblage à un niveau et une planification à une ou plusieurs périodes ou (ii) des systèmes d'assemblage à deux niveaux et une planification sur une période unique.

[Yano, 1987a, b] a été parmi les premiers à étudier les effets de la variabilité des délais. Elle a étudié le cas des systèmes de production en série avec une planification mono-période. Ce travail a été généralisé par [Elhafsi, 2002]. Dans une troisième étude, [Yano, 1987c] a étudié un système d'assemblage à un niveau et une planification sur une période. Un composant est assemblé à partir de deux composants. Le délai du premier composant suit une loi de Poisson et le deuxième une loi binomiale négative. La fonction objectif minimise l'espérance mathématique du coût total qui est composé des coûts de rupture et de stockage. Un algorithme basé sur une exploration des propriétés de la fonction objectif détermine les délais planifiés optimaux qui minimise l'espérance mathématique du coût total. [Kumar, 1989] a étudié le cas de la planification à une seule période et des systèmes d'assemblage à un niveau. Comme [Yano, 1987a], il a minimisé l'espérance mathématique du coût total et proposé des dates optimales de lancement des commandes. Une analyse exacte a été proposée avec des hypothèses de types particuliers de distribution (normale, uniforme et exponentielle). Quelques années plus tard, [Chu et al., 1993] se sont penchés sur le cas d'un système d'assemblage à un niveau. Les auteurs ont prouvé la convexité de l'espérance mathématique du coût total et proposé un algorithme itératif pour la minimiser. Pour résoudre ce problème dans le cas d'un modèle de production en deux étapes, [Hegedus et Hopp, 2001] ont proposé une approche basée sur une formulation analytique de type Newsboy. La principale faiblesse de ces études est qu'elles sont limitées à des systèmes d'assemblage à un ou deux niveaux.

Pour le cas des modèles multi-période et des systèmes d'assemblage à un niveau, [Ould-Louly et Dolgui, 2002a, b, 2004, 2009, 2011, 2013, Ould-Louly et al., 2008b, a] et [Shojaie et al., 2015] ont présenté des modèles mathématiques avec des délais aléatoires, une demande de produit fini constante (la même pour toutes les périodes) et une capacité d'assemblage illimitée. Le même problème a été étudié par [Hnaien et Afsar, 2017] pour modéliser le problème avec une approche robuste et un ensemble de scénarios. Une analyse de complexité est fournie. Elle prouve que le problème est NP-difficile même lorsqu'il existe deux scénarios.

Dans le cas de la planification sur une période et des systèmes d'assemblage à deux niveaux, [Tang et Grubbstrom, 2003] ont modélisé à la fois les temps de processus et les délais d'approvisionnement pour les composants. Ils ont considéré une demande fixe, une capacité illimitée et une date de livraison connue. Plus tard, [Hnaien et al., 2009] et [Fallah Jamshidi et al., 2011] ont traité le même problème en utilisant des algorithmes génétiques pour minimiser l'espérance mathématique du coût total.

Quelques années plus tard, [Hnaien et al., 2016] ont étudié le cas d'un modèle à une seule période pour un système d'assemblage à un niveau avec une demande et des délais stochastiques. Un modèle analytique et une approche par séparation et évaluation ont été proposé pour optimiser les quantités de composants et les délais planifiés. Bien que les auteurs aient limité leur modèle à une période, il peut être étendu à la planification sur plusieurs périodes pour tenir compte des compromis possibles entre les stocks de différentes périodes. En s'appuyant sur ces travaux, [Borodin et al., 2016] ont proposé une nouvelle approche basée sur la contrainte de hasard et un modèle linéaire pour résoudre le même problème. Cependant, ils se sont limités aux systèmes d'assemblage à un niveau en raison de la difficulté de modéliser la dépendance entre les niveaux. Cette difficulté a été surmontée dans l'étude faite par [Ben-Ammar et al., 2020].

[Atan et al., 2016] ont récemment envisagé un processus d'assemblage final alimenté par un processus parallèle en plusieurs étapes. Chaque étape avait un temps de traitement stochastique et le système était contrôlé par des délais planifiés à chaque étape. Les auteurs ont développé une heuristique itérative pour optimiser les délais prévus pour les différentes étapes et pour minimiser l'espérance mathématique du coût total.

2.2 Système de désassemblage

Le problème de planification des opérations de désassemblage est le problème de la détermination de la quantité et de date de désassemblage des produits en fin de vie tout en satisfaisant la demande de leurs pièces/composants sur un horizon de planification fini. La planification de désassemblage est l'un des principaux problèmes de planification de la production dans les systèmes de désassemblage. En d'autres termes, à partir de sa solution, il est possible de déterminer quels produits, combien et quand les démonter [Slama et al., 2020c].

En général, la planification de désassemblage peut être considéré comme une version inversée du problème de planification des opération d'assemblage, puisqu'il s'agit essentiellement d'une forme inversée du problème ordinaire de MRP [Lee et al., 2004]. La Figure 2 représente une nomenclature de produit à désassembler dans le cas de désassemblage où la demande et les sources d'approvisionnement sont inversées.

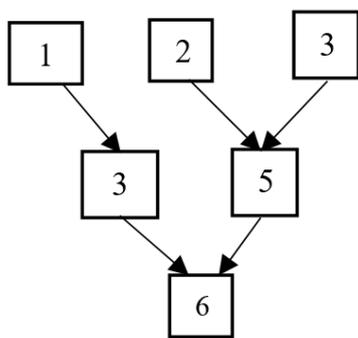


Figure 2. Système de désassemblage à trois niveaux.

Comme le souligne la littérature, la planification du désassemblage ne peut être similaire à un problème de planification de la production. Concernant sa conception, le processus d'assemblage converge vers une seule source de demande (produit final), tandis que le processus de désassemblage diverge vers de multiples sources de demande (pièces ou sous-ensemble). En raison de la structure divergente du désassemblage, la complexité des problèmes connexes augmente considérablement avec le nombre de types de produits à désassembler [Prakash et al., 2012, Slama et al., 2019a].

La planification du désassemblage et la gestion des stocks sont importantes pour les entreprises afin de fournir aux clients des composants usagés à des prix compétitifs. Pour atteindre cet objectif, l'une des priorités des planificateurs est de réduire le niveau prévu des stocks dans un environnement incertain.

La revue de la littérature en contexte stochastique est très limitée et peut être divisée en quatre catégories en fonction des incertitudes : (i) l'incertitude de la demande en composants [voir Barba-Gutiérrez et Adenso-Díaz, 2009 et Kim et al., 2010], (ii) l'incertitude du rendement de désassemblage. Le rendement aléatoire se produit lorsque la différence entre les quantités d'articles planifiées et celles obtenues après le désassemblage n'est pas connue (voir par exemple les travaux de [Inderfurth et Langella, 2005, Inderfurth et al., 2015]), (iii) l'incertitude simultanée du rendement et de la demande [Liu et Zhang, 2018] et (iv) l'incertitude du délai de désassemblage. Ce dernier peut être défini comme la différence de temps entre le lancement d'un ordre de désassemblage et la réception des éléments démontés. Il peut être exprimé en nombre de périodes (jours, semaine, mois, etc.).

En se basant sur la littérature, le délai de désassemblage du produit en fin de vie dans les systèmes de désassemblage est généralement supposé constant ou égal à zéro [Kim et al., 2003, 2006a, b, 2009]. Dans la réalité industrielle, le délai de désassemblage est loin d'être constant en raison du niveau élevé d'incertitude du processus de désassemblage. Cette incertitude peut être causée par des problèmes techniques (absentéisme, capacité de démontage limitée, problème de qualité, etc.).

Peu de travaux sont proposés dans la littérature pour aider à décider le plan optimal de désassemblage quand les délais de désassemblage sont aléatoires. [Slama et al., 2019b] sont les premiers à traiter ce type de problème. Les réalisations possibles du processus stochastiques décrivant des paramètres incertains tout au long de l'horizon décisionnel sont représentés via un ensemble de scénarios. Un modèle de

programmation linéaire mixte en nombres entiers est proposé pour traiter un seul type de produit et un système de désassemblage à deux niveaux. La planification est faite sur un horizon de planification multi-période. La capacité de désassemblage est supposée infinie. Le modèle proposé est utilisé pour déterminer la quantité et la date optimale de lancement de l'opération de désassemblage, afin de minimiser le coût total moyen sur l'horizon de planification.

Récemment, [Slama et al., 2020a] ont développé un modèle pour analyser un système de désassemblage mono-période quand les délais de désassemblage sont incertains. Le cas d'un seul type de produit en fin de vie qui est décomposé en plusieurs articles est considéré. En premier lieu, une expression analytique de la fonction du coût est développée. Cette fonction minimise la somme des coûts de stockage et de rupture des composants. En deuxième lieu, une approche basée sur la formule Newsboy est proposée pour minimiser l'espérance mathématique du coût total et optimiser la date de début du processus de désassemblage. Cette recherche a étudié l'effet de l'incertitude des délais sur le problème de planification du désassemblage. Les analyses montrent que la variabilité du délai de désassemblage affecte fortement la performance du système. A notre connaissance, cette étude a été la première à examiner les systèmes de désassemblage en utilisant le modèle Newsboy. Pour une étude approfondie des problèmes de Newsboy, les lecteurs peuvent se référer aux recherches proposées par [Khouja, 1999].

Par la suite, [Slama et al., 2020b] étend le travail de [Slama et al., 2020a] en intégrant la contrainte de la capacité de désassemblage à chaque période de l'horizon de planification. Les délais stochastiques sont modélisés par des intervalles fermés en utilisant l'approche basée sur les scénarios. Ces délais ont été pris comme des variables aléatoires indépendantes avec des distributions de probabilité quelconques connues. L'objectif est de proposer le meilleur plan de désassemblage garantissant le meilleur compromis entre les coûts de setup et la somme d'espérance des coûts de stockage et de rupture de stock sur un horizon de planification fini. Le problème étudié est formulé comme un modèle de Programmation Linéaire Mixte en Nombres Entiers (PLMNE) stochastique sous un ensemble de scénarios possibles. Les résultats montrent que, pour un nombre particulier de composants/périodes ou des délais, PLMNE peut fournir un nombre exponentiel de scénarios. Pour résoudre des problèmes à grande échelle, les auteurs ont couplé cette approche avec l'échantillonnage de Monte Carlo. Les résultats montrent que la méthode d'échantillonnage proposée fournit une bonne approximation, et les tests montrent la convergence de l'estimateur de l'Approximation Moyenne de l'Echantillon.

2.3 Système d'assemblage/désassemblage

La réalisation d'un processus efficace d'assemblage/désassemblage est un facteur clé pour la compétitivité d'une entreprise performante. Les réglementations environnementales et la pression des clients pour fabriquer des produits plus respectueux de l'environnement, obligent les entreprises à intégrer davantage d'assemblage/désassemblage dans l'environnement de fabrication. Cependant, l'efficacité des processus d'assemblage/désassemblage existants est encore faible en

raison de la difficulté inhérente à développer des systèmes entièrement automatisés et de la complexité de l'exécution des opérations pour une large gamme de produits [Guide V.D.R., 2000].

Dans ce contexte, [Guiras et al., 2018] sont les premiers à étudier une optimisation du système de planification du démontage, de la remise à neuf et d'assemblage. Le système de désassemblage-assemblage est soumis à des pannes aléatoires au cours de l'horizon de planification, ce qui nous conduit à intégrer une politique de maintenance pour faire face à l'indisponibilité du système. Un modèle mathématique est proposé pour formuler le problème. Un algorithme d'optimisation basé sur un algorithme génétique et une simulation est développé pour trouver la date optimale de lancement de produit en fin de vie ainsi que les dates de lancement optimales pour acheter de nouveaux composants afin de satisfaire la demande des clients à une date de livraison donnée. Cette étude est très intéressante. Cependant, nous remarquons une lacune dans le nombre d'études tenant compte des incertitudes dans un processus de désassemblage suivi d'un processus d'assemblage.

3 PERSPECTIVES DE RECHERCHE

La crise actuelle de coronavirus a eu un impact critique sur les chaînes d'approvisionnement. En particulier, de nombreuses entreprises n'ont pas pu obtenir les composants nécessaires à leur production, souvent en raison de la fermeture des frontières. De telles perturbations sont le résultat de la mondialisation, où les fournisseurs se sont répandus dans le monde entier. L'une des principales recommandations pour éviter l'interruption des flux de matières et renforcer la résilience de la production est de s'appuyer sur des sources d'approvisionnement locales. La crise des coronavirus montre clairement que nous ne pouvons pas nous permettre d'être totalement dépendants des fournisseurs de loin. Si la mondialisation peut avoir apporté de nombreux avantages, il est désormais nécessaire de pouvoir se limiter aux ressources déjà disponibles sur nos marchés respectifs.

Dans ce contexte, la récupération et le reconditionnement ou la remise à neuf des produits usagés est considéré comme une solution la plus économique et écologique. Alors, la plupart des systèmes de reconditionnement sont plus respectueux à l'environnement que les systèmes de production traditionnels. En effet, le processus de reconditionnement implique la récupération et la remise en bonne état des produits usagés au lieu de consommer des matières premières. De plus, il permet de réduire les émissions de carbone, en raison que la quantité de carbone pour produire un produit neuf est plus élevée que celle pour reconditionner un produit usagé. En effet, comme il est connu, un produit neuf est fabriqué à partir de la matière première et nécessite des différentes procédures pour le finir. Donc, pour un produit neuf le billon carbone est plus important que celui pour un produit reconditionné [Turki et al., 2018]. Généralement dans un système de reconditionnement on a principalement le processus suivant : désassemblage du produit usagé, remise en état les composants désassemblés et l'assemblage des composants (voir Figure 3).

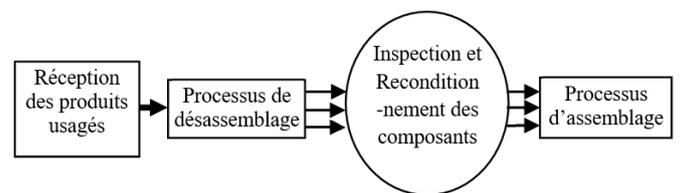


Figure 3. Système de reconditionnement.

Dans plusieurs cas de reconditionnement, comme par exemple le reconditionnement des pièces d'automobile ou les produits électroménager. Il est nécessaire de désassembler chaque produit pour inspecter et remettre en état chaque composant. Parfois il est nécessaire de remplacer un composant très abimé par un composant neuf. Finalement, les composants sont assemblés pour obtenir un produit reconditionné. La nomenclature de l'assemblage est l'inverse de celle de désassemblage.

L'économie circulaire peut apporter une contribution majeure à l'exploitation de ce potentiel : les produits sont ramenés dans les usines en fin de cycle de vie, au lieu d'aller en décharge. De retour dans les usines, les produits usagés sont démontés en pièces. De cette façon, un flux de matières circulaire indépendant est créé à l'écart des fournisseurs mondiaux.

4 CONCLUSION

Cette recherche est principalement motivée par les tendances actuelles du monde industriel. Comme mentionné précédemment, il existe aujourd'hui une forte tendance à la sensibilisation à l'environnement et à la conservation des ressources et de l'énergie. Cette étude a examiné les problèmes de planification des opérations des systèmes d'assemblage et de désassemblage sous incertitude de délais. Dans cette étude, nous avons passé en revue les travaux de recherche sur ce problème pour chaque système et analysé les différents modèles et méthodes de résolution utilisés pour les résoudre.

Les conclusions de cette étude révèlent de nouvelles idées sur la manière dont les options de valorisation peuvent être intégrées de manière appropriée dans les environnements de production traditionnels pour converger vers une économie respectueuse de l'environnement. Un accent particulier a été mis sur l'économie circulaire. Sur la base de l'ensemble des documents rassemblés, l'analyse documentaire présentée dans ce document peut être qualifiée de subjective. Malgré cela, nous pensons que les publications sélectionnées donnent une vue complète de l'état des connaissances pour les chercheurs.

5 REMERCIEMENT

Ce travail a été réalisé dans le cadre du projet RMS et les auteurs remercient chaleureusement la région Pays de la Loire pour le financement.

6 REFERENCES

Atan, Z., de Kok, T., Dellaert, N. P., van Boxel, R., & Janssen, F. (2016). Setting planned lead times in customer-order driven assembly systems. *Manufacturing*

- & *Service Operations Management*, 18, 122–140
- Barba-Gutiérrez, Y., & Adenso-Díaz, B. (2009). Reverse MRP under uncertain and imprecise demand. *International Journal of Advanced Manufacturing Technology*, 40(3–4), 413–424.
- Brennan, L., Gupta, S. M., & Taleb, K. N. (1994). Operations planning issues in an assembly/disassembly environment. *International Journal of Operations & Production Management*.
- Ben-Ammar O., & Dolgui A. (2018). Optimal order release dates for two-level assembly systems with stochastic lead times at each level, *International Journal of Production Research*, 56(12), p. 4226-4242.
- Ben-Ammar, O., Castagliola, P., Dolgui, A., & Hnaien, F. (2020). A hybrid genetic algorithm for a multilevel assembly replenishment-planning problem with stochastic lead times. *Computers & Industrial Engineering*, 149, 106794.
- Ben-Ammar, O., Dolgui, A., and Wu, D.D. (2018). Planned lead times optimization for multi-level assembly systems under uncertainties. *Omega*, 78, 39–56.
- Borodin, V., Dolgui, A., Hnaien, F., & Labadie, N. (2016). Component replenishment planning for a single-level assembly system under random lead times: A chance constrained programming approach. *International Journal of Production Economics*, 181, 79–86.
- Chauhan S., Dolgui A., & Proth J.-M. (2009). A continuous model for supply planning of assembly systems with stochastic component procurement times. *International Journal of Production Economics*, 120, p. 411-417.
- Chu, C., Proth, J.-M., & Xie, X. (1993). Supply management in assembly systems. *Naval Research Logistics (NRL)*, 40, 933–949
- Daniel, V., Guide, R., & Srivastava, R. (1997). An evaluation of order release strategies in a remanufacturing environment. *Computers and Operations Research*, 24(1), 37–47.
- Dolgui A. (2002). A model of joint control of reserves in automatic control systems of production. *Automation and Remote Control*, 62, p. 2020-2026.
- Dolgui, A., Ben-Ammar, O., Hnaien, F., & Ould-Louly, M.-A. (2013). A state of the art on supply planning and inventory control under lead time uncertainty. *Studies in Informatics and Control*, 22, 255–268.
- Dolgui A., & Ould Louly M. (2002). A model for supply planning under lead time uncertainty. *International Journal of Production Economics*, 78, p. 145-152.
- Elhafsi, M. (2002). Optimal leadtimes planning in serial production systems with earliness and tardiness costs. *IIE Transactions*, 34, 233–243.
- Fallah-Jamshidi, S., Karimi, N., & Zandieh, M. (2011). A hybrid multi-objective genetic algorithm for planning order release date in two-level assembly system with random lead times. *Expert Systems with Applications*, 38, 13549–13554.
- Galantucci, L. M., Percoco, G., & Spina, R. (2004). Assembly and disassembly planning by using fuzzy logic & genetic algorithms. *International Journal of Advanced Robotic Systems*, 1(2), 7.
- Guide Jr, V. D. R. (2000). Production planning and control for remanufacturing: industry practice and research needs. *Journal of operations Management*, 18(4), 467–483.
- Guiras, Z., Turki, S., Rezg, N., & Dolgui, A. (2018). Optimization of two-level disassembly/ remanufacturing/ assembly system with an integrated maintenance strategy. *Applied Sciences*, 8(5), 666.
- Hegedus, M. G., & Hopp, W. J. (2001). Setting procurement safety lead-times for assembly systems. *International Journal of Production Research*, 39, 3459–3478.
- Hnaien, F., & Afsar, H.-M. (2017). Robust single-item lot-sizing problems with discrete-scenario lead time. *International Journal of Production Economics*, 185, 223–229.
- Hnaien, F., Delorme, X., & Dolgui, A. (2009). Genetic algorithm for supply planning in two-level assembly systems with random lead times. *Engineering Applications of Artificial Intelligence*, 22, 906–915.
- Hnaien, F., Delorme, X., & Dolgui, A. (2010). Multi-objective optimization for inventory control in two-level assembly systems under uncertainty of lead times. *Computers & Operations Research*, 37, 1835–1843.
- Hnaien, F., Dolgui, A., & Wu, D. D. (2016). Single-period inventory model for one-level assembly system with stochastic lead times and demand. *International Journal of Production Research*, 54, 186–203.
- Inderfurth, K., & Langella, I. M. (2006). Heuristics for solving disassemble-to-order problems with stochastic yields. *OR Spectrum*, 28(1), 73–99.
- Inderfurth, K., Vogelgesang, S., & Langella, I. M. (2015). How yield process misspecification affects the solution of disassemble-to-order problems. *International Journal of Production Economics*, 169, 56–67.
- Khouja, M. (1999). The single-period (news-vendor) problem: literature review and suggestions for future research. *Omega*, 27(5), 537–553.
- Kim, H. J., & Xirouchakis, P. (2010). Capacitated disassembly scheduling with random demand. *International Journal of Production Research*, 48(23), 7177–7194.
- Kim, H. J., Lee, D. H., & Xirouchakis, P. (2006a). Two-phase heuristic for disassembly scheduling with multiple product types and parts commonality. *International Journal of Production Research*, 44(1), 195–212.
- Kim, H. J., Lee, D. H., & Xirouchakis, P. (2006b). A Lagrangean heuristic algorithm for disassembly scheduling with capacity constraints. *Journal of the Operational Research Society*, 57(10), 1231–1240.
- Kim, H. J., Lee, D. H., Xirouchakis, P., & Kwon, O. K. (2009). A branch and bound algorithm for disassembly scheduling with assembly product structure. *Journal of the Operational Research Society*, 60(3), 419–430.
- Kim, H. J., Lee, D. H., Xirouchakis, P., & Züst, R. (2003). Disassembly scheduling with multiple product types. *CIRP*
- Kleindorfer, P. R., & Saad, G. H. (2005). Managing disruption risks in supply chains. *Production and Operations Management*, 14(1), 53-68.
- Kumar A. (1989). Component Inventory Costs in an Assembly Problem with Uncertain Supplier Lead-Times, *IIE Transactions*, 21, p. 112-121.
- Lee, D. H., & Xirouchakis, P. (2004). A two-stage heuristic for disassembly scheduling with assembly product structure. *Journal of the Operational Research Society*, 55(3), 287–297.
- Liu, K., & Zhang, Z. H. (2018). Capacitated disassembly scheduling under stochastic yield and demand. *European Journal of Operational Research*, 269(1), 244–257.
- Kumar, A. (1989). Component inventory costs in an assembly problem with uncertain supplier lead-times. *IIE*

- Transactions*, 21, 112–121
- Ould Louly M. -A., & Dolgui A. (2002a). Generalized newsboy model to compute the optimal planned lead times in assembly systems. *International Journal of Production Research*, 40(17), p. 4401-4414.
- Ould Louly M. -A., & Dolgui A. (2002b). Supply Planning Optimization under Uncertainties. *International Journal of Agile Manufacturing*, 5(1), pp. 17-26.
- Ould Louly M., & Dolgui A. (2004). The MPS parameterization under lead time uncertainty. *International Journal of Production Economics*, 90(04), p. 369-376.
- Ould Louly M. -A., & Dolgui A. (2009). Calculating safety stocks for assembly systems with random component procurement lead times: A branch and bound algorithm. *European Journal of Operational Research*, 199(3), pp. 723-731.
- Ould Louly M. -A., Dolgui A., & Hnaien, F. (2008a). Supply planning for single-level assembly system with stochastic component delivery times and service level constraint. *International Journal of Production Economics*, 115(1), p. 236-247.
- Ould Louly M. -A., Dolgui A., & Hnaien, F. (2008b). Optimal supply planning in MRP environments for assembly systems with random component procurement times. *International Journal of Production Research*, 46(19), p. 5441-5467.
- Ould Louly M. -A., & Dolgui A. (2013). Optimal MRP parameters for a single item inventory with random replenishment lead time, POQ policy and service level constraint. *International Journal of Production Economics*, 143(1), p. 35-40.
- Prakash, P., Ceglarek, D., and Tiwari, M. (2012). Constraint-based simulated annealing (CBSA) approach to solve the disassembly scheduling problem. *The International Journal of Advanced Manufacturing Technology*, 60(9-12), 1125–1137.
- Proth J.-M., Mauroy G., Wardi Y., & Chu C. (1997). Supply management for cost minimization in assembly systems with random component yield times. *Journal of Intelligent Manufacturing*, 8(5), p. 358-403.
- Sakiani R., Ghomi S.F., & Zandieh, M. (2012). Multi-objective supply planning for two-level assembly systems with stochastic lead times. *Computers & Operations Research*, 39(7), p. 1325-1332.
- Shojaie, S. H., Bahoosh, A., & Pourhassan, M. (2015). A study on MRP with using leads time, order quality and service level over a single inventory. *Jurnal UMP Social Sciences and Technology Management*, 3(1), 235–239.
- Slama, I., Ben-Ammar, O., Dolgui, A., and Masmoudi, F. (2020a). Newsboy problem with two-level disassembly system and stochastic lead time. *ROADEF*.
- Slama, I., Ben-Ammar, O., Dolgui, A., and Masmoudi, F. (2020b). A stochastic model for a two-level disassembly lot-sizing problem under random lead time. *In IFIP International Conference on Advances in Production Management Systems*, pages 275–283. Springer.
- Slama, I., Ben-Ammar, O., Masmoudi, F., and Dolgui, A. (2019a). Disassembly scheduling problem: literature review and future research directions. *IFAC-PapersOnLine*, 52(13), 601–606.
- Slama, I., Ben-Ammar, O., Masmoudi, F., and Dolgui, A. (2019b). Scenario-based stochastic linear programming model for multi-period disassembly lot-sizing problems under random lead time. *IFAC-PapersOnLine*, 52(13), 595–600.
- Slama, I., Ben-Ammar, O., Dolgui, A., and Masmoudi, F. (2020c). New mixed integer approach to solve a multi-level capacitated disassembly lot-sizing problem with defective items and backlogging. *Journal of Manufacturing Systems*, 56 :50–57.
- Suzanne, E., Absi, N., & Borodin, V. (2020). Towards Circular Economy in Production Planning: Challenges and Opportunities. *European Journal of Operational Research*.
- Snyder, L.V., Atan, Z., Peng, P., Rong, Y., Schmitt, A.J., & Sinoysal B. (2016). OR/MS models for supply chain disruptions: a review. *IIE Transactions*, 48, p. 89-109.
- Speier, C., Whipple, J.M., Closs, D.J., & Voss, M.D. (2011). Global supply chain design considerations: Mitigating product safety and security risks. *Journal of Operations Management*, 29, 721–736.
- Tang, O., & Grubbström, R. W. (2003). The detailed coordination problem in a two-level assembly system with stochastic lead times. *International journal of production economics*, 81, 415-429.
- Tighazoui, A., Turki, S., Sauvey, C., & Sauer, N. (2019). Optimal design of a manufacturing-remanufacturing-transport system within a reverse logistics chain. *The International Journal of Advanced Manufacturing Technology*, 101(5-8), 1773-1791.
- Turki, S., Sauvey, C., & Rezg, N. (2018). Modelling and optimization of a manufacturing/remanufacturing system with storage facility under carbon cap and trade policy. *Journal of Cleaner Production*, 193, 441-458.
- Yano, C. A. (1987a). Setting planned lead times in serial production systems with tardiness costs. *Management Science*, 676 33, 95–106.
- Yano, C. A. (1987b). Setting planned lead times in serial production systems with tardiness costs. *Management Science*, 678 33, 95–106.
- Yano, C. A. (1987c). Stochastic lead times in two-level assembly systems. *IIE Transactions*, 19, 371–378.

PLANIFICATION DYNAMIQUE MULTILIGNES ET MULTIGAMMES AVEC MATIÈRES HÉTÉROGÈNES : UN MODÈLE DE BLENDING APPLIQUÉ À UNE CHAÎNE LOGISTIQUE MINIÈRE

M. BAMOUMEN^{a,b}, V. HOVELAQUE^{a,b}, V. GIARD^{b,c}

^a Univ-Rennes, CNRS, CREM - UMR 6211, F-35000 Rennes, France

^b Paris-Dauphine, PSL Research University, 75016, Paris, France

^c Mines-Université Mohammed VI Polytechnique, Benguerir, Maroc

Mouna.Bamoumen@emines.um6p.ma, Vincent.Hovelaque@univ-rennes1.fr, Vincent.Giard@dauphine.psl.eu

Résumé – La mission du site minier de Khouribga (Groupe OCP) est de constituer des produits intermédiaires ou finis de qualité chimique imposée (outputs), à partir de minerais de qualité hétérogène dans le temps et dans l'espace (multi-gisements), en utilisant un système de transformation multi-lignes constitué de cinq laveries. Ce défi est abordé à travers deux processus séquentiels : un traitement mécanique de lavage et un mélange de minerais. Afin de répondre aux demandes internes (usines chimiques) et externes (export), notre modèle est basé sur deux configurations alternatives : (1) l'utilisation de toutes les lignes pour constituer un mélange à destination des usines, (2) le fractionnement en deux sous-ensembles de lignes afin de constituer un mélange pour les usines et un mélange pour l'export. Le modèle mathématique, vise à définir le programme synchronisé de prélèvements des minerais-source couplés aux gammes à utiliser par laverie, pour satisfaire un programme d'expédition de lots à travers le pipeline principal, en tenant compte des temps d'acheminement et des disponibilités des minerais. L'exemple illustratif confirme l'intérêt de cette approche dans un objectif de satisfaction des clients (quantité, qualité, délai) en utilisant au mieux les ressources disponibles.

Mots clés - Blending dynamique, gammes alternatives, multi-lignes, intrants hétérogènes, planification.

Keywords – **blending, planification, alternative ranges, heterogeneous inputs.**

1 INTRODUCTION

Ce papier complète une première analyse de la planification de la production du site de Khouribga de l'OCP, principale entreprise mondiale d'extraction du phosphate (Bamoumen *et al*, 2020). Ce site, décrit Figure 1, a comme principal objectif d'alimenter les usines chimiques de l'OCP (client interne) et de satisfaire les demandes à l'exportation (clients externes) par des lots de produits (Outputs externes ou internes) fabriqués à partir de minerais (Inputs) et envoyés en continu par un pipeline long de 187 kms (*slurry pipeline*) à partir d'une station de tête (ST) vers une station terminale (SL). Chaque output doit respecter une charte qualité (CQ) à travers des contraintes de composition, de type [Min-Max], portant sur un ensemble de composants chimiques. Les minerais extraits, qualifiés ici d'Inputs Primaires (IP), sont caractérisés par ces mêmes composants, mais vérifient rarement les contraintes de composition des outputs. La solution classique d'un *blending* à sec d'IP, avant éventuel traitement complémentaire en laverie, pour obtenir un output conforme (Azzamouri *et al*, 2020), n'est techniquement pas possible sur ce site de Khouribga, qui comporte cinq laveries et où une laverie est alimentée par une séquence de lots d'IP. Le passage en laverie d'un lot d'IP produit un lot d'input intermédiaire (II) de caractéristiques volumiques et de composition qui dépend du choix de l'une des trois gammes de lavage possibles. Dans un premier modèle (Bamoumen *et al*, 2020), nous avons traité la question de

l'alimentation de la ST pour le principal produit demandé par les usines chimiques (output interne) en déterminant l'alimentation simultanée des cuves à partir de cinq flux, chacun provenant d'une laverie, et en assurant que le mélange entrant dans le *slurry pipeline* (en sortie des cuves) vérifiait la CQ. Dans cet article, nous proposons de traiter deux configurations possibles : un traitement mono production (output interne) mobilisant les cinq laveries (sur la base du travail réalisé dans Bamoumen *et al*, 2020) et un traitement bi-production permettant de combiner la fabrication simultanée de deux outputs, l'un interne et l'autre externe à partir de deux groupes de laveries prédéterminés. Le fait de travailler sur un sous-ensemble de laveries restreint fortement les combinaisons possibles satisfaisant la CQ des deux outputs, sachant que l'approvisionnement continu des usines chimiques avec un output unique est une contrainte forte imposée par l'industriel. En raison de la différence entre le débit de production et celui d'alimentation des usines chimiques, les deux configurations de production génèrent soit une saturation des cuves de la SL par l'output interne (mono production), soit un désamorçage lié aux cuves vides de l'output interne. Pour résoudre ce problème, notre modèle permet (1) de garantir une alimentation continue des usines chimiques avec un output interne entrant dans le *slurry pipeline* qui vérifie la CQ, et (2) de satisfaire la charte qualité de l'autre output (lié à un client externe) une fois que tous les lots d'II répondant à une commande sont arrivés aux cuves de la SL situées en aval du pipeline.

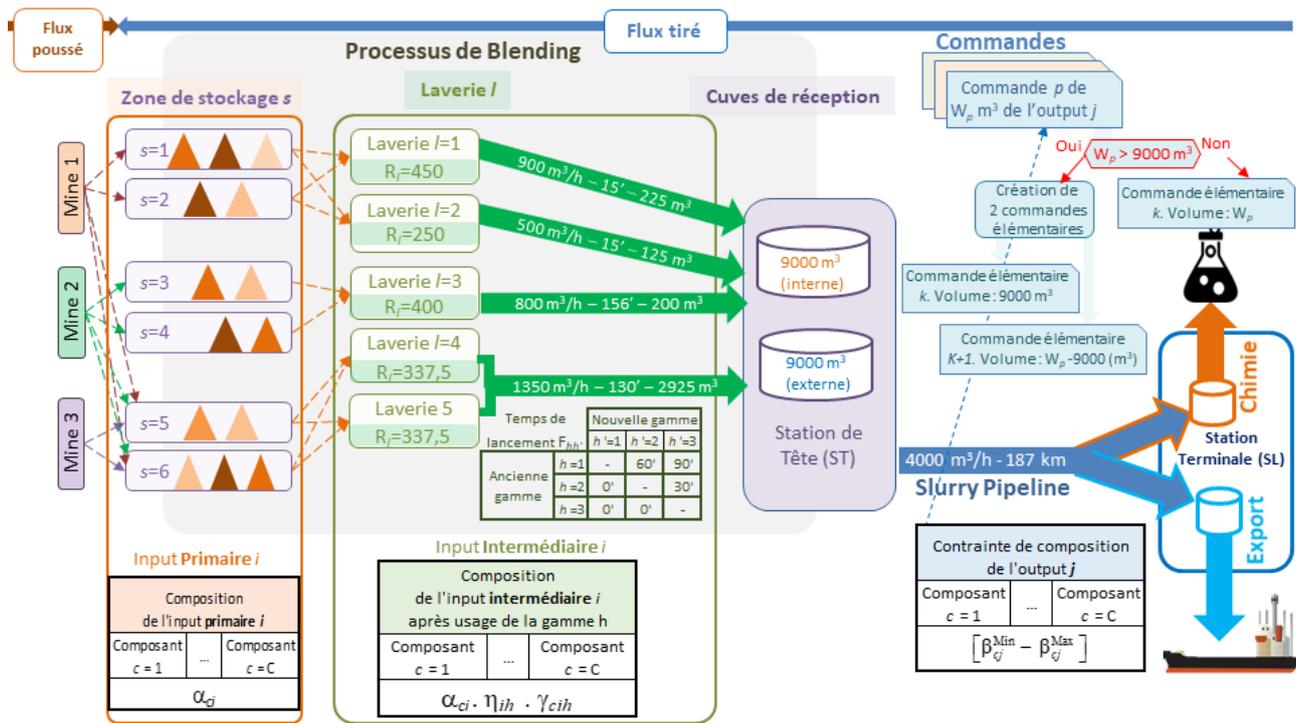


Figure 1 : Schéma de flux de la chaîne logistique étudiée

Le problème de blending étudié dans ce papier reprend les principales caractéristiques exprimées dans (Bamoumen *et al*, 2020), à savoir : *i*) Les laveries traitent successivement des lots d'IP (pas de mélange en entrée de la laverie) ; *ii*) plusieurs gammes existent pour obtenir un lot d'II ; elles impactent à la fois la composition de l'II et la réduction de poids en sortie de laverie ; *iii*) le changement de gamme peut s'accompagner d'un temps de lancement qui stoppe la production de la laverie ; *iv*) les pipes d'alimentation de la ST contiennent des quantités d'II dépendant du débit et de l'éloignement de la laverie ; en conséquence, le programme de production d'II en laverie doit tenir compte du temps de transport pour assurer une alimentation synchronisée de la station de tête par les lots

d'II en provenance des laveries ; *v*) la production étant continue, il faut tenir compte de résidus présents dans les laveries, dans les pipes secondaires et dans la station de tête (voir Figure 1) lors du lancement en fabrication d'un nouvel II ; *vi*) les disponibilités et caractéristiques chimiques des IP varient au cours du temps en fonction des prélèvements et de l'extraction minière (à flux poussés).

En tenant compte de l'ensemble de ces contraintes, l'aval détermine le programme d'expédition de la ST vers le *slurry pipeline* ainsi que le programme de production affecté aux lignes (voir Figure 2), le problème à traiter étant celui du choix des IP et des gammes permettant de respecter la CQ.

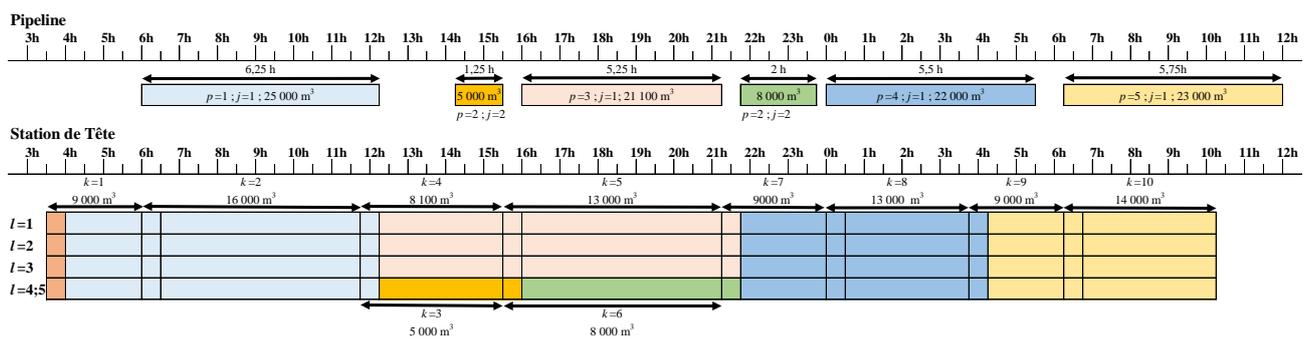


Figure 2 : Programme d'expédition et de production

Ce problème original de blending est formulé par un programme mathématique non linéaire en nombres entiers, visant à déterminer simultanément les quantités d'IP à prélever et leurs gammes de traitement en laverie, en tenant compte de l'ensemble des caractéristiques citées précédemment (§2). Une application numérique est fournie en §3, avant une brève conclusion.

2 DESCRIPTION DU PROBLEME ET FORMULATION MATHEMATIQUE

Le problème consiste à spécifier la production de P commandes exécutées sur L=5 laveries, à partir d'un ensemble de I inputs. Chaque commande p est caractérisée par un output à fabriquer parmi J outputs et un volume D_p (exprimé en m³).

Les laveries peuvent concourir simultanément à la production d'une commande ou être partagées en deux sous-ensembles, travaillant en parallèle sur une même durée, pour traiter simultanément deux commandes. Pour la commande interne, le mélange obtenu dans la ST respecte la CQ. Pour la commande externe, le respect de la CQ est assuré en aval, lorsque tous les lots d'II sont arrivés dans la cuve accueillant la production destinée à l'export. Dans le cas d'une mono production le volume de la première commande élémentaire correspond au remplissage d'une cuve de la ST (9000 m³) par des II provenant simultanément de l'ensemble des laveries, sans vidage préalable de la cuve ; la seconde commande élémentaire, d'un volume $D_p - 9000$ est constituée dès le début du vidage du lot précédent dans le *slurry pipeline*, le vidage et le remplissage pouvant donc s'effectuer simultanément pour ce second lot élémentaire, puisque la contrainte de satisfaction de la charte qualité est vérifiée dès l'entrée du mélange dans la ST. Dans le cas d'une bi production (un sous-ensemble de laveries affectées à une commande interne et l'autre à une commande export), le volume d'une commande élémentaire interne est contraint par la durée de production d'une commande élémentaire externe et par l'intervalle de temps nécessaire pour nettoyer le *slurry pipe* par de l'eau entre deux expéditions successives. Dans le cas d'une bi-production, le volume de la première commande élémentaire interne peut ne pas atteindre 9000 m³, et dépend du volume produit par les laveries mobilisées pendant la production d'une commande élémentaire externe. La satisfaction d'une commande externe passe par une scission de cette commande en plusieurs lots d'un volume inférieur à 9000 m³ mais obligatoirement supérieur à 4000 m³. Ces lots (ordonnés sur plusieurs périodes disjointes) alimentent in fine par le *slurry pipeline* l'une des deux cuves de la station terminale. Comme énoncé précédemment, c'est le mélange de ces lots successifs qui assure la CQ au niveau de ces cuves. Ainsi, chaque lot est associé à un sous-ensemble prédéterminé de laveries, les laveries restantes étant affectées à la satisfaction d'un lot d'un output interne. Dans cette configuration dite de bi-production, chacune des deux commandes ne génère qu'une commande élémentaire. Dans les deux configurations, la première commande élémentaire va progressivement remplir une cuve vide de la ST en commençant par accueillir les résidus des II précédents présents dans les laveries affectées à celle-ci jusqu'à saturation des cuves de la ST ou à la fin de la production demandée, puis les nouveaux II. Pour la seconde commande élémentaire liée à une mono production (donc à un output interne), vidage et remplissage peuvent être simultanés, les contraintes de composition étant respectées dès le mélange effectué à l'entrée de la cuve.

En amont de chaque ligne de laverie se trouve un sous-ensemble de I IP, répartis dans l'une des S=6 zones de stockage selon le gisement d'où il a été extrait. Pour la constitution d'une commande élémentaire, chaque laverie est alimentée par un seul IP, donnant lieu à un II à l'aide d'une seule gamme h , parmi $H=3$ gammes possibles. Chaque gamme engendre (i) des pertes de matières suivant un coefficient de réduction propre à cette gamme, et (ii) une augmentation ou diminution de la concentration des composants chimiques de l'IP traité, suivant des coefficients de distorsion relatifs à cette gamme. Le traitement d'un IP dans une laverie dure environ 30', quelle que soit la laverie. En conséquence, lors du changement d'IP alimentant la laverie l , sa cuve contient un lot de l'IP précédent, de volume R_l , équivalent à un résidu d'une demi-heure

de production de la laverie en question (donc dépendant du débit de cette laverie).

Le changement de gamme sur une laverie peut donner lieu à un temps de lancement pendant lequel l'IP précédent séjourne dans les installations avant reprise de l'alimentation (voir Figure 1). Ces temps de lancement sont pris en compte dans la distribution des volumes à produire sur chaque laverie, afin de maintenir une synchronisation de production de la commande élémentaire en question. Ainsi, si une laverie est momentanément neutralisée en raison d'un temps de lancement, la production équivalente à ce temps est répartie sur les autres laveries.

Par ailleurs, le critère d'optimisation retenu dans le modèle vise à minimiser les coûts de production des outputs (extraction et traitements en laverie) tout en cherchant à stabiliser la composition des outputs commandés par le client interne. Pour ce dernier, la composition d'une commande interne doit être la plus proche d'une structure cible pour éviter des coûts de réglages élevés (difficilement quantifiables actuellement), ce qui conduit à introduire dans le critère d'optimisation la valorisation de la déviance par un coût de pénalité arbitraire qui permet d'orienter la solution retenue dans l'ensemble des solutions respectant les contraintes du problème (voir Bara et al. 2020) ; nous discuterons au §3.2 l'impact du choix de cette pénalité.

Le modèle détermine pour chaque laverie le couple « quantité d'IP à traiter / gamme à utiliser » pour obtenir le lot de l'II utilisé pour fabriquer une commande élémentaire, en tenant compte des impératifs de synchronisation et d'éventuels temps de lancements.

2.1 Notations

Les poids sont exprimés en tonne, les volumes en m³ et les coûts en euro, à partir de coûts conventionnels pour des raisons évidentes de confidentialité. La conversion de la tonne en m³ est assurée par un coefficient noté θ , qui exprime le produit de la densité (en tonne/m³) par un indicateur nommé taux de solide (en %), exprimant la concentration du produit dans l'eau.

Indices	
i	indice de l'input ($i = 1 \dots I$).
c	indice du composant chimique ($c = 1 \dots C$).
s	indice de la zone de stockage des inputs ($s = 1 \dots S$)
h	indice de la gamme ($h = 1 \dots H$).
l	indice de la laverie ($l = 1 \dots L$).
p	indice de la commande à fabriquer ($p = 1 \dots P$).
k	indice de la commande élémentaire à fabriquer ($k = 0 \dots K$). La commande fictive $k = 0$ repère la dernière commande produite juste avant le début du programme de production, en postulant que toutes les laveries sont mobilisées.
j	indice de l'output ($j = 1 \dots J$).
n	Indice d'opérateur ($n = 1 \dots N$)
Paramètres	
θ	coefficient de conversion du poids (tonne) en volume (m ³).
ρ_l	minimum entre les débits de production et de transfert pour une laverie l .
δ_l	numéro de l'IP présent dans la laverie l au début de

	la fabrication de la première commande élémentaire.
δ_l	numéro de la gamme utilisée dans la laverie l avant le début de la fabrication de la première commande élémentaire.
Π_{sl}	booléen valant 1 si la zone de stockage s alimente la laverie l , et 0 sinon.
Ψ_i	numéro de la zone de stockage de l'IP i .
S_{i0}	quantité de l'IP i disponible au stock amont au début de l'horizon considéré (tonne).
α_{ci}	proportion du composant c dans le poids de l'IP i .
η_{ih}	coefficient de réduction de l'IP i par la gamme h . Si la gamme h n'est pas utilisable pour traiter l'IP i , alors $\eta_{ih} = 0$.
γ_{cih}	coefficient de distorsion du composant c de l'IP i par la gamme h .
R_l	volume résiduel dans la laverie l au début de la fabrication d'une commande élémentaire.
λ_k	numéro d'output de la commande élémentaire k ($j = \lambda_k$).
E_k	numéro de commande à laquelle se rattache la commande élémentaire k .
D_k	volume de la commande élémentaire k à fabriquer.
G_{E_k}	booléen valant 1 si la commande $p = E_k$ émane d'un client interne, 0 sinon.
$\beta_{c\lambda_k}^{\text{Min}}$	borne inférieure de la proportion (%) du composant c dans le poids de l'output $j = \lambda_k$.
$\beta_{c\lambda_k}^{\text{Max}}$	borne supérieure de la proportion (%) du composant c dans le poids de l'output $j = \lambda_k$.
$\tau_{c\lambda_k}$	valeur cible de la proportion (%) du composant c dans le poids de l'output $j = \lambda_k$. Cette valeur cible n'est utilisée que si l'output j est commandé par un client interne ($G_{E_k} = 1$).
B_{kl}	booléen valant 1 si la commande élémentaire k fait appel à la laverie l , 0 sinon.
A_k	booléen valant 1 si la commande élémentaire k mobilise toutes les laveries disponibles, 0 sinon.
W_{kl}	numéro de la commande qui précède la commande k dans la séquence des commandes exécutées sur la ligne l .
$F_{hh'}$	temps de lancement lors du passage d'une gamme h à une gamme h' .
σ_c	coût unitaire (par m^3) de la pénalité associée à l'écart absolu entre la quantité cible souhaitée $\tau_{c\lambda_k}$ du composant c de l'output λ_k et celle obtenue par le mélange produisant cet output.
χ	coût d'extraction d'une tonne d'IP.
ϕ_h	coût de traitement d'une tonne d'IP par la gamme h .
Γ_n	Expression d'un prédicat spécifique n (voir ci-après)
Variables	
ω_{ihlk}	variable de décision - binaire valant 1 si l'IP i est traité par la gamme h sur la laverie l pour produire la commande élémentaire k , 0 sinon.
y_{ilk}	variable intermédiaire - binaire valant 1 s'il y a changement d'IP à traiter dans la laverie l entre les commandes élémentaires $k-1$ et k , 0 sinon.
$z_{hh'lk}$	variable intermédiaire - binaire valant 1 s'il y a changement d'une gamme h à la gamme $h' \neq h$ sur

	la laverie l entre les commandes élémentaires $k-1$ et k , générant un temps de changement positif, 0 sinon.
S_{ik}	variable intermédiaire - stock de l'IP i (tonne) à la fin de fabrication de la commande élémentaire k , dans la zone de stockage Ψ_i (stock secondaire).
T_k	variable intermédiaire - durée de production du lot élémentaire k sur les laveries.
v_{kl}	variable intermédiaire - volume à produire sur la laverie l pour la constitution de la commande élémentaire k .
μ_{clk}	variable intermédiaire - poids du composant c dans le volume v_{kl} produit par la laverie l pour la constitution de la commande élémentaire k .
Δ_{ck}	variable intermédiaire - écart absolu du composant c de sa valeur cible dans la commande élémentaire k , passée par un client interne ($G_{E_k} = 1$).

2.2 Description des prédicats

La notion de prédicat permet d'établir des restrictions sur l'existence de certaines variables, et a donc l'avantage de limiter drastiquement le nombre de variables dans le modèle d'optimisation.

Une commande élémentaire k possède quatre caractéristiques : (i) elle porte sur l'output $j = \lambda_k$; (ii) rattachée à la commande $p = E_k$, elle répond à un client interne si $G_{E_k} = 1$ et un client externe si $G_{E_k} = 0$; (iii) elle ne peut être exécutée que sur les lignes l telles que $B_{kl} = 1$; (iv) elle est produite par toutes les laveries disponibles si $A_k = 1$ (mono production); dans ce cas d'une bi-production ($A_k = 0$), il n'est pas possible d'avoir de changement de gamme impliquant un temps de lancement positif. Afin de faciliter la lecture des relations mathématiques de la section suivante, les prédicats seront notés sous la forme Γ_n , où n est le numéro du prédicat.

Ainsi, l'existence de la variable ω_{ihlk} est conditionnée par le prédicat $\Gamma_1 \equiv (B_{kl} \cdot \Pi_{\Psi_i} \cdot \eta_{ih} = 1)$, ce produit de paramètre remplaçant $B_{kl} = 1 \wedge \Pi_{\Psi_i} = 1 \wedge \eta_{ih} = 1$. En effet, la commande élémentaire k ne peut charger la ligne l que si $B_{kl} = 1$, ne peut mobiliser l'input i que s'il est utilisable par la ligne l , soit $\Pi_{\Psi_i} = 1$, et que si la gamme h est utilisable pour traiter l'input i , soit $\eta_{ih} = 1$. Pour des raisons équivalentes, l'existence de la variable y_{ilk} est conditionnée par le prédicat $\Gamma_2 \equiv (B_{kl} \cdot \Pi_{\Psi_i} = 1)$.

L'existence de la variable $z_{hh'lk}$ est conditionnée par le prédicat $\Gamma_3 \equiv (k \cdot F_{hh'} \cdot B_{kl} \cdot A_k > 0)$ qui simplifie l'expression $k > 0 \wedge F_{hh'} > 0 \wedge B_{kl} = 1 \wedge A_k = 1$. Cette variable étant définie pour la relation entre les commandes élémentaires $k-1$ et k , elle ne peut exister pour $k = 0$. De plus, le changement de gamme à temps non nul ne peut exister que si $F_{hh'} > 0$ et si $A_k = 1$ (hypothèse sur le non-changement de gamme à temps positif dans le cas d'une bi-production). Le booléen B_{kl} exprime enfin la possibilité d'utiliser la laverie l pour exécuter la commande élémentaire k .

La pénalité appliquée à une déviance de produit fabriqué par rapport à une cible prédéterminée ne s'applique que pour les produits liés à des commandes internes. Ainsi, l'existence de la variable Δ_{ck} est conditionnée par le prédicat $\Gamma_4 \equiv (G_{E_k} = 1)$.

Le problème du blending sur le site de Khouribga peut être modélisé comme suit :

$$\text{Min} \left(\begin{array}{l} \sum_{ck|\Gamma_4} \sigma_c \cdot \Delta_{ck} \\ + \sum_{ihlk|\Gamma_1} (\chi + \phi_h) \cdot \omega_{ihlk} \cdot v_{kl} \cdot (1 - \eta_{ih}) \cdot (\theta / \eta_{ih}) \end{array} \right) \quad (1)$$

Sous contraintes :

$$\sum_{h,i} \omega_{ihlk|\Gamma_1} = 1; \forall l, k | k > 0 \quad (2)$$

$$\begin{cases} \sum_{i,h,l | l < 3 \wedge \psi_i = s \wedge \Gamma_1} \omega_{ihlk} \leq 1, \forall k, s | s \in \{1, 2\} \wedge k > 0 \\ \sum_{i,h,l | l > 3 \wedge \psi_i = s \wedge \Gamma_1} \omega_{ihlk} \leq 1, \forall k, s | s \in \{5, 6\} \wedge k > 0 \end{cases} \quad (3)$$

$$\begin{cases} z_{hh'lk} \geq \left(\sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ih'lk} + \sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}} \right) - 1 \\ z_{hh'lk} \leq \frac{\sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ih'lk} + \sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}}}{2} \end{cases} \quad \forall h, h', l, k | \Gamma_3 \quad (4)$$

$$\begin{cases} \sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ih'lk} + \sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}} \leq 1 \\ \forall h, h', l, k | k > 0 \wedge F_{hh'} > 0 \wedge B_{kl} = 1 \wedge A_k = 0 \end{cases} \quad (4')$$

$$T_k = (D_k + \sum_{h,h',l | F_{hh'} > 0 \wedge B_{kl} = 1 \wedge A_k = 1} F_{hh'} \cdot z_{hh'lk} \cdot \rho_l) / \sum_l \rho_l \cdot B_{kl}, \quad \forall k | k > 0 \quad (5)$$

$$v_{kl} = (T_k - \sum_{h,h',l | F_{hh'} > 0 \wedge A_k = 1} F_{hh'} \cdot z_{hh'lk}) \cdot \rho_l, \quad \forall l, k | k > 0 \wedge B_{kl} = 1 \quad (6)$$

$$S_{ik} = S_{i0} - \sum_{h,l,k | k' \leq k \wedge \Pi_{\psi_i} \cdot \eta_{ih} > 0} v_{k'l} \cdot \omega_{ihlk} \cdot \theta / \eta_{ih}; \forall i, k | k > 0 \quad (7)$$

$$S_{ik} \geq 0; \forall i, k | k > 0 \quad (8)$$

$$\begin{cases} y_{ilk} \geq \frac{\sum_{h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlk} - \sum_{h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}}}{2} - \frac{1}{4} \\ y_{ilk} \leq \frac{\sum_{h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlk} - \sum_{h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}}}{2} + \frac{1}{2} \end{cases} \quad (9)$$

$$\begin{cases} y_{ilk} \geq 0 \\ \forall i, l, k | k > 0 \wedge \Gamma_2 \end{cases}$$

$$\mu_{clk} = \left(\begin{array}{l} \sum_{i,h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} R_l \cdot \omega_{ihlW_{kl}} \cdot \alpha_{ci} \cdot \gamma_{cih} \\ + \sum_{i,h|\Pi_{\psi_i} \cdot \eta_{ih} > 0} (v_{kl} - R_l) \cdot \omega_{ihlk} \cdot \alpha_{ci} \cdot \gamma_{cih} \end{array} \right) \quad \forall c, l, k | k > 0 \wedge B_{kl} = 1 \quad (10)$$

$$\beta_{c\lambda_k}^{\text{Min}} \cdot \sum_l v_{kl} \leq \sum_{l|B_{kl}=1} \mu_{clk} \leq \beta_{c\lambda_k}^{\text{Max}} \cdot \sum_l v_{kl} \quad \forall c, k | k > 0 \wedge c < 4 \wedge \Gamma_4 \quad (11)$$

$$\beta_{c\lambda_k}^{\text{Min}} \cdot \sum_{l|B_{kl}=1 \wedge c=1} \mu_{clk} \leq 218,5 \cdot \sum_{l|B_{kl}=1 \wedge c=4} \mu_{clk} \leq \beta_{c\lambda_k}^{\text{Max}} \cdot \sum_{l|B_{kl}=1 \wedge c=1} \mu_{clk}; \quad \forall k | k > 0 \wedge \Gamma_4 \quad (11')$$

$$\beta_{c\lambda_k}^{\text{Min}} \cdot \sum_{k|E_k=p} D_k \leq \sum_{k,l|E_k=p \wedge B_{kl}=1} \mu_{clk} \leq \beta_{c\lambda_k}^{\text{Max}} \cdot \sum_{k|E_k=p} D_k \quad \forall c, p | G_p = 0 \quad (12)$$

$$\begin{cases} \Delta_{ck} \geq \sum_{l|B_{kl}=1} \mu_{clk} - \tau_{c\lambda_k} \cdot \sum_l v_{kl} \\ \Delta_{ck} \geq \tau_{c\lambda_k} \cdot \sum_l v_{kl} - \sum_{l|B_{kl}=1} \mu_{clk} \end{cases}; \forall c, k | k > 0 \wedge \Gamma_4 \quad (13)$$

$$\begin{cases} \omega_{i,h=2,l=5,k} \leq 1 - \sum_{h>1} \omega_{ih,l=4,k}; \forall i, k \\ |k > 0 \wedge (\Pi_{\psi_i,l=4} \cdot \eta_{ih=2} > 0 \vee \Pi_{\psi_i,l=4} \cdot \eta_{ih=3} > 0) \wedge \Pi_{\psi_i,l=5} \cdot \eta_{ih=2} > 0 \end{cases} \quad (14)$$

La fonction-objectif (1) est une somme (i) du coût de la stabilité de la composition des commandes élémentaires k autour des valeurs cibles fixées pour les composants considérés ; ce coût est propre à chaque composant chimique ; (ii) du coût de perte de matière liée à l'utilisation des gammes ; ce coût se compose du coût d'une tonne d'IP extrait, noté χ , et du coût d'une tonne d'IP traitée par la gamme h , noté ϕ_h .

La relation (2) impose, pour la fabrication d'une commande élémentaire k , de traiter un seul IP i sur une laverie l avec une seule gamme h .

Les deux relations (3) permettent de respecter les contraintes de convoyage de telle sorte que chacune des deux laveries $l=1$ et $l=2$, et donc $l < 3$, soient alimentées par une seule zone parmi $s=1$ et $s=2$ (respectivement chacune des laveries $l=4$ et $l=5$, et donc $l > 3$, soient alimentées par une seule zone parmi $s=5$ et $s=6$).

Les deux relations (4) modélisent l'existence d'un changement de gamme, sur chaque laverie l , entre deux commandes élémentaires consécutives (dont le rang de la première est W_{kl} et celui de la suivante est k). La variable $\omega_{ihlk=0}$ est initialisée à 1 pour $i = \delta_i$ et $h = \delta'_i$ et à 0 dans les autres cas. Il y a passage de la gamme h à la gamme h' ($h \neq h'$) si (i) le changement de gamme avec temps de lancement positif est autorisé ($A_k = 1$),

et (ii) la somme $\sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ih'lk} + \sum_{i|\Pi_{\psi_i} \cdot \eta_{ih} > 0} \omega_{ihlW_{kl}}$ vaut 2. Si par contre cette somme vaut 1 ou 0, alors la laverie l opère les deux commandes successives avec la même gamme. Ainsi, la variable binaire $z_{hh'lk}$ vérifie les conditions suivantes :

$z_{hh'lk} = 1$ si la somme précédente vaut 2 et $z_{hh'lk} = 0$ si la somme précédente est < 2 .

Dans le cas d'une bi-production ($A_k = 0$), tout changement de gamme impliquant un temps de lancement positif est interdit pour assurer la synchronisation des deux commandes exécutées en parallèle dans le programme d'expédition des commandes déjà établi. En utilisant l'explication précédente, cette contrainte est modélisée dans la relation (4').

La relation (5) permet de déterminer la durée de production d'une commande élémentaire k , en répartissant le volume total D_k de celle-ci sur les laveries concernées par cette commande (telles que $B_{kl} = 1$), compte tenu d'éventuels temps de lancement $\sum_{h,h',l | F_{hh'} \cdot B_{kl} > 0} F_{hh'} \cdot z_{hh'lk} \cdot \rho_l$ dans le cas où toutes les

laveries sont mobilisées pour une même commande élémentaire ($A_k = 1$).

La relation (6) détermine le volume à produire sur la laverie l pour la fabrication de la commande élémentaire k , en tenant compte d'éventuels temps de lancement positifs dans le cas d'une mono-production

La relation (7) calcule le stock d'un IP i à la fin de la production d'une commande élémentaire k , en retranchant du stock initial, le cumul des volumes consommés de cet IP i ,

ramené à la tonne en multipliant par le coefficient θ/η_{ih} . Cette relation est non linéaire en raison du produit de deux variables $v_{kl} \cdot \omega_{ihlk}$. La variable v_{kl} étant intermédiaire, celle-ci peut être remplacée par son expression en fonction de $z_{ih'lk}$, faisant apparaître le produit de deux variables binaires pouvant être linéarisés¹.

La relation (8) impose de ne pas consommer plus d'IP que les quantités disponibles en stock.

Sur une laverie l , un passage de l'IP i' à l'IP i ($i \neq i'$), entre deux commandes élémentaires consécutives W_{kl} et k , a lieu quand $\sum_h \omega_{ihlk} - \sum_h \omega_{ihW_{kl}} = 1$ ou $\sum_h \omega_{i'hlk} - \sum_h \omega_{i'hlW_{kl}} = -1$.

En effet, l'expression $\sum_h \omega_{ihlk} - \sum_h \omega_{ihW_{kl}} = 1$ signifie que l'IP i n'est pas utilisé sur la laverie l pour la fabrication de la commande W_{kl} mais l'est pour la fabrication de la commande k (et inversement pour l'expression $\sum_h \omega_{i'hlk} - \sum_h \omega_{i'hlW_{kl}} = -1$). Si par contre cette différence vaut 0, il n'y a pas de changement d'IP sur la laverie l entre deux commandes consécutives. Pour ne pas comptabiliser le changement d'IP deux fois, la variable binaire y_{ilk} doit vérifier les conditions suivantes : si $\sum_h \omega_{ihlk} - \sum_h \omega_{ihW_{kl}} = 1$ alors $y_{ilk} = 1$; si $\sum_h \omega_{ihlk} - \sum_h \omega_{ihW_{kl}}$ vaut 0 ou -1 alors $y_{ilk} = 0$. Ces conditions sont être exprimées par les relations linéaires (9).

Les relations (10) calculent le poids μ_{clk} du composant c dans le volume v_{kl} produit par la laverie l pour la fabrication d'une commande k compte tenu du résidu des laversies résultant de la commande élémentaire précédente W_{kl} sur cette ligne. À noter que ces relations sont également non linéaires mais linéarisées comme précédemment.

Les deux relations suivantes ont pour but de vérifier à l'entrée de la station de tête, la charte qualité imposée pour le client interne ($G_{E_k} = 1$). La relation (11) impose le respect des bornes supérieures et inférieures de chaque composant $c < 4$. La contrainte technique imposée par l'industriel reliant les composants $c=1$ et $c=4$ est exprimée par la relation (11'). La relation (12) impose que le mélange obtenu en aval du pipeline (station terminale) vérifie la charte qualité pour les commandes destinées à l'export ($G_{E_k} = 0$).

Les relations (13) calculent les écarts absolus entre la composition de la commande élémentaire et la valeur cible des composants demandée par un client interne.

Enfin, la relation (14) impose le respect de conditions d'opérabilité des gammes sur certaines laversies. Dans notre cas, si la laverie $l=4$ utilise la gamme $h=2$, alors la laverie $l=5$ ne peut opérer qu'avec la gamme $h=1$.

3 ILLUSTRATION

Dans la présente section, nous présentons dans une première partie les données d'entrée pour le test, suivie des résultats fournis par optimisation linéaire du modèle élaboré dans la

section §2, avant de finir avec la génération du programme horaire de production des lignes sur un horizon temporel.

3.1 Données du test

Le modèle mathématique linéaire a été testé avec des données proches de celles employées par le Groupe OCP dans la planification de sa production du site minier de Khouribga (les données structurelles sont réelles et les données concernant l'état initial sont proches de la réalité). Dans ce test, 18 IP répartis sur les 6 zones de stockage ont été retenus. Pour chaque IP sont connus les caractéristiques chimiques, les coefficients de réduction (Tableau 1, voir annexe) et les coefficients de distorsion des composants selon les trois gammes possibles.

L'étude a été réalisée sur deux outputs qui sont actuellement expédiés par le *slurry pipeline* ; l'output $j=1$ est commandé par un client interne. Les contraintes de composition de ces outputs pour les quatre composants chimiques retenus sont fournies dans le Tableau 2.

Dans l'exemple retenu, cinq commandes sont programmées sur un horizon d'une journée environ (Figure 2) et décomposées en deux commandes élémentaires, soit dix commandes élémentaires à fabriquer au total (Tableau 3). À noter que les commandes $p=2$ et $p=3$ sont exécutées en parallèle. Le Tableau 4 fourni l'affectation des laversies à l'ensemble des commandes élémentaires ($B_{kl} = 1$), ainsi que la commande antécédente (W_{kl}) pour chaque commande élémentaire exécutée sur la laverie qui lui est affectée.

Les coûts de la fonction-objectif sont factices pour des raisons évidentes de confidentialité. Il convient de distinguer (Bara et al, 2020) *i*) ceux qui correspondent à une pénalité destinée à orienter la solution optimale (sans impact sur le respect des contraintes) pour prendre en compte l'impact de la solution sur la performance des lignes de production d'acide ; on a retenu un scénario central avec une pénalité identique pour tous les composants ($\sigma_c = \sigma = 100$), que l'on a complété par l'étude de scénarios dans lesquels la pénalité varie de 0 à 10 000 ; *ii*) des coûts-standards qui reflètent des charges réelles observées dans le compte de résultats : le coût unitaire d'extraction ($\chi = 3$) et celui de traitement en laverie, lié à la gamme utilisée ($\varphi_{h=1} = 2$; $\varphi_{h=2} = 4$; $\varphi_{h=3} = 5$). Dans l'analyse des résultats, on dissociera la partie de charges réelles de celle, fictive pour les laversies, de la pénalité visant à respecter les préférences du maillon aval de la chaîne logistique.

3.2 Solution obtenue

Nous présentons ici les résultats d'optimisation obtenus avec le modèleur Xpress. Le temps de résolution est de moins de dix minutes pour une pénalité de stabilité nulle avec une génération de 5717 variables et 10069 contraintes ; et environ dix-huit minutes pour $\sigma_c = \sigma = 100$, avec 5749 variables et 10102 contraintes générées (qui restent stables pour les pénalités positives). Ces temps de résolution sont obtenus sur un serveur d'une vitesse d'exécution de 3.5 GHz et d'une mémoire installée de 64 Go.

Deux scénarios A et B sont étudiés selon la quantité initialement disponible des IP (ligne 2 et 3 du Tableau 1). Pour les deux scénarios, le Tableau 5 présente, pour chaque laverie l , les IP à prélever et leur zone de stockage s , la quantité de prélèvement (en tonne) et la gamme h à appliquer pour la fabrication de chacun des huit commandes élémentaires.

¹ La linéarisation est obtenue par la propriété suivante : soient x , w et z trois variables binaires. Alors l'expression $z = x \cdot w$ est équivalente aux deux inégalités suivantes : $z \geq x + w - 1$ et $z \leq \frac{x + w}{2}$

Lorsque les IP sont présents en quantité suffisante (scénario A), on remarque que les trois commandes portant sur l'output $j=1$ et faisant appel à l'ensemble des laveries, sont fabriquées à partir des mêmes mélanges d'IP, avec deux changements de gammes entre les commandes élémentaires $k=1$ et $k=2$ (on passe de $h=1$ à $h=2$ sur la laverie $l=1$) ; et entre $k=3$ et $k=6$ (on passe de $h=2$ à $h=1$ sur la laverie $l=4$). En limitant la quantité des IP les plus consommés (scénario B), le modèle de blending propose différents mélanges d'IP pour les mêmes commandes mentionnées plus haut, engendrant une dégradation de la fonction-objectif (420 388 pour le scénario A contre 476 044 pour le scénario B). Cette dégradation est dû : 1) à l'augmentation de la quantité d'IP traitée par les laveries (chiffrée à 2% environ) par suite de l'utilisation de gammes engendrant un coût plus fort (lié à la consommation de ressources) que dans le scénario A (environ 11% d'augmentation de coût de production) ; 2) à l'utilisation d'IP ayant une composition causant l'augmentation de l'écart par rapport aux valeurs cibles fixées par les CQ des outputs de près de 17% par rapport au scénario A. Dans la suite de cette illustration, nous présentons les résultats obtenus pour le scénario B uniquement.

Le modèle permet aussi de déterminer, pour chaque commande élémentaire, le volume à produire sur chaque laverie (en colonnes du Tableau 6) compte tenu des temps de changement s'ils ont lieu (variable intermédiaire), ainsi que la composition chimique en sortie de chaque laverie.

Les cellules en rouges indiquent un non-respect de la borne supérieure ou inférieure pour le composant du lot de l'II produit par la laverie en colonne. Sur les quarante lots d'II produits pour la fabrication des dix commandes élémentaires, un quart seulement respecte totalement les spécificités de tous les composants en sortie de la laverie. Cela étant, par le mélange des flux provenant des laveries affectées, toutes les commandes élémentaires portant sur l'output interne $j=1$ respectent bien les caractéristiques demandées (Tableau 3). Pour la

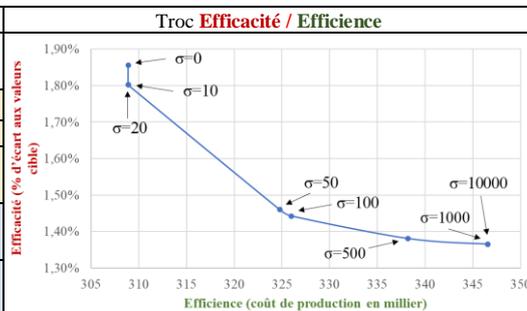
commande portant sur l'output externe $j=2$, on remarque que le respect des spécificités demandées en composition ne se fait qu'en mélangeant les deux commandes élémentaires $k=3$ et $k=6$ en aval du pipeline (dans une cuve de la SL). Cette approche fournit alors une meilleure flexibilité au niveau des laveries (IP à prélever et gamme à utiliser) tout en opérant à un niveau de performance optimal de composition et de coût de production.

En conséquence, comme la fonction-objectif proposée agrège un coût lié à la composition du produit (non mesurable) et un coût lié à la production (mesurable), le modèle permet de fournir des éléments de négociation aux deux maillons de la chaîne étudiée (le client qui est le complexe chimique et le fournisseur qui est le site minier). En modifiant le coût de la stabilité (fixé dans l'exemple numérique à $\sigma=100$ pour tous les composants chimiques c , i.e. $\sigma_c=100, \forall c \in C$) pour prendre successivement les valeurs $\{0, 10, 20, 50, 100, 500, 1000 \text{ et } 10000\}$, nous obtenons les différents coûts optimaux dans le Tableau 7.

L'introduction d'une pénalité génère une baisse allant jusqu'à 26% de l'écart en m^3 aux valeurs cibles par rapport à un coût nul de stabilité pour $\sigma \geq 500$ (cette baisse étant de 22% pour $\sigma=100$). Néanmoins, cette réduction de l'écart n'est possible que grâce à une augmentation de la quantité prélevée des IPs allant jusqu'à 2% (resp. 1%) par rapport à $\sigma=0$ pour $\sigma \geq 500$ (resp. $\sigma=100$), qui est équivalente à une augmentation du coût réel de production allant jusqu'à 12% (resp. 6%) par rapport à $\sigma=0$ pour $\sigma \geq 500$ (resp. $\sigma=100$). Cette illustration fournit ainsi les bases d'une négociation entre deux maillons d'une CL. En effet, la perte d'efficacité (augmentation du coût réel) en amont n'est acceptable à un niveau supérieur de supervision de la CL, qu'à condition que l'amélioration de l'efficacité (ici amélioration de la qualité) en aval se traduise par des économies au moins égales

Tableau 7 : Coûts de la fonction-objectif en fonction du coût unitaire de stabilité de la composition

Pénalité σ	0	10	20	50	100	500	1000	10000
Fonction-objectif (10^3)	309	328	346	401	476	1056	1767	14549
Ecart (m^3)	1930	1874	1874	1518	1500	1436	1420	1420
Ecart %	1,86%	1,80%	1,80%	1,46%	1,44%	1,38%	1,37%	1,37%
Coût stabilité (10^3)	0	19	37	76	150	718	1420	14203
Quantité IP utilisée (10^3 t)	140	140	140	141	141	143	144	144
Coût production total (10^3)	309	309	309	325	326	338	347	347



3.3 Programme horaire de production

En plus des données de sortie présentées plus haut, le modèle permet de calculer la durée de production de chaque commande élémentaire, qui est identique pour toutes les laveries. Cette durée est calculée comme le quotient du volume de la commande élémentaire considérée par rapport à la somme des débits de production des laveries (voir première ligne du Tableau 8). Pour le lot $k=1$ par exemple, cette durée est de $9000/3550 = 2,54$ h.

La première étape pour établir le programme horaire de production des laveries consiste à déterminer la date d'arrivée des commandes programmées à la ST. Connaissant la date d'expédition de chaque commande d'output (4^e ligne du Tableau 8), on retranche la durée de production de sa première

commande élémentaire. On obtient ainsi les dates d'arrivée à la ST (dernière ligne du Tableau 8).

Pour établir le programme horaire de production des laveries, il faut déterminer les dates de lancement d'OF pour la fabrication des commandes d'output programmées. Pour garantir la synchronisation de l'arrivée des flux à la ST, on retranche des dates d'arrivée à la ST, les durées de transport résultant de l'espacement entre laveries et la ST (Tableau 9).

Connaissant ainsi ces deux dates, on peut schématiser les arrivées des lots à la ST (Gantt milieu de la Figure 3) et les lancements d'ordres de fabrication sur les laveries (Gantt inférieure de la Figure 3). Comme expliqué précédemment, l'instauration d'un régime permanent nécessite une demi-heure environ lors de l'introduction d'un nouvel IP, c'est bien ce résidu qui arrive en premier à 3h30 pour $k=1$, suivi une

de mi-heure après de l'II choisi par le modèle. Les dates de lancements d'OF sur le Gantt inférieure sont décalées d'après

le Tableau 9, pour tenir compte des durées de transport par pipelines secondaires à la ST.

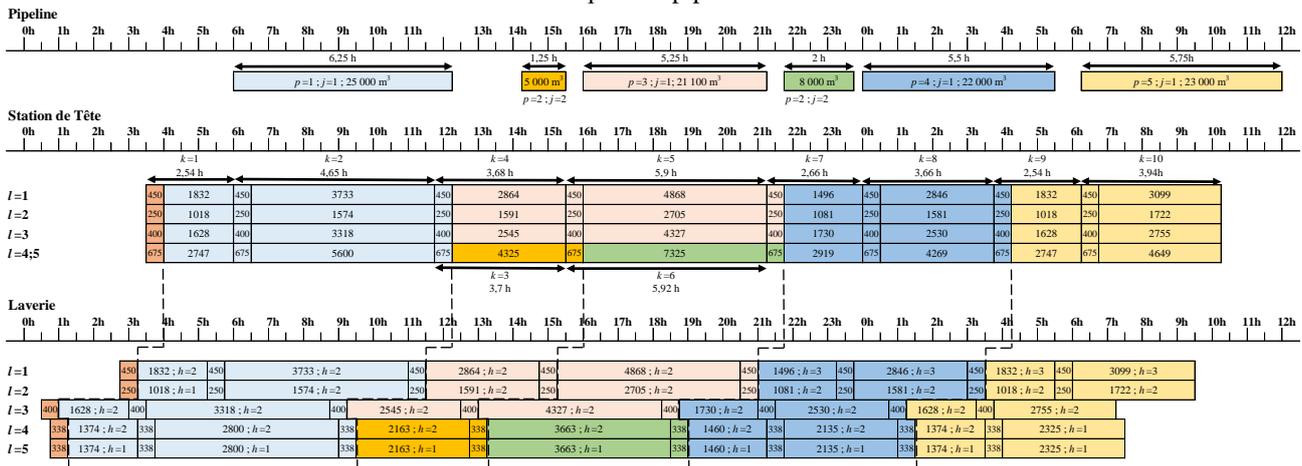


Figure 3 : Synchronisation des flux en sortie de laverie, en entrée de la ST et en sortie de la ST

4 CONCLUSION

L'organisation actuelle de la production des laveries est différente de celle proposée ici. Ce premier travail a trouvé un écho favorable à l'OCF et plusieurs autres études de cas seront indispensables pour convaincre les décideurs de tester cette approche novatrice. Ce modèle implique une disponibilité d'IP connue en début de production, ce qui est acceptable pour un programme d'une ou deux journées. Une extension de ce modèle est en cours, pour définir dynamiquement les alimentations des stocks d'IP, à partir de minerais dans les mines prêts à être expédiés, en fonction des commandes à exécuter. Cette extension permettra, d'une part, de travailler avec un programme glissant de production hebdomadaire et, d'autre part, de repousser à la mine (au niveau du stock de minerai prêt à partir), le point de découplage entre flux tirés et poussés.

REFERENCES

- Azzamouri, A., M. Bamoumen, H. Hilali, V. Hovelaque, V. Giard. 2020. "Flexibility of dynamic blending with alternative routings combined with security stocks: a new approach in a mining supply chain", *International Journal of operation research*. DOI: [10.1080/00207543.2020.1814443](https://doi.org/10.1080/00207543.2020.1814443)
- Bamoumen, M., V. Hovelaque, V. Giard. 2020. "Un modèle de blending pour la planification dynamique d'une chaîne logistique minière", MOSIM'20, Agadir, Maroc. <https://hal.archives-ouvertes.fr/hal-03005623v1>
- Bara, B., Gautier, F., Giard, V. 2020, "Usage des pénalités dans les modèles économiques de management industriel", MOSIM'20, Agadir, Maroc. <https://hal.archives-ouvertes.fr/hal-03005697v1>

ANNEXE

Tableau 1

Zone de stockage s	s=1			s=2			s=3			s=4			s=5			s=6			
	i=1	i=2	i=3	i=4	i=5	i=6	i=7	i=8	i=9	i=10	i=11	i=12	i=13	i=14	i=15	i=16	i=17	i=18	
Quantité initiale du scénario A (10 ³ tonne)	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	
Quantité initiale du scénario B (10 ³ tonne)	12	16	12	12	16	12	12	12	12	12	12	16	12	12	12	12	12	12	
Composition chimique	c=1	56.0%	##	55.0%	54.0%	61.0%	62.0%	61.4%	58.6%	57.0%	59.0%	54.4%	61.0%	64.5%	61.5%	60.3%	55.5%	60.5%	61.0%
	c=2	1.5%	##	1.7%	1.3%	1.3%	1.6%	1.0%	1.1%	1.3%	1.2%	1.2%	1.4%	0.7%	0.9%	1.1%	0.9%	0.7%	1.1%
	c=3	3.3%	##	3.5%	3.4%	2.9%	3.2%	4.7%	5.6%	3.9%	5.0%	5.6%	4.1%	4.6%	5.1%	4.4%	7.0%	5.7%	4.8%
	c=4	0.0020%	##	0.0022%	0.0018%	0.0017%	0.0019%	0.0021%	0.0022%	0.0023%	0.0022%	0.0025%	0.0021%	0.0022%	0.0020%	0.0020%	0.0021%	0.0018%	0.0020%
Coefficient de réduction	h=1	0.64	1	0.64	0.64	0.69	0.69	0.72	0.67	0.66	0.67	0.66	0.67	0.78	0.71	0.71	0.69	0.71	0.71
	h=2	0.58	1	0.58	0.58	0.63	0.63	0.71	0.66	0.65	0.66	0.65	0.66	0.75	0.63	0.63	0.6	0.63	0.63
	h=3	0.55	1	0.55	0.55	0.58	0.58	-	-	0.51	-	0.51	-	-	-	-	0.52	-	-

Tableau 2

Composant c	Caractéristiques de l'output j=1			Caractéristiques de l'output j=2		
	Valeur minimale β_{c1}^{Min}	Valeur maximale β_{c1}^{Max}	Valeur cible τ_{c1}	Valeur minimale β_{c2}^{Min}	Valeur maximale β_{c2}^{Max}	Valeur cible τ_{c2}
c=1 (P ₂ O ₅)	65.0%	100%	66.6%	65.0%	66.0%	65.0%
c=2	0%	0.85%	0.85%	-	-	-
c=3	-	-	-	2.2%	3.9%	3.2%
c=4	0 mg/P ₂ O ₅	65 mg/P ₂ O ₅	60 mg/P ₂ O ₅	0.0018%	0.0025%	0.0025%

Tableau 3

	Volume total des commandes d'output (m ³)									
	Output j=1	Output j=2	Output j=1							
	25,000	13,000	21,100	22,000	23,000	23,000	23,000	23,000	23,000	23,000
	Commande élémentaire k									
	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
E _k	1	1	2	3	3	2	4	4	5	5
W _k (m ³)	9,000	16,000	5,000	8,100	13,000	8,000	9,000	13,000	9,000	14,000
W _k · θ (tonne)	8,064	14,336	4,480	7,258	11,648	7,168	8,064	11,648	8,064	12,544

Tableau 4

		B_{ij}					W_{ij}					
		l=1	l=2	l=3	l=4	l=5	A_k	l=1	l=2	l=3	l=4	l=5
Commande élémentaire k	k=1	1	1	1	1	1	1	0	0	0	0	0
	k=2	1	1	1	1	1	1	1	1	1	1	1
	k=3	0	0	0	1	0	0	0	0	0	2	2
	k=4	1	1	1	0	0	0	2	2	2	0	0
	k=5	1	1	1	0	0	0	4	4	4	0	0
	k=6	0	0	0	1	1	0	0	0	0	3	3
	k=7	1	1	1	1	1	1	5	5	5	6	6
	k=8	1	1	1	1	1	1	7	7	7	7	7
	k=9	1	1	1	1	1	1	8	8	8	8	8
	k=10	1	1	1	1	1	1	9	9	9	9	9

Tableau 8

Commande p	p=1	p=2	p=3	p=4	p=5
Commande élémentaire k	k=1	k=2	k=3	k=4	k=5
Durée de production (heure)	2,54	4,65	3,70	5,93	3,68
TOTAL (heure)	7,18	9,63	9,59	2,66	6,32
Date d'expédition de la commande d'output	6h00	14h15	21h45	16h00	0h00
Date d'arrivée de la commande d'output à la ST	3h30	11h45	15h30	11h45	21h15

Tableau 9

		Laverie l				
		l=1	l=2	l=3	l=4	l=5
Date de lancement	k=1	3h15	3h15	0h55	1h20	
d'OF pour les commandes d'output	k=3	11h30	11h30	9h10	9h35	
	k=4	11h30	11h30	9h10	9h35	
	k=6	15h15	15h15	12h55	13h20	
	k=7	21h00	21h00	18h40	19h05	
	k=9	3h30	3h30	1h10	1h35	

Tableau 5

Laverie l	Commande Élémentaire k	Scénario A											Scénario B										
		Output j=1		Output j=2	Output j=1		Output j=2	Output j=1		Output j=2	Output j=1		Output j=2	Output j=1		Output j=2	Output j=1		Output j=2				
		k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10		
		s=2/i=5	s=2/i=5	s=1/i=2	s=1/i=2	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=2/i=5	s=1/i=2	s=1/i=2	s=2/i=6	s=2/i=6	s=1/i=2	s=2/i=6	s=1/i=2		
	Quantité prélevée (tonne)	3245	5949	4713	7564	3245	4687	3245	5048	3245	5949	4713	7564	3006	5901	3525	5483						
	Gamme h utilisée	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2						
	Quantité prélevée (tonne)	1646	2594	2618	4202	1803	2604	1803	2804	1646	2594	2618	4202	1893	2604	1803	2804						
	Gamme h utilisée	h=1	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=1	h=2	h=2	h=2	h=2	h=2	h=2	h=2						
	Input primaire i	s=4/i=12	s=4/i=12	s=3/i=8	s=4/i=10	s=4/i=12	s=4/i=12	s=4/i=12	s=4/i=12	s=4/i=12	s=4/i=12	s=4/i=12	s=4/i=12	s=3/i=8	s=4/i=10	s=4/i=12	s=4/i=12						
	Quantité prélevée (tonne)	2753	5048	3999	6418	2753	3977	2753	4283	2753	5048	3999	6418	2891	3977	2753	4283						
	Gamme h utilisée	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2	h=2						
	Input primaire i	s=6/i=17	s=5/i=13	s=5/i=13	s=5/i=13	s=5/i=17	s=5/i=13	s=5/i=13	s=5/i=13	s=6/i=17	s=5/i=13	s=5/i=13	s=5/i=14	s=5/i=13	s=5/i=13	s=5/i=15	s=5/i=13						
	Quantité prélevée (tonne)	2434	3748	2987		5048	1966	2839	1966	3058	2434	3748	3556	4779	2147	3515	2044						
	Gamme h utilisée	h=2	h=2	h=2		h=1	h=1	h=1	h=1	h=1	h=2	h=2	h=2	h=2	h=2	h=2	h=2						
	Input primaire i	s=5/i=13	s=6/i=17	s=6/i=17	s=6/i=18	s=6/i=13	s=6/i=17	s=6/i=17	s=6/i=17	s=5/i=13	s=6/i=17	s=6/i=18	s=6/i=18	s=6/i=18	s=6/i=17	s=6/i=18	s=5/i=13						
	Quantité prélevée (tonne)	1966	3959	3155		4595	2160	3119	2160	3389	1966	3959	3155	5048	2268	3119	2160						
	Gamme h utilisée	h=1	h=1	h=1		h=1	h=1	h=1	h=1	h=1	h=1	h=1	h=1	h=1	h=1	h=1	h=1						
	Total des prélèvements pour la commande élémentaire k (tonnes)	12044	21299	6142	11330	18184	9643	11927	17227	11927	18552	12044	21299	6710	11330	18184	9827	12204	18307	12285	19289		

Tableau 6

Commande élémentaire k	Composant	Laverie					TOTAL
		l=1	l=2	l=3	l=4	l=5	
k=1	Volume total (m ³)	2282	1268	2028	1711	1711	9000
	Volume résiduel du lot k-1 (m ³)	450	250	400	337.5	337.5	1775
	c=1 (%)	68.28%	64.11%	66.61%	65.72%	65.58%	66.32%
	c=2 (%)	0.83%	0.74%	1.07%	0.43%	0.78%	0.79%
	c=3 (%)	1.47%	1.98%	2.02%	2.90%	2.78%	2.19%
	c=4 (mg/P ₂ O ₅)	55.63	69.75	54.62	72.15	79.52	64.94
Gamme		h=3	h=2	h=2	h=1	h=1	
k=3	Volume total (m ³)	5318	2955	4727			13000
	Volume résiduel du lot k-1 (m ³)	450	250	400			1100
	c=1 (%)	67.38%	68.60%	64.39%			66.57%
	c=2 (%)	0.67%	0.89%	0.91%			0.81%
	c=3 (%)	1.55%	1.46%	2.32%			1.81%
	c=4 (mg/P ₂ O ₅)	62.41	55.40	59.20			59.64
Gamme		h=1	h=2	h=1			
k=5	Volume total (m ³)	1946	1331	2130	1797	1797	9000
	Volume résiduel du lot k-1 (m ³)	450	250	400	337.5	337.5	1775
	c=1 (%)	70.30%	67.59%	63.03%	64.32%	60.71%	65.07%
	c=2 (%)	0.89%	0.71%	0.99%	0.51%	0.47%	0.73%
	c=3 (%)	1.32%	1.53%	1.87%	2.89%	2.63%	2.06%
	c=4 (mg/P ₂ O ₅)	56.96	60.99	55.50	77.94	76.29	64.98
Gamme		h=1	h=2	h=1	h=1	h=1	
k=7	Volume total (m ³)	3296	1831	2930	2472	2472	13000
	Volume résiduel du lot k-1 (m ³)	450	250	400	337.5	337.5	1775
	c=1 (%)	71.18%	67.38%	64.73%	65.40%	63.65%	66.66%
	c=2 (%)	0.96%	0.67%	0.94%	0.61%	0.45%	0.75%
	c=3 (%)	1.24%	1.55%	2.24%	2.33%	2.85%	2.02%
	c=4 (mg/P ₂ O ₅)	55.41	62.41	58.53	78.67	74.77	64.93
Gamme		h=1	h=2	h=1	h=1	h=1	
k=9	Volume total (m ³)	3549	1972	3155	2662	2662	14000
	Volume résiduel du lot k-1 (m ³)	450	250	400	337.5	337.5	1775
	c=1 (%)	69.88%	67.52%	66.61%	64.60%	63.53%	66.60%
	c=2 (%)	0.76%	0.70%	1.07%	0.68%	0.70%	0.79%
	c=3 (%)	1.14%	1.54%	1.88%	2.48%	2.23%	1.82%
	c=4 (mg/P ₂ O ₅)	51.25	61.45	54.63	81.81	83.23	64.89
Gamme		h=1	h=2	h=1	h=1	h=1	

A multi-product scheduling heuristic for a multiple-source slurry pipeline in the phosphate industry

Résumé – Les problèmes d’ordonnement des pipelines retiennent de plus en plus l’attention des chercheurs en raison des économies potentielles sur les coûts opérationnels. Dans cet article, nous proposons une heuristique d’ordonnement multi-produits dans un slurry pipeline à sources multiples. Le slurry pipeline est un pipeline adapté au transport d’un mélange d’un concentré de minerai et d’eau (slurry). Dans cet article, nous souhaitons résoudre un problème d’ordonnement réel dans l’industrie du phosphate. L’algorithme d’ordonnement s’exécute en deux itérations : en commençant par l’ordonnement d’un produit dominant qui représente plus de 90% de la demande totale, et en

MOUAD SIDKI¹, NIKOLAY TCHERNEV^(1,2), SELWA ELFIRDOUSSI¹

¹ EMINES School of Industrial Management, Mohammed VI Polytechnic University
LOT 660 Hay Moulay Rachid, 43150 Benguerir, MAROC
{Mouad.Sidki, Nikolay.Tchernev, Selwa.Elfirdoussi}@emines.um6p.ma

² LIMOS, UMR CNRS 6158, Université Clermont Auvergne
1 rue de Chebarde, 63178 Aubière, FRANCE
nikolay.tchernev@uca.fr

terminant par l’ordonnement du reste des produits dans un ordre de priorité. En raison de la nature non linéaire du problème, une formulation en temps discret est choisie pour intégrer toutes les contraintes industrielles nécessaires. L’objectif de l’algorithme est de prendre des décisions de production dans les laveries, des décisions de transport dans les pipelines et des décisions de consommation dans la station terminale, pour chaque période de temps. D’autre part, à chaque période de temps, les contraintes de stockage ne doivent pas être violées.

Abstract – Pipelines scheduling problems are getting more attention due to the potential operational cost savings. A slurry pipeline is a pipeline suitable for transporting a mixture of an ore concentrate and water (slurry). In this paper, we propose a heuristic for the multi-product scheduling problem in a multiple-source slurry pipeline. We are interested in solving a real-world scheduling problem in the phosphate industry. The scheduling algorithm runs in two iterations: starting by scheduling a dominant product that makes more than 90% of the total demand, and ending by scheduling the rest of the products in a priority order. Due to the non-linear nature of the problem, a discrete time formulation is chosen to integrate all the necessary industrial constraints. The objective of the algorithm is to take production decisions in the washing stations, transportation decisions in the pipelines and consumption decisions in the terminal station, for every time period. On the other hand, at each time period storage constraints should not be violated.

Mots clés – Ordonnement, Pipeline multi-produits, Sources multiples, Heuristique, Phosphate.

Keywords – Scheduling, Multiproduct pipeline, Multiple Sources, Heuristic, Phosphate.

1 INTRODUCTION

Slurry pipelines are adapted pipelines for transporting a mixture of an ore concentrate and water (slurry). They are one of the most reliable and cost-effective ways of transporting large amounts of minerals over large distances. They ensure a continuous slurry delivery to terminal stations, which is more difficult to achieve using traditional transportation means such as rails and trucks. However, the multi-product scheduling of a single-source pipeline subject to delivery time-windows is NP-complete [Jittamai, 2004].

The Multi-product pipeline scheduling problem comes in a range of complexity. One simple problem would be with one source and one destination. More challenging problems have multiple sources and/or multiple destinations and additional industry specific constraints. The oil industry pipelines are the most common in the literature, and despite of the potential operational cost reduction, the transportation scheduling is

done manually in many cases due to the complexity of the real-world problem.

In the literature, one-to-one and one-to-many topologies are the most studied as they are the most common. The many-to-one topology has fewer publications. In this paper, we will be interested in proposing a solution to the many-to-one multi-product straight pipeline scheduling problem, in the phosphate industry, applied to a real-world operational problem. As illustrated in Figure 1, in this problem, each washing station (WS_i) is linked to the Head station via the use of straight secondary pipelines (SP_i). A washing station, for example (WS_1), may have its independent tank (SR_1) if the flow rate of production is different from the flow rate of transportation using its corresponding secondary pipeline (SP_1), otherwise, there is no storage in the washing station. The head station is linked to the terminal station using the slurry pipeline. We also

consider that each tank has a set of forbidden products to be stored in it.

As multi-product pipeline scheduling has received increasing attention in the last decades, different rigorous approaches have been proposed. We propose to split those approaches in

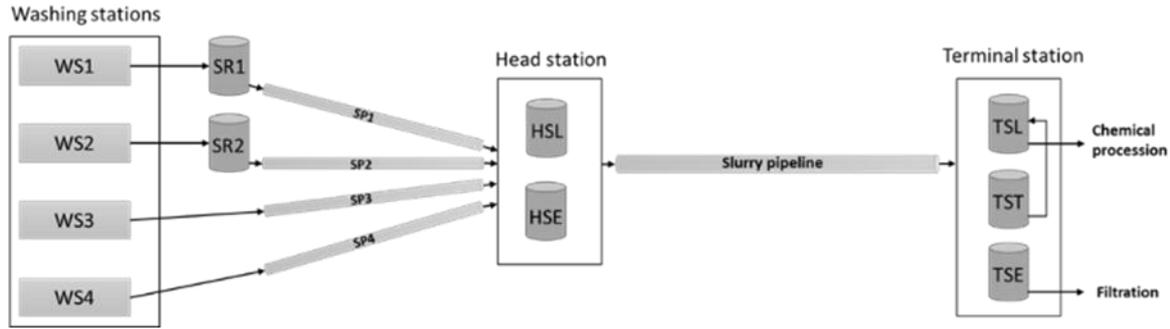


Figure 1. Multiple sources slurry pipeline

to minimizing the scale of the problem and so its complexity. [Cafaro and Cerdá, 2010] introduced a MILP continuous formulation for the scheduling of unidirectional pipeline networks that allows simultaneous batch injections at multiple nodes. This model proved to increase the use of the pipeline capacity, and decrease the overall time needed to meet depot demands. On the other hand, a discrete formulation aims to divide the scheduling horizon into multiple time intervals and the pipeline into multiple packages of equal volume each containing a single product. This approach can be used to simplify the non-linear relations between the problem variables. [Herrán et al., 2010] uses a discrete MIP mathematical approach to plan the transportation of multiple petroleum products in a multi-pipeline system. In this paper, the discretization approach is used and the scheduling horizon is divided into uniform time intervals. The nature of the industrial problem studied allows a formulation with a constant flow, as it mildly vary. For this reason, the packages are of uniform sizes.

Real world operational pipeline scheduling problems can be impractical to solve using mathematical optimization when the solution time becomes very important. For this reason, some approaches make use of knowledge-based heuristics search [Sasikumar et al., 1997] and discrete-event simulation [Mori et al., 2007] to reduce the solution space explored.

[Relvas et al., 2009] used a heuristic approach for batch sequencing on a multiproduct oil distribution system. The heuristic generates feasible batch sequences used by a MILP model to generate the scheduling solution and its inventory management. The used approach was tested on three real scenarios varying on the length of the scheduling horizon.

(Haoran et al., 2018) presents a solution for the multi-product pipeline detailed optimal scheduling problem using a self-learning approach. Using fuzzy clustering analysis, ant colony optimization algorithm and a MILP model, the algorithm proposed by the authors generates a historical data base containing a set of problems and solutions. Based on the given problem to solve, fuzzy clustering is used to provide the closest initial solution to the problem from the data base. The MILP model and the ant colony algorithm are used then to find the optimal solution. As a final step, the solution and the problem instance are stored in the data base. This approach makes use of the historical data base, which can be self-generated, to find the closest initial solution to the optimal solution, therefore the calculation converges relatively faster. This approach was tested on real world scenarios and proved to have a big impact on the CPU time of the solution.

two categories based on time and volume formulation. On the first hand, we have a continuous-time formulation where the scheduling horizon is divided based on the start and end time of each task. This formulation proves to be the most reliable when it comes

In this paper, we propose a heuristic for the multi-product scheduling problem in a multiple-source slurry pipeline. The scheduling algorithm runs in two iterations: starting by scheduling a dominant product that makes more that 90% of the total demand, and ending by scheduling the rest of the products in a priority order.

The rest of the paper is organized as follows. Section 2 presents a problem definition. In Section 3 initialization parameters are formally introduced. Section 4 introduces the decision variables. In section 5 the proposed heuristic algorithms are presented. Section 6 presents the computational results for a real case study. Finally, Section 7 presents conclusions and future research which may result from this work.

2 PROBLEM DEFINITION

This paper addresses the multiple-source pipeline scheduling problem in the phosphate industry. Solid phosphate is processed by washing stations to produce a final product from the set $P = \{P_1, P_2, \dots, |P|\}$. We can differentiate two types of products. The first type is a dominant product P_1 , used by multiple customers in the downstream for chemical processing. The tanks of the terminal station dedicated to P_1 storage must be filled at all times since the chemical processes require a continuous flow of phosphate pulp to operate normally. The second type of products is filtered to reduce the amount of water it contains, then it is dried and stored. The second type products are present in the set $P - \{P_1\}$. All the tanks can allow the storage of only one of the two product types. P_1 can be stored in all tanks but not in HSE and TSE, and the rest of the products can be stored in all tanks but not in HSL, TSL and TST. In addition, only one product P_i can be stored in a tank at a given time. Therefore, tanks must be empty before storing a different product. In reality, due to technical reasons, the storage level of a tank must always be between a lower limit, strictly superior to zero, and upper limit. Therefore, when the product in a tank needs to be replaced by a new one, there is still a quantity of the preceding product present in it. Therefore, depending on the new product relative to the previous one, a contamination of the new product may happen. $Q = \{Q_1, Q_2, \dots, |P|\}$ is the set of the quality scores in order of the products present in P . when $Q_{\text{new}} > Q_{\text{old}}$, a rinsing operation is performed for the tanks risking to cause a contamination, starting from the Head Station. RQ is the quantity of the rinsing batch, which is used to rinse all the tanks risking a contamination. We consider $S = \{SR_1, SR_1,$

HSL, HSE, TSL, TST, TSE } to be the set containing all the tanks. Without loss of generality id considered only the useful storage volume in a tank. Therefore, 0 is the lower limit and S_{max} is a set containing the useful storage capacities $\{SR_{1,max}, \dots, HSL_{max}, \dots, TSE_{max}\}$. $W = \{W_1, W_2, \dots, |W|\}$ is a set containing all the washing stations. Each washing station is able to produce all the products in P . WF is the production flow matrix of all the products in P for all washing stations in W . Therefore, $WF_{i,w}$ $i \in \{1, \dots, |P|\}, w \in \{1, \dots, |W|\}$ is the production flow of product P_i in the washing station w . In addition, washing stations have a setup time when changing the quality in production. As a simplification we consider WST to be the setup time of all the washing stations. In each time period, the choice of the washing stations to produce a product is done using the matrix WP . $WP_{i,w}$ $i \in \{1, \dots, |P|\}, w \in \{1, \dots, |W|\}$ is a score given to the production of P_i in W_w . The washing station with the highest score is chosen first to produce P_i . In addition, a minimal production batch quantity needs to be respected. $WQ = \{WQ_1, WQ_2, \dots, |W|\}$ is a set containing the minimal production quantities of all the washing stations.

Due to some technical constraints, such as the pumping power of the secondary pipelines present in the set $SP = \{SP_1, SP_2, \dots, |SP|\}$ and the topology of the territory between the washing stations and the head station, the flow of the secondary pipelines can't be always adjusted to be equal to the flow of production. Therefore, if two flows are different, a tank is needed to link between the washing station and its secondary pipeline. $SPF = \{SPF_1, SPF_2, \dots, |SP|\}$ is a set containing the transportation flows of the secondary pipelines. In addition, different products in the same secondary pipeline needs to be separated by a minimal quantity of water. $SPW = \{SPW_1, SPW_2, \dots, |SP|\}$ is the minimal water separation quantities. The volume of a secondary pipeline, its transportation flow, its maintenance profile and the shutdown decisions made can be used to link each discharging time period in the head station to its corresponding charging time period in the washing station. $SPV = \{SPV_1, SPV_2, \dots, |SP|\}$ contains the volumes of the secondary pipelines. Moreover, the secondary pipelines have a minimum batch quantity constraint to be respected. $SPmin = \{SPmin_1, SPmin_2, \dots, |SP|\}$ is a set containing the minimal batch quantities allowed in the secondary pipelines.

In the head station, tank HSL is used to store only P_1 and HSE is used to store the rest of the products. A valve system is used to choose between the two tanks to load the slurry pipeline PL of volume PV and a flow equal to PF . Similar to the secondary pipelines, the slurry pipeline has a minimal batch size condition to be respected. We consider P_{min} as the minimal allowed batch size. In addition, due to industrial constraints, a maximal batch size constraint is added. P_{max} is the maximal batch size allowed. Furthermore, two batches of different products need to be separated by water, the minimal water separation batch size is PW . The slurry pipeline in this study is different from the oil industry pipelines, as no shutdown is allowed beside the preventive maintenance. Therefore, water is charged when no pulp is taken from the head station tanks. Infinite source of water is considered.

In the terminal station, TSL and TST are used to store P_1 . The customers of P_1 are linked to TSL , which can be also filled from TST using a pipeline. The flow of this pipeline is considered to be constant and equal to TF . The pulp transfer operation between the two tanks is important in the case of a high P_1 demand, as it ensures a longer time without the need to

discharge a new batch in TSL . We consider this pipeline ("transfer pipeline") to have a transportation duration equal to zero. Moreover, TSE is the tank dedicated to the products other than P_1 , it is linked to the filtration unit of a flow equal to FF . A valve system enables the operators to chose which tank to discharge a batch in. We consider STT_i $i \in \{2, \dots, |P| - 1\}$ to be the cumulative filtered quantity of product P_i at the end of the scheduling horizon.

A discrete time is chosen to solve this scheduling problem. The time horizon is divided into equal time intervals. Furthermore, each pipeline is divided into equal packages. We consider t_{start} as the effective start of the scheduling and t_{end} as the end of the time horizon. We also consider t_0 to be the start of the initialization period that ends at t_{start} . Moreover, $T = \{1, 2, \dots, |T|\}$ contains all the time intervals in the scheduling horizon.

3 INITIALIZATION PARAMETERS

Each tank is initialized at t_{start} with a quantity and a product id . $\{1, 2, \dots, |P|\}$ are the ids corresponding to the products in P . The tanks initialization data is stored in SI where $SI_b = \{i, TSQ_b\}$ $b \in \{1, \dots, |S|\}$ $i \in \{1, \dots, |P|\}$ is the initialisation of tank b with the product P_i and an initial quantity equal to TSQ_b .

The customers demand initialization depends on the type of the product. For P_1 , the demand is modelled as a set containing quantities to be taken from TSL at each time period. LD is a vector of size $|T|$ containing P_1 demand. For the other product, we consider the delivery due dates to be at the end of the time horizon as we are interested in a short-term scheduling problem. Therefore, the demand for a product P_i is modelled in the matrix ED , of size $(|P| - 1) \times 2$, as a quantity and a priority couple for each product. For $ED_i = \{EDS_i, EDQ_i\}$ $i \in \{2, \dots, |P|\}$, EDS_i is the priority score of P_i (the product with the highest priority score is scheduled first) and EDQ_i is the quantity of product P_i to be filtered at the end of the time horizon.

In the phosphate industry, preventive maintenance is done frequently. Therefore, for a scheduling plan to be functional in the real world, it needs to take planed shutdowns in consideration. In this study, we consider the preventive maintenance of the washing stations and the secondary pipelines. The maintenance of a unit is modelled as a vector of size $|T|$ of a binary nature (equal to 1 at a time period if a shutdown should be performed, otherwise it is equal to 0). We denote the washing stations maintenance plan as the Matrix WM of size $|W| \times |T|$, containing the maintenance vectors of all washing station. WM_w $w \in \{1, \dots, |W|\}$ is the maintenance schedule of size $|T|$ of the washing station w . At the other hand, the maintenance plan of the secondary pipelines is modelled in the same manner. SPM_s $s \in \{1, \dots, |SP|\}$ is the maintenance set, of size $|T|$, of the secondary pipeline s .

Pipelines are initialized, before t_{start} , with the charging schedule of the batches that occupies their inner volume. And the washing stations are initialized with the residual scheduling plan still to be performed. For a secondary pipeline, the slurry pipeline or a washing stations, we consider an initialization vector of size $|T|$, equal for a time interval t to $i \in \{1, \dots, |P|\}$ (the charging or the production of P_i is done at t), or it can be equal to the id of shutdown Ids (a shutdown should be performed at t). For the secondary pipelines and the slurry pipeline water can be charged, we consider Idw to be the id corresponding to the water charging operation at t . We consider $SPI = \{SPI_1, SPI_2, \dots, |SP|\}$, of size $|SP| \times |T|$, to be

the initialization decision matrix of the secondary pipelines. PI ($|T|$) is the initialisation vector of the slurry pipeline and $WI = \{WI_1, WI_2, \dots, |W|\}$ ($|W| \times |T|$) is the initialisation matrix of the washing stations.

4 DECISION VARIABLES

To achieve a scheduling plan for the washing stations, the secondary pipelines, the slurry pipeline and the transfer pipeline, a decision needs to be made for every unit at each time period starting from t_{start} . For the tanks, we should indicate, at each time period, the product and its quantity stored. Therefore, we consider $SD_{b,t} = \{SDid_{b,t}, SDq_{b,t}\}$ $b \in \{1, \dots, |S|\}$, $t \in \{1, \dots, |T|\}$ containing the product's id $SDid_{b,t}$ and its quantity $SDq_{b,t}$ present in the tank S_b at a time period t . The tanks storage profile depends on the pipelines and the washing stations decisions. For the washing stations, we consider $WD_{w,t}$ $w \in \{1, \dots, |W|\}$, $t \in \{1, \dots, |T|\}$ to be the decision id corresponding to the scheduling to be performed in the washing station W_w at the time interval t . For the secondary pipelines, $SPD_{s,t}$ $s \in \{1, \dots, |SP|\}$, $t \in \{1, \dots, |T|\}$ corresponds to the decision to be made for the secondary pipeline SP_s at the time interval t . In addition, PD_t $t \in \{1, \dots, |T|\}$ corresponds to the decisions to be made for the slurry pipeline and TD_t $t \in \{1, \dots, |T|\}$ indicates if at a time period t , the transfer operation is performed or not (equal to 1 if yes otherwise is equal to 0). FD_t $t \in \{1, \dots, |T|\}$ is the decision to run the filtration operation or not at t (equal to 1 if yes otherwise is equal to 0). All the batches in the pipelines are linked to a tank to be stored in. The batches in the secondary pipelines are stored in HSL if the product present in the batch is P_1 , otherwise it is stored in HSE . The transfer pipeline fills TSL from TST . And for the slurry pipeline, $P - \{P_1\}$ are stored in TSE and P_1 is stored either in TSL or in TST depending on the wanted tank. Let be $TSTD_t$ $t \in \{1, \dots, |T|\}$, a binary value, that indicate if what is charged in the slurry pipeline at t will be discharged in TST ($TSTD_t = 1$), otherwise it will be discharged in TSL ($TSTD_t = 0$).

5 SCHEDULING HEURISTICS

The algorithms that follows were chosen due the nature of the problem described above. An analysis of the real-world parameters concludes the following points :

- The yearly demand for P_1 is above 90% of the total demand.
- The problem is storage sensitive. For example, during a peak in demand of the product P_1 , TSL and TST can't cover the client's needs for more than 4 hours.
- To not lose in precision and for the proposed schedule to be valid, the time intervals should be less than 15 minutes, as it is the time to charge the slurry pipeline with the minimal water separation quantity PW . Therefore, $Step = 1/4$ h is chosen to be the time intervals length.
- A seven days scheduling horizon is more than enough for the operational needs.
- The scheduling solution must take a relatively short amount of time to be computed, as the initial parameters can change at any time.

For the details described above, a tow steps scheduling heuristic is considered to solve the operational scheduling problem. P_1 is scheduled first, during the whole scheduling horizon, with an aim to maximize the time between the end of the discharge of a batch and the start of discharging the

following batch. The transfer operation is used when no P_1 batch is discharged in TSL . This rule makes it possible to send important batch sizes of products in $P - \{P_1\}$ during the time windows available, after the scheduling of P_1 is achieved. In the first scheduling step, the choice of the washing stations to produce P_1 depends on the storage level in TST . When TST reaches a critical storage level $TSTc$, the total flow arriving to the head station from the washing stations should be greater and close to the consumption flow of P_1 in the terminal station. When TST is filled to its maximum or if any constraint may be violated, this production mode ends. In the normal case when TST level is above $TSTc$, the washing stations are chosen to have a total production flow, arriving to the head station, that is inferior and close to the demand flow in the terminal station. This rule ensures that, for the washing station with the lowest production score of P_1 , the time between a shutdown and a turn-on is maximized, leaving a relatively important interval of time for the other products to be scheduled. The washing stations dedicated for P_1 production, at a time period t , are chosen one by one using the production priority vector WP_1 until the total flow achieves the wanted goal. We consider $WSset_t$ to be the Ids set of the chosen washing stations at t to produce P_1 . We should mention that the flow of a product i , arriving to the head station from a given washing station w_w , is equal to $\min(WF_{i,w}, SPF_1)$. On the other hand, products in $P - \{P_1\}$ are scheduled by their order of priority. For example, the product P_i with the highest score EDS_i is scheduled first until its demand is satisfied. The scheduling continues until the time horizon ends or all the demand is satisfied.

5.1 Storage management

The storage levels of the tanks, at each time period, are based on the washing stations and the pipelines decisions made as follows. First at each time period t , are determined: (i) a charging time of the secondary pipelines relative to its discharge ; (ii) a discharging time of the slurry pipeline relative to its charging at t . Let $SPT_{s,t}$ be the time period of charging a product or water in the secondary pipeline s relative to its discharge in the head station at t and PT_t is the time period of discharging a product or water in the terminal station relative to its charging in the head station during t . Taking into account the shutdowns and the preventive maintenance of the pipelines $SPT_{s,t}$ and PT_t are determined considering all the transportation timing calculations using the algorithm "Transport_timing_algorithm". After computing the transportation timing for the pipelines, tanks storage levels can be calculated for the scheduling horizon based on the pipelines, the washing stations and the filtration discussions. Let "Storage_algorithm" be the algorithm that computes the quantity and the id of the product stored, for each tank and each time period.

5.2 Washing stations and secondary pipelines scheduling

Based on the difference between the production flow of a washing station and the flow of the associated secondary pipeline, a different scheduling approach for each case is adopted. We consider WT_w $w \in \{1, \dots, |W|\}$ to be the type of the difference between the tow flows. For instance, if $WF_{i,w} < SPF_w$, for a given product P_i , $WT_w = 1$ and if $WF_{i,w} > SPF_w$ the type $WT_w = 2$, otherwise, $WT_w = 3$.

The following algorithm (Algorithm 1) schedules a washing station W_w and its corresponding secondary pipeline SP_w during each time period t .

WS_Scheduler_algorithm			
Input			
i	:		a product id
w	:		a washing station id
t	:		a time period
WT_w	:		W_w flow type
$SPSP_w$:		SP_w shutdown profile
Begin			
1.		If $SPSP_{w,SPT_{w,t}} = 0$ Do	
2.		$V \leftarrow SPSP_w$	
3.		WS_Scheduler(type $\leftarrow WT_w$, time $\leftarrow SPT_{w,t}$, i , w)	
5.		If not $SP_w = V$ Do	
6.		Transport_timing_algorithm	
7.		EndIf	
8.		Storage_algorithm	
9.		EndIf	
End			

Algorithm 1. Washing stations and secondary pipelines scheduling

In Algorithm 1, $SPSP_w |T|$ is the shutdown profile of the secondary pipeline w . (if $SPD_{w,t} = Ids$ or $SPM_{w,t} = 1$, $SPSP_{w,t} = 1$, otherwise $SPSP_{w,t} = 0$). In line 1, the algorithm verifies if $SPSP_{w,SPT_{w,t}} = 0$. If it is true, the scheduling of SP_w and W_w is possible. In other terms, if SP_w is not discharging in the head station during t , no equivalent charging time period exists ($SPT_{w,t} = 0$). Therefore, neither the production nor the transportation can be scheduled.

In line 2, the current state of $SPSP_w$ is stored in the vector V , it is compared with its new state (line 5) after the scheduling is achieved using WS_Scheduler (wash station scheduling). If there is a change in the shutdown profile, Transport_timing_algorithm is executed to establish new links between the charging and discharging time periods of the pipeline SP_w . Moreover, Storage_algorithm is executed at the end to recalculate the tanks storage levels based on the decisions made in $WD_{w,t}$ and $SPD_{w,t}$.

Based on the value of WT_w , WS_Scheduler adopt a specific scheduling strategy. When $WT_w = 1$ the tank S_w needs to attain a certain storage level S_{min_w} before stating the charging operation of a new batch. For a product P_i , S_{min_w} depends on $WF_{i,w}$, SPF_w and the minimal batch size $SPmin_w$. Therefore, $S_{min_w} = (1 - \frac{WF_{i,w}}{SPF_w})SPmin_w$. If $SPD_{w,SPT_{w,t}} = 0$, $SPD_{w,SPT_{w,t-1}} \neq i$ (figure 2, Step 0), and it is possible to schedule a P_1 batch b in W_w that ends at $SPT_{w,t}$ of size S_{min_w} , the batch b is planned in the washing station W_w and $SPD_{w,SPT_{w,t}} = i$ (figure 2, Step 1). In addition, if $SPD_{w,SPT_{w,t}} = 0$, $SPD_{w,SPT_{w,t-1}} = i$ and $SDq_{w,SPT_{w,t}} > SPF_1 \times Step$, the decision variables WD_w does not change and $SPD_{w,SPT_{w,t}} = i$ (figure 2, Step 2). Otherwise, if $SDq_{w,SPT_{w,t}} < SPF_1 \times Step$ a new batch of the size $SPF_1 \times Step$ is scheduled before b (figure 2, Step 3). Furthermore, if a batch of SP_w ends, a shutdown duration is scheduled after the end of the batch, equivalent to the time it takes W_w to fill up to S_{min_w} (figure 2, Step 4). The objective of this scheduling rules is to send the biggest batches possible and always end with a storage level in $S_{w,close}$ to 0. Therefore, the tank can accept a new product of different quality.

For the case when $WT_w = 2$, if the tank S_w attains its maximal level during $SPT_{w,t}$, $WD_{w,t} = Ids$ and $SPD_{w,SPT_{w,t}} = i$. The shutdown continues for a known period of time $WT2S$.

otherwise, $WD_{w,t} = i$ and $SPD_{w,SPT_{w,t}} = i$. At last, for $WT_w = 3$, $WD_{w,t} = i$ and $SPD_{w,SPT_{w,t}} = I$ for all cases.

It should be noted that WS_Scheduler verifies at all times if the decisions that could be made violates the lower and upper limits constraints of the tanks. If it is the case, no decisions are made.

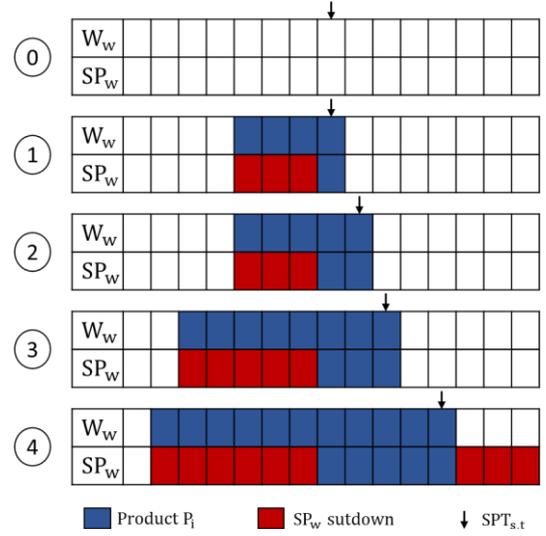


Figure 2. Type 1 washing station scheduling

5.3 General scheduling heuristic

The general scheduling algorithm runs in two iterations. Firstly, P_1 is scheduled on the chosen washing stations, secondary pipelines, the slurry pipeline and transfer the pipeline for every time period of the scheduling horizon. Secondly the rest of the products are scheduled in a priority order defined by EDS_i $i \in \{2, \dots, |P|\}$, until all the demand is satisfied or the scheduling horizon ends. We consider that the washing station W_w , having the lowest score $WP_{1,w}$ of P_1 production, to be associated with the production of products in $P - \{P_1\}$. Algorithm 2 starts by initializing its parameters with a given instance. Its main objective is to outputs a detailed scheduling plan for all the units, starting from the washing stations end ending in the terminal station. It is designed to be a decision assistant tool, to replace the manual scheduling which can be short sighted and time consuming.

In lines 1 and 2, transportation timing logic is computed and the initial storage levels are calculated and stored in SD . Afterward, P_1 is scheduled between line 3 and line 20. As we are interested in scheduling the charging operations of the slurry pipeline, the last time period t scheduled would be $t_{end} - P_{travel}$, as it takes P_{travel} times steps for a batch to be discharged in the terminal station. Therefore, for every time period, starting from t_{start} and ending at $t_{end} - P_{travel}$, Production_mode_algorithm selects a set of washing stations for P_1 to be scheduled in based on the logic described before. Then, WS_Scheduler_algorithm(WT_w , t , 1, w) is executed to schedule P_1 in every selected washing station w for the time period t (line 6). Afterword, if P_1 _Scheduling_is_possible(t) is true, P_1 is scheduled during t in the slurry pipeline, otherwise, if it is possible to schedule the transfer operation at PT_t (without violating TSL and TST capacity constraints), the transfer of P_1 at PT_t is achieved. P_1 _Scheduling_is_possible(t) returns True if specific conditions are met : If P_1 is not scheduled during $t - 1$ ($PD_{t-1} = 0$), the storage level of HSL should exceed a certain value $HSLmin$ equivalent to the minimal batch size P_{min} allowed in the slurry pipeline and the

storage level of TSL should be lower than $TSL_{\max} - HSL_{\min}$. This condition ensures the possibility of sending a batch of a volume superior to P_{\min} at all times. On the other hand, if $PD_{t-1} = 1$, the decision $PD_t = 1$ is taken if no storage constraints can be violated. When TSL attains its maximal level, $TSTD_{PT_t} = 1$ (for 1 hour, after that, the flow is reversed back to TSL). If the slurry pipeline batch volume attains the value of P_{\max} , the charging ends. Moreover, in line 19, $Correction_algorithm(t)$ is used to correct the minimal batch size violations in the secondary pipelines. As an example, if a batch of a secondary pipeline SP_w end at $SPT_{w,t} - 1$ without respecting the minimal batch size constraints, the batch scheduling decisions are deleted for SP_w and W_w . Then, PD and TD are updated using the decision rules described above. The deletions may cause storage perturbations of TSL during the time periods before PT_t , a simple security storage level constraint is enough to cover the possible sold out. At last, $Storage_algorithm$ is executed to update the storage levels (line 15) and the minimal water separation quantity is planned on the slurry pipeline and every secondary pipeline using

$Plan_water_separation$ (line 21).

Products in $P - \{P_1\}$ are scheduled in a different manner (lines from 22 to 38). For every time period in $\{t_{\text{start}}, \dots, t_{\text{end}} - P_{\text{travel}}\}$, a product P_i , with the biggest score EDS_i , verifying the condition $EDQ_i < STT_i$, is chosen to be produced in the dedicated washing station W_w (lowest score for P_1 production). $Product_choice$ outputs the chosen product's id (line 23). Then, $WS_Feasible_batch_start(SPT_{w,t}, w, i)$ (line 24) is used to verify if it is feasible to start the production of a batch ba of minimal size of P_i in W_w starting from $SPT_{w,t}$. If it is the case, HSE should be empty or contain the same product P_i and there should be enough storage space in HSE . In lines 25 and 26, the batch ba is scheduled starting from $SPT_{w,t}$ until $SPT_{w,t} + WS_Steps_in_batch(i, w)$ ($WS_Steps_in_batch(i, w)$ is the time steps needed to schedule ba in SP_w). on the other hand, a new loop starts for the slurry pipeline and the filtration scheduling, starting from t and ending at $t_{\text{end}} - P_{\text{travel}}$ (line 28). Then, the maximal size batch possible (respecting the minimal size constraint and of $P_Steps_in_batch$ time periods) is scheduled in PL starting from the time period j . j verifies the

General_Scheduling_Algorithm

Input

I : Problem initialisation

Output

{SD, WD, SPD, PD, TD, FD }

Begin

```

1. Transport_timing_algorithm
2. Storage_algorithm
3. For all t in {t_start, ..., t_end - P_travel }Do
4.   WSset_t ← Production_mode_algorithm
5.   For all w in WSset_t Do
6.     WS_Scheduler_algorithm(WT_w, t, 1, w)
7.   EndFor
8.   If P_1_Scheduling_is_possible(t) = True Do
9.     PD_t ← 1
10.    Storage_algorithm
11.   EndIf
12.   Else Do
13.     If SDq_{6,t} ≥ TF × Step and TF × Step + SDq_{5,t} ≤ TSL_max Do
14.       TD_{PT_t} ← 1
15.       Storage_algorithm
16.     EndIf
17.   EndElse
18.   Storage_algorithm
19.   Correction_algorithm(t)

```

Algorithm 2. General scheduling algorithm

```

21. Plan_water_separation
22. For all t in {t_start, ..., t_end - P_travel }Do
23.   i ← Product_choice
24.   If WS_Feasible_batch_start(SPT_{w,t}, w, i) = True Do
25.     For j in {t, ..., t + WS_Steps_in_batch(i, w)} Do
26.       WS_Scheduler_algorithm(WT_w, j, i, w)
27.     EndFor
28.     For j in {t, ..., t_end - P_travel } Do
29.       If P_Feasible_batch_start(j, i) = True Do
30.         For v in {j, ..., j + P_Steps_in_batch} Do
31.           PD_j ← i
32.           Storage_algorithm
33.         EndFor
34.         Replan_filtration_algorithm
35.       EndIf
36.     EndFor
37.   EndIf
38. EndFor
End

```

condition $P_Feasible_batch_start(j,i) = True$, if the scheduling is feasible (lines 29 to 33). If a new batch is scheduled in the slurry pipeline, the filtration is rescheduled for every v from t_{start} to t_{end} using `Replan_filtration_algorithm` (line 34). The scheduling of the filtration operation is done as soon as possible (for every v in $\{t_{start}, \dots, t_{end}\}$, if $SDq_{7,v} > FF \times Step$ then $FD_v = 1$).

In order to test the problem formulation and the scheduling algorithm proposed, several real world instances were studied during the time period of several months and validated by the industrial scheduling team as feasible scheduling solutions. We suggest in this section a case study of a real-world scenario with a scheduling horizon of 8 days.

6.1 Parameters initialization

What follows are the parameters initializing the case study.

6 CASE STUDY

Day	7/13	7/14	7/15	7/16	7/17	7/18	7/19	7/20	7/21
Flow (m ³ /h)	3387	3364	3245	3337	3031	3566	3156	3201	3475

Table 1. Daily P1 demand flow

Unit	Product	Start	End
SP ₁	P ₁	7/13/2020 13:30:00	7/13/2020 15:45:00
SP ₂	P ₁	7/13/2020 00:30:00	7/13/2020 15:45:00
SP ₃	P ₁	7/13/2020 15:30:00	7/13/2020 15:45:00
SP ₄	P ₁	7/13/2020 15:45:00	7/13/2020 15:45:00
PL	P ₁	7/12/2020 13:30:00	7/12/2020 15:15:00
PL	Water	7/12/2020 15:30:00	7/12/2020 16:45:00
PL	P ₁	7/12/2020 17:00:00	7/13/2020 00:15:00
PL	Water	7/13/2020 00:30:00	7/13/2020 01:45:00
PL	P ₁	7/13/2020 02:00:00	7/13/2020 14:30:00
PL	Water	7/13/2020 14:45:00	7/13/2020 15:45:00

Table 2. Pipelines batch initialization

Unit	Start	End
W ₄	7/13/2020 16:00:00	7/15/2020 08:00:00
SP ₄	7/13/2020 16:00:00	7/15/2020 08:00:00

Table 3. Preventive maintenance time intervals

$P = \{P_1, P_2, P_3\}$
 $Q = \{4, 3, 8\}$
 $RQ = 4000 \text{ m}^3$
 $S = \{SR1, SR2, HSL, HSE, TSL, TST, TSE\}$
 $S_{max} = \{4620, 9240, 9240, 9240, 9240, 13860, 13860\}$
 $W = \{W_1, W_2, W_3, W_4\}$
 $WF = \{\{900, 1550, 900, 500\}, \{900, 1550, 900, 500\}, \{900, 1550, 900, 500\}\}$
 $WST = 1h$
 $WP = \{\{3, 10, 8, 7\}, \{7, 0, 3, 3\}, \{7, 0, 3, 3\}\}$
 $WQ = \{4000, 4000, 4000, 4000\}$
 $SP = \{SP_1, SP_2, SP_3, SP_4\}$
 $SPF = \{1350, 1400, 900, 500\}$
 $SPW = \{500, 500, 500, 500\}$
 $SPV = \{3529, 2369, 226, 226\}$
 $SPmin = \{4000, 4000, 4000, 4000\}$

$\{PV, P_{min}, P_{min}, PW\} = \{106059, 4000, 4000, 80000, 1000\}$
 $PF = 4000$
 $\{TF, FF\} = \{1000, 600\}$
 $t_{start} = 7/13/2020 16:00:00$
 $t_{end} = 7/21/2020 16:00:00$
 $t_0 = 7/11/2020 16:00:00$
 $SI_b = \{\{1,280\}, \{1,6060\}, \{1,4140\}, \{2,2760\}, \{1,8060\}, \{1,780\}, \{2,0\}\}$
 $ED = \{\{7,21250\}, \{6,9000\}\}$
 In addition, Table 1, Table 2 and table 3 presents in order : the daily P₁ demand flow, Pipelines batch initialization and Preventive maintenance time intervals.

6.2 Results

The Gantt diagram in Figure 3 illustrates the scheduling of all the units. It shows the charging schedule of the pipelines and the production and filtration schedules. On each time

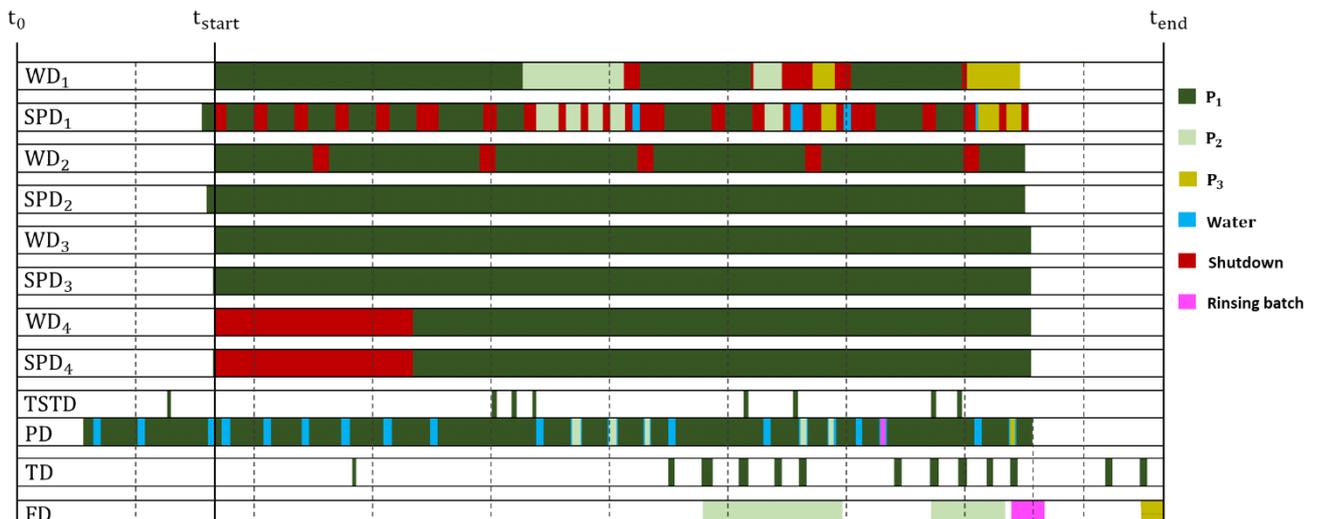


Figure 3. Scheduling diagram

period, the decision vectors are equal to a product Id, shutdown Id, water id, 0 or 1 for the binary vectors or to a rinsing product id. In this problem we have : a preventive maintenance for W_4 and SP_4 for 16h, a rinsing operation and 3 products to be scheduled. W_1 was chosen for P_2 and P_3 production as it has the lower priority score for P_1 production. To avoid the contamination of P_3 , the first P_3 batch scheduled in W_1 is considered a rinsing batch. The production of P_3 can continue only if HSE is empty. The rinsing batch is charged in the slurry pipeline to arrive to the terminal station after TSE is emptied from P_2 . TSTD is the decision to store P_1 in TST. TSTD = 1 during one hour (4 time periods) after TSL attains its maximal storage capacity. PTT is the transportation time in the slurry pipeline (26h and 31min).

The conducted numerical experiments were carried on an AMD Ryzen 7 3700X CPU, 3.59 GHz computer with 16 GB of RAM running under Windows 10. The proposed scheduling algorithm was coded in python and the results were illustrated in Excel. The scheduling algorithm took 287 seconds to compute the scheduling solution. The sold-out of P_1 is 3082 m^3 (0.49% of the total P_1 demand). $\{STT_2, STT_3\} = \{25950 \text{ m}^3, 2550 \text{ m}^3\}$, therefor, P_2 demand is fully satisfied and the sold-out of P_3 is 6450 m^3 . The scheduling ends with a full production of P_3 and filled HSE and TSE.

7 CONCLUSION

The main objective of this work was to provide an operationally feasible scheduling solution for the multi-product scheduling problem, in a multiple-source slurry pipeline, in the phosphate industry. The scheduling algorithm takes into consideration multiple non-linear constraints, such as tanks rinsing constraints and water separation constraint between different products in the pipelines. The algorithm aims to satisfy clients demand while having the least switch offs of the washing stations. During a period of 3 months, the scheduling solutions were tested on the real world with satisfying results.

Future works will focus on finding the optimal solution to the problem as a baseline to compare the performance of the scheduling algorithm proposed in this article.

8 REFERENCES

- Jittamai, P. (2004). *Analysis of oil pipeline distribution of multiple products subject to delivery time windows*. (Doctoral dissertation, Texas A&M University).
- Cafaro, D. C., & Cerdá, J. (2010). Operational scheduling of refined products pipeline networks with simultaneous batch injections. *Computers & Chemical Engineering*, 34(10), 1687-1704.
- Herrán, A., de la Cruz, J. M., & De Andrés, B. (2010). A mathematical model for planning transportation of multiple petroleum products in a multi-pipeline system. *Computers & chemical engineering*, 34(3), 401-413.
- Sasikumar, M., Prakash, P. R., Patil, S. M., & Ramani, S. (1997). PIPES : A heuristic search model for pipeline schedule generation. *Knowledge-Based Systems*, 10(3), 169-175.
- Mori, F. M., Lüders, R., de Arruda, L. V. R., Yamamoto, L., Bonacin, M. V., Polli, H. L., ... & de Jesus Bernardo, L. F. (2007). Simulating the operational scheduling of a realworld pipeline network. *In Computer Aided Chemical Engineering*, 24, 691-696.
- Relvas, S., Barbosa-Póvoa, A. P. F., & Matos, H. A. (2009). Heuristic batch sequencing on a multiproduct oil distribution system. *Computers & Chemical Engineering*, 33(3), 712-730.
- Haoran, Z., Yongtu, L., Qi, L., Yun, S., & Xiaohan, Y. (2018). A self-learning approach for optimal detailed scheduling of multi-product pipeline. *Journal of Computational and Applied Mathematics*, 327, 41-63.

Session 1.2 : Design et Pilotage des systèmes de production de biens et de services

Chairman : Yves Sallez - Bernard Penz

Etude prospective d'un environnement d'aide à la décision rendu configurable par des modèles

LIWEN ZHANG^{1,3}, HERVE PINGAUD², ELYES LAMINE¹, FRANCK FONTANILI¹, CHRISTOPHE BORTOLASO³, MUSTAPHA DERRAS³,

¹ Centre Génie Industriel, Université de Toulouse-IMT Mines Albi
Campus Jarlard, 81000 Albi, France
{prénom.nom}@mines-albi.fr

² CNRS LGC, Université de Toulouse-INA Champollion
Place de Verdun, 81012 Albi, France
herve.pingaud@univ-jfc.fr

³ Berger-Levrault
Rue Jean Rostand, 31670 Labège, France
{prénom.nom}@berger-levrault.com

Résumé – Les circuits de décision, le rythme de prise de décision, autant que le poids de celles-ci, prennent de nouvelles dimensions dans les organisations. Nous nous intéressons à la capacité du décideur à inclure les caractéristiques propres à son écosystème, son contexte et ses enjeux, souvent changeants, dans ses pratiques. Les décideurs sont souvent des experts métiers à qui nous essayons de procurer la capacité de concevoir et adapter à moindre effort leurs outils de prise de décision. A cette fin, la création d'un environnement d'aide à la prise de décision utilisant une approche d'Ingénierie Dirigée par les Modèles (IDM) est proposée. Ce travail exploratoire cherche à lier des modèles de besoin de prise de décision à des modèles logiques d'appel à un solveur en optimisation combinatoire. Les architectures fonctionnelles et logiques de cet environnement sont décrites sur un cas d'études, le problème du voyageur de commerce. Nous expliquons les hypothèses et les connaissances véhiculées par les modèles peuplant la chaîne d'ingénierie en suivant les principes de l'architecture MDA. Plusieurs règles de transformation entre modèles y sont définies. Un prototype de cet environnement a été développé à l'aide d'un outil de méta-modélisation. La preuve de concept est sa capacité à reformuler des problèmes en ligne.

Abstract – Decision-making paths, frequency of decision-making, as much as their critical weight, have new prominences in today organizations. We are interested here in the decision-maker's capacity to include the characteristics specific to his/her ecosystem, its context and its often-changing issues in the way he/she practices. Decision-makers are often business experts to whom we want to help the design and adaption of their decision-making tools. The creation of a decision support framework using a Model Driven Engineering (MDE) approach is proposed. It seeks to link models of decision-making needs to logical models required by a combinatorial optimization solver. To do this, the functional and logical architectures of this framework are described on a case study, the traveling salesman problem. The assumptions and knowledge conveyed by the models all along the engineering chain are proposed. The design of the environment follows the principles of the MDA architecture. Several transformation rules between models are then explained. A prototype of this environment was developed using a meta modeler software. It provides a meaningful proof of concept, we mean the capacity to change problem formulations on the fly.

Mots clés - Ingénierie Dirigée par les Modèles, Aide à la décision, Gestion des opérations, Recherche Opérationnelle, Problème du Voyageur de Commerce, Outils de méta modélisation

Keywords – Model Driven Engineering, Decision making, Operations Management, Operations Research, Travelling Salesman Problem, Meta modeler.

1 ANALYSE DES ATTENTES EN AIDE A LA DECISION

Nos travaux s'inscrivent dans le faisceau des transitions sociétale et numérique des systèmes de production de biens ou de services, et en particulier le mouvement de fond que constitue Industrie 4.0. Nous le revisitons sous l'angle d'une évolution fournie par le numérique sur la capacité décisionnelle des systèmes.

Actuellement, dans une société connectée à de multiples échelles territoriales, une forte pression s'exerce sur les organisations productives pour régulièrement coordonner leurs activités et piloter leurs flux de produits et services. Poursuivre les bons objectifs est toujours le cap, mais ces objectifs sont souvent évolutifs (accès aux marchés, maîtrise de performance, réduction des risques, ...) et portent sur des périmètres qui sont à l'échelle de réseaux collaboratifs ce qui dresse parfois des

barrières en termes d'autonomie de décision. Ce constat vaut pour de multiples domaines d'application, citons la chaîne logistique manufacturière ou encore la gestion de tournées de professionnels libéraux dans les services à la personne, à titre d'exemples. Le caractère stratégique de tels sujets ouvre nécessairement un débat récurrent sur la culture du pilotage dans sa propre organisation et sur l'aptitude à répondre aux sollicitations des multiples réseaux dans lesquels elle s'inscrit. C'est là que se trouve une partie de l'art du management. Dans ce contexte, l'amélioration continue qui irrigue nos démarches qualité n'est pas seulement confinée dans les processus opérationnels et supports, elle impacte les processus de pilotage. Elle interroge sur les modes de pensée, les activités de prises de décision, les pratiques et les outils d'aide à la décision.

Nous savons de longue date que les avantages concurrentiels, tout comme la quête de valeur ajoutée, ne peuvent prendre une réalité tangible que par un savoir-faire en aide à la décision. Le génie industriel a joué un rôle important dans cette prise de conscience. Par exemple, la planification ne fait pas que traduire une volonté stratégique, c'est une fonction qui peut devenir une arme stratégique. Rappelons si besoin est l'importance donnée à la tension des flux, au juste à temps, ou encore à la satisfaction des usagers ou des agents de production dans les travaux de recherche en gestion des opérations. C'est la raison pour laquelle cette planification est abordée avec une formulation « maison », tenant compte de spécificités locales dans une organisation donnée.

La tendance est à introduire de plus en plus de connaissances métier dans la prise de décision.

Enfin, la société est devenue une société de l'exigence immédiate. Les circuits d'approvisionnement, de production et de distribution ne doivent pas seulement poursuivre des critères de performance économique et sociale, ils doivent accepter des changements de la demande et savoir y répondre. L'agilité est devenue un maître-mot. Mais le poids de ces incertitudes dans les espaces de prise de décision conduit inexorablement à plus de flexibilité dans les usages de différents outils de calcul et invite à développer un nouvel esprit de continuité des activités de prise de décision, prolongeant ainsi logiquement le concept de continuité des activités opérationnelles.

Pour toutes ces raisons, un décideur appréciera d'avoir le contrôle sur la formulation de ses modèles de prise de décision et de pouvoir faire bon usage de cette capacité de contrôle lorsque le contexte est très dynamique. Cela est d'autant plus factuel que beaucoup d'organisations ont gagné en maturité dans cet exercice. Toujours en planification, ce sont les progrès enregistrés grâce à la diffusion du MRP II (Manufacturing Resource Planning II) et des facilités prodiguées par l'usage courant des progiciels de gestion intégrés qui sont à l'origine de cela. Le progiciel est paramétré pour représenter les processus relatifs à une organisation connue de la production. Modifier les paramètres est une opération lourde et à risque. D'autres outils se sont imposés dans le paysage de l'aide à la décision. Ce sont ceux qui sont mobilisés par les experts et chercheurs en aide à la décision. Ils leur offrent des services numériques encapsulant des techniques de recherche opérationnelle, appelés solveurs. Beaucoup de solveurs sont aujourd'hui pris en main en quelques heures et équipés de bibliothèques d'algorithmes dits classiques qui permettent de travailler par réutilisation de code existant.

Ces deux catégories d'outils encadrent un périmètre dans lequel le décideur en gestion des opérations évolue actuellement. D'un côté des outils accédant facilement aux données d'entreprise, mais pour lesquels les formulations de problème sont relativement fermées, même s'il existe quelques paramétrages possibles. De l'autre, des outils ouverts qui permettent de reformuler les problèmes d'aide à la décision, de faire face à une certaine complexité, mais qui requièrent une connaissance plus fine de la recherche opérationnelle, sans être vraiment branchés en ligne sur les données de l'organisation. A ce stade de notre analyse, l'idée a germé d'étudier une troisième voie ouvrant de nouvelles perspectives de travail pour le décideur.

Une analyse bibliographique nous a montré qu'il n'existe pas à ce jour de travaux de recherche équivalents sur une application de l'Ingénierie Dirigée par les Modèles (IDM) allant de la génération personnalisée de modèle d'aide à la décision à leur résolution passant avec un couplage systématique. Notre travail est donc prospectif, nous ouvrons une voie.

Cette nouvelle voie ne dispose pas de la maturité et de l'expérience des deux autres, évoquées ci-dessus. Par contre, il cherche à ouvrir au maximum les possibilités de maîtriser l'instruction des problèmes d'aide à la décision à partir de connaissances métier et sans à avoir à traiter directement la conception de code pour l'appel au solveur.

Ainsi, nous émettons l'hypothèse qu'avec le temps, le verrou scientifique se déplace d'un savoir-faire factuel en aide à la décision vers un besoin nouveau de prise de décision rapide. Les outils doivent être conçus pour faire face à la variété des situations à traiter et pour pouvoir prendre des décisions circonstanciées en contexte de pression temporelle (ce qui revient à considérer que le nombre d'interlocuteurs dans la chaîne décisionnelle est réduit).

L'objectif poursuivi est donc de chercher comment donner au décideur non expert en recherche opérationnelle la faculté de faire une spécification en ligne et à moindre effort de ses problèmes, puis de les résoudre avec les bonnes données.

Une telle ligne de pensée invite à revisiter la place de la décision dans les modèles d'entreprise. Car cette ambition ne peut s'affranchir de la quantité et de la qualité des connaissances métier représentatives de la réalité du fonctionnement de l'organisation. Elles sont la source du raisonnement d'un décideur. Or celles-ci étant nombreuses, une solution viable ne peut pas s'imaginer sous la seule forme de l'émancipation d'une solution technique existante, aussi riche soit elle. Cette émancipation se heurterait à des interfaces trop sophistiquées quand il faudrait mobiliser ces connaissances métier et les données afférentes. Il en découle une exigence qui est de bien structurer le passage de la connaissance métier à la connaissance logique dont un solveur a besoin pour réaliser des calculs afin de trouver des solutions. A cette fin de prendre une décision, conscients de l'impératif d'utilisation structurée des connaissances métier, nous proposons une nouvelle option dans la pratique d'un décideur avec un environnement d'aide à la décision dont le fonctionnement relève des théories de l'ingénierie dirigée par les modèles (IDM).

Puisque cette pratique exigera de conjuguer une relation subtile entre (1) les connaissances et l'expertise métier et (2) la capacité à décider, c'est d'une logique de flux de connaissances entre modèle orienté métier, modèle mathématique d'aide à la décision et modèle logique de calcul de solutions dont il sera question dans cet article. Notre posture consiste à étudier un enchaînement de ces trois modèles pour relier par voie numérique l'espace de formulation de problèmes et l'espace de résolution

Notre recherche vise d'abord à imaginer l'architecture fonctionnelle d'un tel environnement de travail. Du fait du caractère très exploratoire du sujet, nous l'expliquons au moyen d'un cas d'études très simple, à vocation uniquement pédagogique. Il s'agit de différentes variantes du fameux voyageur de commerce. Ces variantes forment autant de problèmes différents sur lesquels nous avons raisonné. Il constituera un fil rouge dans la suite de cet article. Cette architecture fonctionnelle est détaillée au paragraphe 2. Cette présentation nous conduit à discuter les différents sens donnés au terme de modèle pour chacun des trois modèles qui sont utilisés par l'IDM. Puis, au chapitre 3, nous montrons comment différents mécanismes de transformation entre ces modèles permettent de rendre assez systématique le passage de la connaissance métier à la connaissance logique. Nous démontrons la faisabilité de cette proposition en faisant une preuve de concept sur notre prototype au chapitre 4. Enfin, nous tirons les enseignements de cette expérience au chapitre 5.

2 DES MODELES POUR ENCADRER LES VARIANTES DE PROBLEMES DE PRISE DE DECISION

Les problèmes de routage et d'ordonnancement sont des problèmes de prise de décision. Ils nécessitent la détermination de séquences optimales d'activités opérées par des ressources qualifiées. L'optimalité se traduit par des critères d'appréciation à la source de l'écriture d'une fonction objectif. L'ordre entre les activités et le choix des ressources allouées sont assujettis à la satisfaction de règles de fonctionnement qui forment un jeu de contraintes pesant sur l'espace des solutions possibles.

Le problème du voyageur de commerce (abréviation de sa traduction anglaise : TSP) est une classe connue dans l'ensemble de ces problèmes. Rappelons simplement son énoncé le plus répandu. Un voyageur doit visiter une et une seule fois un ensemble donné de villes. Sa tournée est conçue pour le satisfaire, au sens où le TSP est souvent présenté comme un problème de minimisation de la distance parcourue par le voyageur. Cette forme du plus court chemin est un problème NP complet. La figure 1 présente un cas d'études simple, à titre d'illustration de notre théorie. Un ensemble de quatre villes (Paris, Lyon, Nice, Toulouse) sont à visiter. Les distances entre ces villes sont connues. Les données de distance fournies permettent de faire le trajet dans les deux sens entre chaque paire de villes. Il n'y a donc aucun interdit en termes de connectivité entre les villes pour ce voyageur.



Figure 1- Le contexte de notre cas d'étude du TSP

Il existe de nombreuses variations dans la prise de décision autour du TSP. Comme le précisent (Gutin & Punnen, 2006), l'état de l'art couvre un ensemble d'au moins dix configurations différentes. Evoquons de manière synthétique les sources de ces changements de problème en trois parties :

- 1- La fonction traduisant l'objectif à optimiser peut varier pour traduire des moyens de transport en compétition avec des temps de transport différents, voire des coûts de transport.
- 2- Les règles de fonctionnement peuvent changer. Ainsi dans une forme d'évolution, chaque ville est associée à un poids non négatif. La tournée doit alors trouver un ensemble minimal de villes à visiter pour lequel la somme des poids est supérieure à un seuil donné. Dans le même ordre d'idée, le fonctionnement peut relaxer la notion de visite unique en spécifiant au moins une visite par ville. Ou encore, l'alternative à une tournée unique de plusieurs tournées partielles, indépendantes et complémentaires, conduit à des décisions différentes.
- 3- La nature du problème peut changer quand on touche aux hypothèses fondamentales. Ainsi, dans une forme clustérisée du TSP, plusieurs sous problèmes coexistent avec une recherche de coût minimal dans un ensemble des

villes partitionné en plusieurs classes indépendantes. Ce même problème peut se décliner dans un autre but, cherchant une seule ville dans chacune de ces classes. Les auteurs citent aussi une publication où les classes de villes ne sont pas totalement disjointes. Enfin, la couverture des villes peut être imaginée en mobilisant plusieurs voyageurs partant de la même ville.

En recherche opérationnelle, une solution traditionnelle pour résoudre un problème de décision de type TSP est de passer par une formulation mathématique en optimisation sous contraintes. Il s'agit de poser les termes d'une programmation linéaire en nombre entiers. Cette écriture repose sur un graphe orienté traduisant les connaissances métier. Un point symbolise une ville à visiter et un arc représente le trajet entre deux villes. Les variables entières sont des booléens associés aux arcs. La solution se traduit par des variables nulles (=0) pour les arcs non retenus au final, les autres forment la tournée choisie et auront une valeur strictement positive (=1). Sur cette base, la formulation de la fonction traduisant l'objectif qualifie les performances attendues, et les relations de contraintes entre les variables traduisent les règles de fonctionnement. Cette première partie des variations est donc significative d'une nouvelle écriture de la fonction à optimiser. La modification dans l'écriture d'une contrainte, l'ajout ou la suppression d'une contrainte sont les changements induits par la deuxième partie. Dans la troisième partie, le changement est plus important. Il concerne tout à la fois les données, la fonction à optimiser et les contraintes.

Toutefois, nous avons choisi d'aborder sur un même front toutes ces sources de changement du TSP. Ainsi, les variations évoquées devraient pouvoir être traitées dans un même environnement de travail. Ce faisant, et ayant posé une première hypothèse sur la formulation mathématique du problème à traiter, nous devons intégrer les conséquences de celle-ci sur la recherche de solutions. Nous supposons qu'un solveur en programmation linéaire en nombres entiers est disponible au sein de l'environnement de travail. L'appel de ce solveur se fait par un langage spécifique à ce solveur qui permet de lui transmettre les données et les relations mathématiques dont il a besoin.

A ce stade de l'exposé et au moyen du cas d'études, nous avons donc décrit littéralement la nature des décisions à prendre sur un système qui est connu. Ce volet constitue les connaissances métier. Nous allons lui associer un modèle métier. Puis, nous avons expliqué comment traduire ces connaissances en un problème mathématique bien formulé en parlant de données, de variables, d'objectif et de contraintes. Cet ensemble mobilise un vocabulaire de recherche opérationnelle. Ce sont ces connaissances mathématiques que nous consignons dans un modèle d'aide à la décision. Enfin, pour prendre la décision, l'espace des solutions du problème mathématique est parcouru par un solveur qui cherche à assurer la solution, faisable avec un objectif poursuivi optimisé. L'appel du solveur est aussi associé à un modèle que nous appelons modèle logique de calcul des solutions.

Nous décrivons donc le mécanisme de l'aide à la décision au moyen d'une chaîne de connaissances qui évolue depuis un stade préliminaire d'analyse des besoins métier jusqu'à une proposition de décisions satisfaisantes au regard de ces besoins. Trois formes de modèles jalonnent donc séquentiellement cette chaîne. Cette logique d'enchaînement correspond à un processus cognitif que l'expert en recherche opérationnelle exécute. Nous le résumons sur la figure 2. Il est composé de quatre étapes : définir le contexte, exprimer les besoins de prise de décision, identifier et formuler le problème d'aide à la décision, préparer le calcul de la solution. A l'issue de chaque étape, le raisonnement a produit un modèle qui constitue l'entrée de l'étape suivante.

C'est la façon dont ce processus linéaire s'exécute durant tout l'acte de prise de décision.

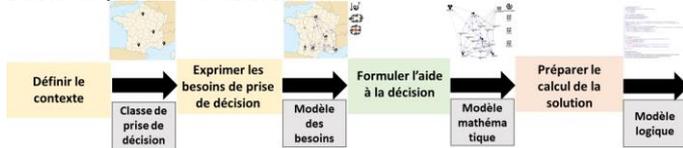


Figure 2- Modèle du processus cognitif de prise de décision

Notre environnement de travail est calqué sur ce processus cognitif. En cela, il cherche à l'explicitier et à le soutenir. Pour faciliter une vue d'ensemble de son architecture fonctionnelle, nous le décrivons en analogie avec une chaîne de production industrielle. Une chaîne de production est « l'ensemble des moyens qualifiés pour opérer une fabrication de produits manufacturés dans un volume et un temps maîtrisés. Elle s'étend des stocks de matières premières jusqu'à ceux de produits finis ». En plaquant cette définition sur notre sujet, nous identifions des stades d'avancement de la fabrication qui va aller de la définition du système et de ses besoins de prise de décision à la production d'un résultat par un logiciel de recherche de solutions appelé en toute connaissance de cause. Nous faisons reposer l'architecture fonctionnelle sur une approche modulaire. Une série de modules assimilés à des ateliers de transformation réalisent les transformations de connaissance. Chaque module correspond à une étape du processus cognitif. Sauf qu'ici, ce n'est pas un bien matériel qui est concerné, mais quelque chose de plus symbolique : une connaissance contenue dans un modèle. La transformation de la connaissance sera effectivement assimilée à une transformation de modèles. Chaque atelier a son propre objectif de transformation. Les flux entre ateliers sont des flux de modèles.

Attention, ce caractère un peu mécanistique des transformations de connaissance ne doit pas laisser supposer que l'humain est hors de la boucle. A chaque étape, il est possible de solliciter l'humain pour apporter les connaissances métier complémentaires en cas de blocage des règles de transformation. C'est une approche classique en IDM, l'apport est normalement minimisé, mais pas forcément inexistant.

Car, en IDM, par principe, l'usage fait d'un modèle quitte la vocation descriptive ou prescriptive qui lui donne la communauté du génie industriel. Il entre dans un rôle très opératoire parce qu'il est une source et un résultat d'un traitement informatique. Nos trois formes de modèles deviennent des artefacts d'ingénierie de la connaissance. Et en filant la métaphore, chaque atelier est équipé d'une ou plusieurs « machines » qui sont des services numériques opérant des traitements à base d'algorithmes, consommant et produisant des flux de connaissance via ces modèles. Tout au long du cycle de vie de l'ingénierie, le contrôle à exercer sur cette chaîne, module par module, correspondra aux règles régissant une transformation en IDM, de modèle à modèle (M à M).

Reprenons l'exemple du TSP pour illustrer le contenu de ces modèles et également leur variété. Les quatre flux de modèles entre les ateliers sont le modèle du contexte, le modèle des besoins de prise de décision, le modèle mathématique d'aide à la décision, le modèle logique de calcul. Dans le cas du TSP, ils seront décrits au prochain paragraphe. Nous choisissons de les présenter dans l'ordre inverse de ce flux. Le type de TSP choisi ici est un calcul cherchant la distance minimale à parcourir par un voyageur qui souhaite faire une seule tournée incluant une et une seule visite par ville.

Puisque nous avons défini le contexte d'un problème de routage et ordonnancement en faisant ce choix d'illustration, la

première étape du processus de la figure 2 n'a pas d'objet autre que de référencer un besoin à exprimer dans cette classe de prise de décision avec un langage métier adapté à ce contexte.

a- Modèle logique de calcul de la décision

Le solveur que nous avons retenu pour notre prototype est IBM CPLEX¹ qui présente un avantage déterminant en IDM, il met à disposition un langage, appelé OPL (pour Optimisation Programming Language) pour que l'utilisateur définisse les problèmes à traiter. Le langage OPL a une structure classique de langage de programmation avec une partie déclarative et une partie exécutable incluses dans un fichier (.mod). Les données d'un calcul sont également renseignées dans un fichier de données (.dat).

Un exemple du code est fourni sur la figure 3. On y distingue une partie déclarative où données (*int*, *range*, *tuple*, *setof*, *int*, lignes 1 à 10) et variables (*dvar*, lignes 11&12) sont distinguées de manière abstraite. Puis, il possède une partie exécutable où les relations explicitant la fonction (*minimize*, lignes 15 à 17) et les contraintes sont définies à l'aide de mots clés (*subject to*, lignes 18 à 25). L'écriture en OPL est vraiment très proche des formulations mathématiques traditionnelles recommandées par l'*American Mathematical Society*, et que la recherche opérationnelle pratique avec rigueur.

L'appel que nous faisons de ce solveur nécessite de fournir ces deux fichiers. La figure 4 décrit le fichier des valeurs des données déclarées. Il est facile de repérer les caractéristiques de notre cas d'études avec les 4 villes et la spécification des distances entre les villes est faite par une liste en bijection avec la liste des arcs.

```

1 int totalNumberOfCityNode=...;
2 int beginNodeIndex=...;
3 range cityNodeSet=beginNodeIndex..totalNumberOfCityNode;
4 tuple routingEdge{
5   int startNodeID;
6   int endNodeID;
7 }
8 setof(routingEdge) totalEdgeTuple=...;
9 int distanceMatrix[totalEdgeTuple]=...;
10 int travellingTimeMatrix[totalEdgeTuple]=...;
11 dvar boolean dvarAssignment[totalEdgeTuple];
12 dvar int+ dvarNoSubTour[cityNodeSet];
13
14
15 dexpr int totalDist = sum ( <i,j> in totalEdgeTuple)distanceMatrix[<i,j>] * dvarAssignment[<i,j>];
16 dexpr int totalTime = sum ( <i,j> in totalEdgeTuple) travellingTimeMatrix[<i,j>] * dvarAssignment[<i,j>];
17 minimize totalDist+totalTime;
18 subject to {
19   forall(i in cityNodeSet: i>1, j in cityNodeSet:j < i) && j!=i )
20     dvarNoSubTour[i] - dvarNoSubTour[j] + (totalNumberOfCityNode - 1) * dvarAssignment[<i,j>] <= totalNumberOfCityNode - 2;
21   forall( j in cityNodeSet, <i,j> in totalEdgeTuple )
22     sum( i in cityNodeSet: i!=j ) dvarAssignment[<i,j>]=1;
23   forall( i in cityNodeSet, <i,j> in totalEdgeTuple )
24     sum( j in cityNodeSet:j!=i ) dvarAssignment[<i,j>]=1;
25 }
26
27
28 main {
29   thisOpModel.generate();
30   cplex.solve();
31   var ofile = new IloOplOutputFile("D:\Users\lilien.zhang\Desktop\tspOplResult.txt");
32   ofile.write(thisOpModel.totalTime);
33   ofile.write(thisOpModel.totalDist);
34   ofile.writeIn(cplex.getSolvedTime());
35   ofile.close();
36   var ofCSV = new IloOplOutputFile("D:\Users\lilien.zhang\Desktop\tspAssResult.csv");
37   for(var i in thisOpModel.dvarAssignment){
38     ofCSV.writeIn(i.startNodeID+","+i.endNodeID+","+thisOpModel.dvarAssignment[i]);
39   }
40   ofCSV.close();
41 }

```

Figure 3 – Modèle logique de calcul en langage OPL – Partie code du modèle (fichier .mod)

```

1 beginNodeIndex=1;
2 totalNumberOfCityNode=4;
3 totalEdgeTuple={<1,2><1,4><1,3><2,4><2,1><2,3><4,2><4,1><4,3><3,1><3,2><3,4>};
4 distanceMatrix=[550,175,750,115,600,450,135,230,200,800,430,180,];
5 travellingTimeMatrix=[0,0,0,0,0,0,0,0,0,0,0,];

```

Figure 4 – Modèle logique de calcul en langage OPL - Partie données du modèle (fichier .dat)

b- Modèle mathématique d'aide à la décision

La formulation du problème de TSP repose sur un graphe orienté duquel sera déduite la formulation mathématique. Pour les besoins de notre environnement, ce graphe est annoté de

¹ Solveur programmation mathématique : <https://www.ibm.com/analytics/cplex-optimizer>

toutes les connaissances dont il sera fait usage pour générer le modèle logique de la figure 3.

Ainsi, le graphe présenté sur la figure 5 est enrichi par ces annotations qui permettent de mémoriser la structure du graphe au niveau de ses arcs (routing edge). Les points sont numérotés à cette fin. Les ensembles homogènes sont aussi présentés (CityNodeSet, routingEdgeSet) et sont reliés à leurs éléments par des lignes pointillées. Enfin, les objets qui sont positionnés sur la droite correspondent à des relations mathématiques dont l'évaluation mobilise des éléments du graphe. On y trouve la fonction à minimiser (min. travel distance), des ensembles de relations portant sur les entrées et sorties de ville (flow in, flow out) qui serviront à brider le nombre de visites autorisées, et un ensemble de relations contraignant la solution à une seule et unique tournée (No Subtour).

A ce stade de développement de notre prototype, nous n'avons pas fait appel à des outils d'écriture formelle, comme Maple, par exemple, pour présenter ces objets mathématiques. Nous avons choisi de mobiliser une bibliothèque de formules pré-écrites où ces relations sont codées. Ce n'est qu'un stade intermédiaire qui nous permet d'entreprendre les premiers tests. La structure modulaire de l'environnement permettra de faire ultérieurement des développements plus détaillés de cette partie. Nous gardons à l'esprit une évidence, l'écriture des contraintes est un savoir-faire et il sera difficile de prétendre à en faire une génération systématique. La philosophie d'une récupération de l'existant n'est qu'une première voie pragmatique que nous retenons.

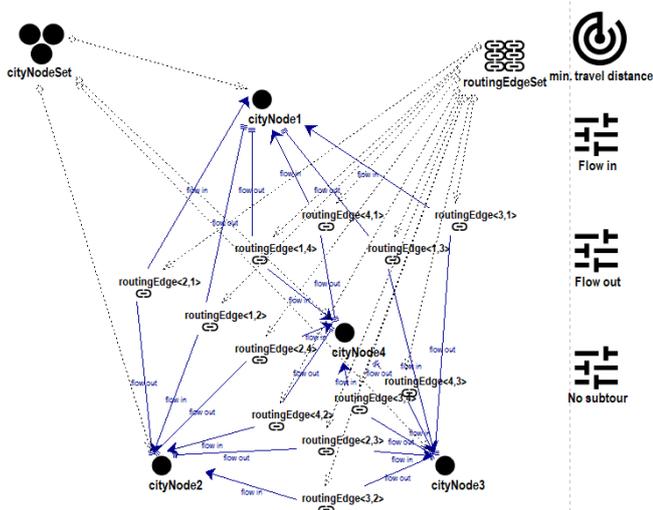


Figure 5 – Modèle mathématique d'aide à la décision inspiré de la théorie des graphes

c- Modèle des besoins de prise de décision

L'utilisation de langage spécifique à un domaine (DSL) a tendance à croître. Elle est une voie différente de la voie classique mobilisant des langages de modélisation d'entreprise, souvent promus par des logiciels BPM. Afin d'éviter toute surcharge en termes de syntaxe de langage, mais aussi avec le souci d'avoir tout degré de liberté sur le plan technique pour écrire et utiliser les connaissances portées par un tel modèle, nous avons retenu cette idée de mettre progressivement au point un langage spécifique au domaine de la prise de décision dans un contexte connu. En faisant cela, nous marchons sur les traces de la méthode GRAI (Graphs with Results and Actions Inter-related) (Chen et al., 1997) qui a de longue date gravé dans le marbre le concept de centre de décision appliqué à la gestion de production.

Récemment, l'OMG a publié la spécification du langage DMN (Decision Model and Notation) (Biard et al., 2015) qui permet

d'enrichir BPMN 2.0 (Allweyer, 2016). Dans la mesure où DMN est encore peu adapté à une prise de décision en gestion des opérations, mais qu'il vise plus à automatiser la gestion de flux d'informations dématérialisés, nous ne nous en sommes pas inspirés. En effet, si ce langage propose des concepts de type critère et variable de décision, il ne propose pas de primitives spécifier explicitement des contraintes de type égalité ou inégalité dans son formalisme de base (Biard, 2015).

Il faut souligner ici que le DSL dépendra vraisemblablement du type de problèmes abordés dans une classification qui est celle de la recherche opérationnelle. C'est cette reconnaissance préliminaire qui doit conduire à sélectionner un DSL adapté à la description du besoin de prendre un certain type de décisions selon le contexte (première étape du processus cognitif de la figure 2). Ici, le contexte est celui d'un problème de routage et d'ordonnancement, le DSL propose des éléments de syntaxe concrète pour aider à représenter ce besoin. Dans un premier temps, nous avons surtout abordé des décisions de nature opérationnelles, les choix tactiques restent un chantier à ouvrir qui dépendait des résultats acquis dans ce premier temps.

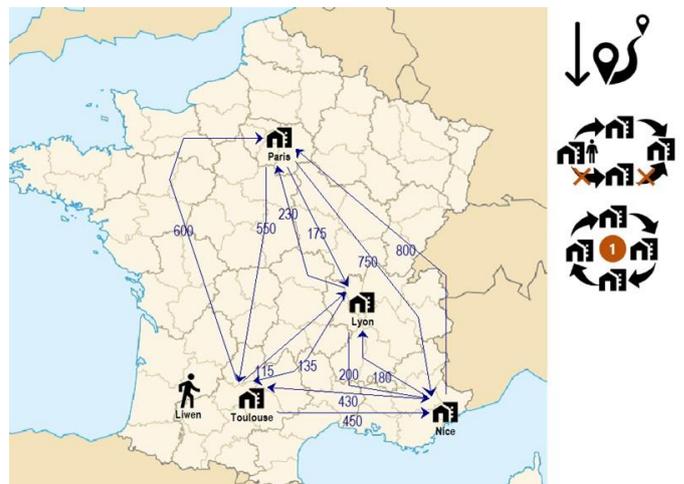


Figure 6 – Modèle des besoins de prise de décision inspiré du centre de décision de la méthode GRAI

Sur la figure 6, la toile de fond est une carte géographique du territoire identique à la figure 1, où on peut positionner de manière réaliste les villes qui prennent la syntaxe concrète d'un symbole de type « maison » avec une sémantique claire (Paris, Lyon, ...). Les flèches qui relient les villes sont des associations de type connexion routière et portent l'attribut de la distance. Le voyageur est symbolisé par une icône « forme humaine en marche », avec le prénom de ce voyageur. Cette icône peut se déplacer, elle est toujours attachée à une ville et cette connexion n'est pas matérialisée graphiquement. Cette partie gauche sur la figure 6 permet de désigner les variables de décision, c'est-à-dire l'ordre dans lequel l'icône humaine va se déplacer pour parcourir les quatre villes. Et par conséquent, les routes qui seront empruntées par lui.

De nouveau, comme pour le paragraphe b, des éléments graphiques sont disposés à droite de la figure. L'objet en tête de liste précise qu'il faut trouver une route entre chaque paire de villes avec une distance minimale à parcourir. Le suivant explique qu'il s'agit d'un flux traversant une et une seule fois toute ville. Le dernier explique qu'il ne fera qu'une et une seule tournée. Ces objets forment le corps explicite du centre de décision (variables, objectif et contraintes).

3 PRINCIPE DES TRANSFORMATIONS DE MODELES

Nos modules au sein de cette architecture produisent des modèles. Mais comment se réalise cette production ? L'IDM a développé une théorie qui permet d'emprunter cette voie et de réaliser de tels pontages entre modèles. Les modèles y ont été d'abord considérés comme des artefacts techniques peuplant le cycle de vie de l'ingénierie logicielle (c'est-à-dire à un niveau applicatif, nous ne sommes plus au niveau fonctionnel). Les modèles prennent ce caractère opérationnel quand ils permettent d'élaborer et de contrôler la production de contenus informatiques. Le génie logiciel a utilisé ces techniques pour la génération de codes d'haute-fidélité sur des systèmes critiques (référence au langage Lustre, par exemple). Mais dans le cadre d'une architecture orientée services, il est aussi possible de maîtriser une exécution conforme à une spécification donnée par des appels réalisés intelligemment grâce à des modèles de services numériques. En visant la production d'un code en langage OPL sur notre prototype, nous adoptons les mêmes démarches.

Ce sont des techniques dites de transformation des modèles (M to M) qui sous-tendent ces capacités de production logicielle. Une transformation des modèles fait d'abord recours à la méta-modélisation. Un méta-modèle est en quelque sorte le moule formel du modèle. Selon (Benaben, 2012), il est : « un modèle qui définit le langage utilisé pour exprimer un modèle ». (Chapron, 2006) est en phase avec cette vision quand il définit un méta-modèle comme : « une description de tous les concepts d'un langage. La méta-modélisation permet de définir directement les concepts manipulés, la sémantique et la syntaxe du langage liée à l'utilisation de ces concepts ». Cette vision est confirmée par (James et al., 2013) et (Karagiannis et al., 2016). Les méthodes de transformation des modèles ont fait l'objet de nombreuses études. (Czarnecki & Helsen, 2003) proposent une classification des approches générales existantes. Proche de notre problématique, (Cuadrado et al., 2014), en mobilisant le langage de transformation ATL, transforment un modèle source qui est un modèle de traitement dématérialisé de l'information (un *workflow* spécifié avec plusieurs options de langage possibles : diagrammes d'activité UML, YAWL et BPMN) en un Réseau de Petri. (Gómez et al., 2018) expliquent comment transformer un modèle physique exprimé en Modelica pour produire le fichier de lancement de sa simulation dynamique. L'environnement EMF (Eclipse Modeling Framework) est utilisé par ces auteurs comme cadre supportant la transformation.

L'OMG a proposé une structuration de l'avancement de processus d'ingénierie dirigée par les modèles, gradué en trois niveaux de modélisation quand on passe du besoin à la solution :

- Le CIM pour Computer Independent Model,
- Le PIM pour Platform Independent Model,
- Le PM pour Platform Model, et le PSM pour Platform Specific Model.

Cette représentation porte le nom d'architecture MDA (Model Driven Architecture) (Mellor et al., 2004). Chaque modèle est écrit avec un langage approprié.

Les problèmes d'interopérabilité entre partenaires impliqués dans différentes configurations de réseaux de collaboration ont été traités selon ces principes architecturaux. Nous citons l'ingénierie de système d'information de médiation comme un domaine où la conception part de modèles métier de type CIM (Mu et al., 2017). Le PIM y représente les connaissances techniques par un modèle de processus en BPMN. Enfin, le PSM est une implantation dans une architecture de services avec une technologie de bus d'entreprise. En amont d'un tel problème, la conception même du réseau de collaboration peut également mobiliser une structure de connaissances métier supplémentaires apportées par des ontologies (Wang et al., 2018).

Dans notre étude, nous avons suivi les mêmes raisonnements :

- Le niveau CIM est une spécification de haut niveau d'une fonction de prise de décision à mettre en œuvre et véhiculant les connaissances métier. C'est le sens que nous donnons à notre modèle de besoins de prise de décision,
- Le niveau PIM est une étape intermédiaire allant vers l'exécution de la fonction décrite par le CIM et mobilisant des connaissances de recherche opérationnelle. Mais à ce stade, aucun environnement de développement n'a été choisi. Toutes les options restent ouvertes. C'est le sens donné à notre modèle mathématique pour l'aide à la décision.
- Le niveau PSM dicte l'implémentation du PIM dans un environnement d'exécution connu qui est nommé PM. En faisant le choix du solveur IBM CPLEX, et, nous instancions le concept de PM et donc du langage OPL. Le PSM correspond aux deux fichiers du modèle logique écrits en OPL.

A chaque étape dans la logique de fonctionnement de notre environnement, c'est-à-dire dans chaque module de son architecture fonctionnelle, une transformation de modèles va donc être opérée (passage CIM-PIM, puis PIM-PSM). En quelque sorte, le module devient un atelier où des opérations sont exécutées. Pour les comprendre, il nous faut d'abord définir les méta-modèles pour nos 3 types de modèles. A partir de cette base, une série de règles de correspondance entre les méta-modèles permet de faire ladite transformation, comme illustrée sur la figure 7.

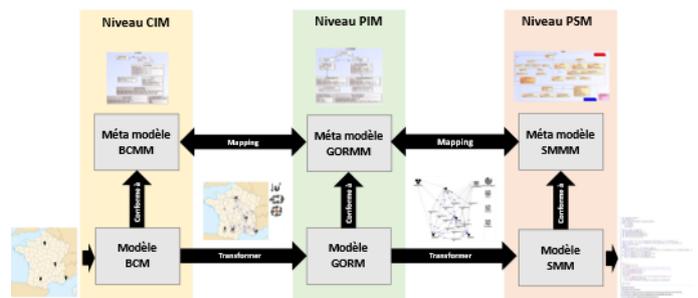


Figure 7 – Principe de la transformation de modèles en IDM

Dans l'écosystème de l'IDM, le principal ingrédient pour soutenir la transformation des modèles est le langage de transformation des modèles (MTL). Ainsi, (Agrawal et al., 2006) introduisent un langage basé sur une représentation graphique pour soutenir la mise en œuvre de transformation particulièrement efficaces qui assurent des traductions entre modèles, ou entre modèles et codes. VMTL est un MTL (Acrejoaie et al., 2018). ATL en est un autre, largement cité dans la littérature (Jouault et al., 2008).

Nous avons choisi la plate-forme de méta-modélisation ADOxx pour réaliser nos développements. Elle possède son propre MTL appelé AdoScript. Cette plateforme est largement utilisée pour construire des preuves de concept (PoC) dans les activités de recherche en ingénierie dirigée par les modèles. Elle a créé une communauté OMILAB (Fill & Karagiannis, 2013) qui partage des expériences et des applications ainsi développées dans une optique de réutilisabilité. Ayant rejoint cette communauté, nous avons pu dresser le constat d'une absence d'application, dans sa panoplie des projets finis ou en cours, ayant un voisinage thématique avec notre travail.

Toutefois, une recherche récente traite de l'usage potentiel de MTL dans la pratique (Burgueño et al., 2019), ainsi que des rôles actuel et futur d'un MTL en intelligence artificielle. Nous nous inscrivons dans la même veine de travaux.

4 APPLICATION SUR LE CAS D'ETUDES

Ce sont ces méta-modèles et ces transformations de modèles que nous décrivons dans ce paragraphe. Le contenu est représentatif du travail fourni pour construire un prototype expérimental dans l'outil ADOxx et en démontrant la pertinence sur notre cas d'études. C'est un mécanisme à base de règles qui régit le MTL.

Nous mobilisons trois types de règles de transformation, règle de « mapping », règle de navigation et règle de « parsing ». Une règle de mapping traite une mise en correspondance entre les deux méta-modèles. Elle pointe des concepts du méta-modèle source et décrit explicitement l'effet de leur combinaison sur des concepts du méta-modèle cible. Une règle de navigation opère des traitements locaux dans le méta-modèle cible pour créer des éléments nouveaux en exploitant des dépendances locales. Une règle de parsing permet de faire des modifications du modèle qui sont induites par des annotations dans le méta-modèle.

La description complète et détaillée des transformations dépasse largement la taille raisonnable de cet article et a, de plus, un caractère très technique. Nous avons donc décidé de présenter ces règles dans leurs grandes lignes.

• Transformation CIM-PIM

Le premier atelier transforme le modèle de besoin de prise de décision (niveau CIM), que nous nommons de manière synthétique BCM, en un modèle mathématique d'aide à la décision (niveau PIM) appelé GORM. La palette graphique du DSL utilisé pour élaborer le BCM est une expression directe de ce qu'est le méta-modèle (cf. figure 6). Nous le présentons de manière plus rigoureuse avec le premier méta-modèle à la source de la transformation, qui est consigné sur la figure 8. Nous l'avons appelé BCMM (Zhang et al., 2020).

Le résultat de la transformation est notre modèle cible GORM (cf. figure 5). Il doit être conforme au méta-modèle de la figure 9, appelé GORMM.

Ce sont des diagrammes de classe UML qui sont utilisés pour décrire ces méta-modèles.

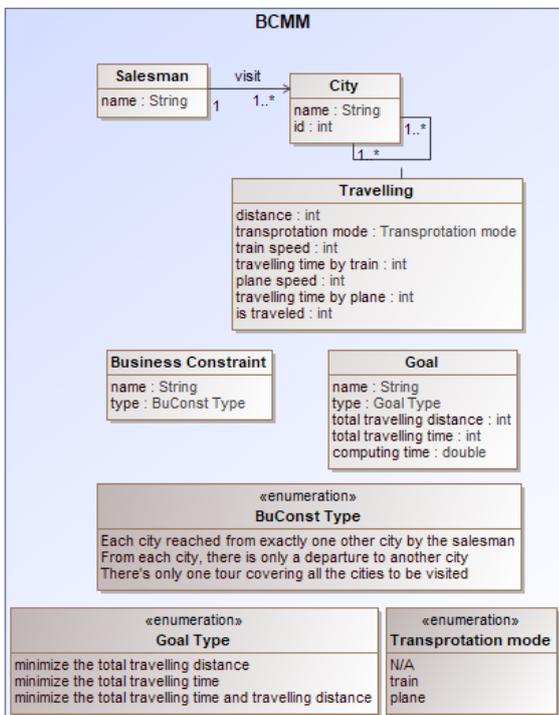


Figure 8 – Méta modèle BCMM (niveau CIM)

Algorithme de transformation CIM-PIM

Règles de mapping entre BCMM et GORMM

Travail sur les relations directes entre concepts

- Chaque ville dans le modèle métier devient un nœud dans le graphe. Son nom (id de City) devient un numéro (id de Node).

Travail sur les maillages de concepts

- Chaque trajet pouvant être emprunté par le voyageur (relation réflexive sur le concept City avec la classe association Travelling) devient un arc dans le graphe. Un arc est un lien dirigé entre deux nœuds. Il prend pour attribut la donnée qui sera utilisée pour poursuivre l'objectif (edgeLength).
- L'objet qui spécifie la traversée d'une et une seule ville dans le modèle métier est désagrégé. Il induit la désignation des arcs comme devenant des variables de décision avec deux indices (dvar Assignment).
- Le « flow in » n'autorisera qu'un trajet à être fait parmi tous les arcs qui entrent dans une ville. Il en va de même pour le « flow out » en sortie de ville.
- L'exigence d'une tournée unique induit la création de variables de décision permettant de tracer la tournée en cours de calcul et donc de contrôler l'absence de sous tournée (dvar NoSubTour).
- Chaque trajet est inducteur d'une contrainte qui n'est que pointée dans son existence à ce stade, mais qui servira à trouver la formulation mathématique associée dans la bibliothèque de formules. Il en va de même pour la tournée unique et pour le choix de la fonction objectif.

Règle de navigation dans le GORMM (niveau PIM)

- Regroupement des données portées par les concepts qui composent le graphe (CityNode et RoutingEdge) dans des ensembles homogènes (CityNodeSet et RoutingEdgeSet) qui seront ensuite utiles pour écrire le code logique dans sa partie déclarative.

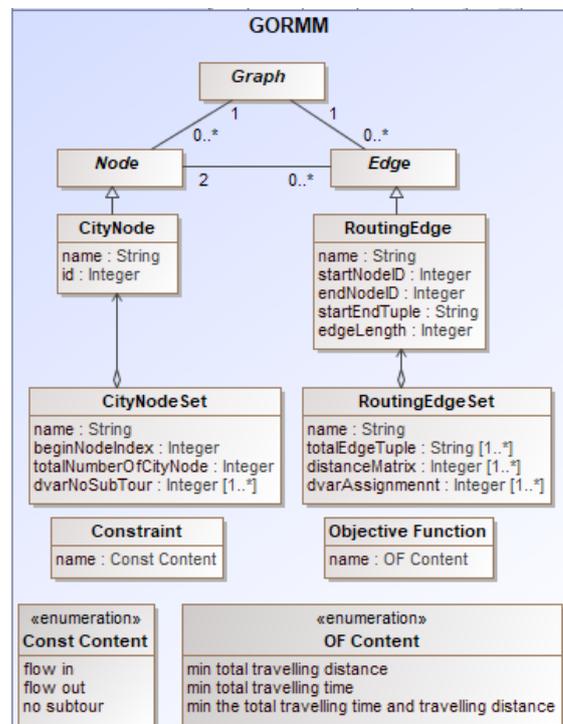


Figure 9 – Méta modèle GORMM (niveau PIM)

• Transformation PIM-PSM

Le deuxième atelier de transformation concerne la génération du modèle logique de calcul de la décision à partir du GORM. Ce modèle est appelé SMM. Nous présentons le méta-modèle

auquel il doit être conforme, appelé SMMM (Assouoko & Denno, 2016), sur la figure 10.

Algorithme de transformation PIM-PSM

Règles de mapping entre GORMM et SMMM

- Développement des types de données dans la partie déclarative du code OPL avec une distinction faite entre paramètres du modèle et variables de décision, chacun ayant ensuite son mot clé dans le langage OPL.
- Utilisation de pointeurs pour construire les lignes de codes des relations mathématiques de la partie exécutable à partir des patrons de formulation extrait d’une bibliothèque de formules en OPL (fonction à optimiser et contraintes).
- Une règle de transformation permet de générer des composants en instanciant des patrons pour lesquels les paramètres attendus sont renseignés à partir des données du GORM.

Règle de navigation dans le SMMM (niveau PSM)

- Assemblage des lignes de déclaration et des lignes d’exécution selon les règles de bonne construction d’un code OPL

Règle de parsing du SMM

- Ajout des mots clés du langage OPL aux parties déclaratives et aux parties exécutables

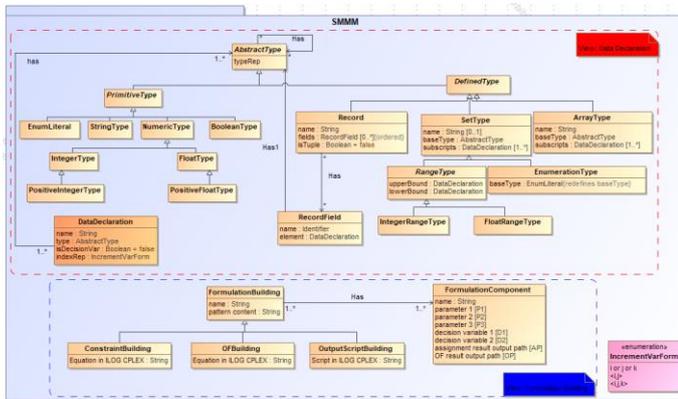


Figure 10 - Méta modèle SMMM (niveau PSM)

L’ensemble de la chaîne de production est résumé sur la figure 11.

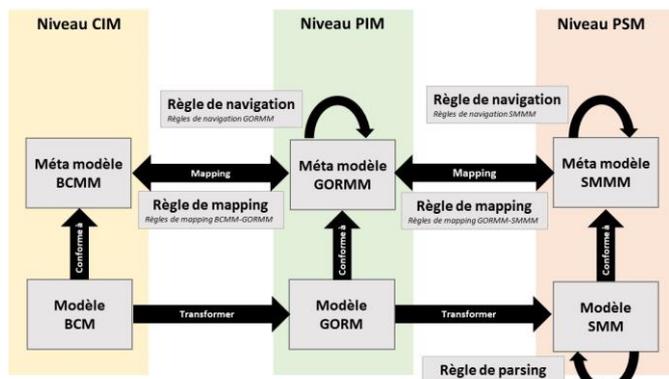


Figure 11- Schématisation de la chaîne de transformations des modèles du niveau CIM à celui du PSM

Le SMM permet de lancer le calcul sur IBM CPLEX. Avec la solution ainsi obtenue, les chemins retenus sont (1,4), (4,3), (3,2), (2,1), présenté sous le forme (id de ville de départ, id de ville d’arrivée). Il en découle une tournée (Toulouse, Paris, Lyon, Nice, Toulouse).

Un décideur ayant contribué à l’élaboration du BCM aura naturellement envie de retrouver des propositions dans le même univers de discours. Il appréciera certainement de disposer d’une interface homme-machine (IHM) afin d’obtenir de manière efficace et compréhensible des connaissances concrètes sur le TSP ainsi résolu. Nous avons donc débuté le développement d’un retour vers le modèle BCM pour visualiser ces résultats de manière graphique en changeant simplement la couleur des chemins retenus entre la série de villes formant la tournée.

5 CHANGEMENT DE PROBLEME

Afin de mettre à l’épreuve notre prototype développé sous ADOxx, nous lui avons soumis un autre problème TSP. Il s’agit maintenant d’une formulation où la fonction à optimiser est un temps minimal de voyage. Ce passage à une caractérisation de chaque trajet entre deux villes par un temps de voyage plutôt qu’une distance est accompagné par une interaction homme-machine dans la conception du BCM. Cela permet de spécifier le mode de transport employé pour chaque trajet. Ainsi, nous avons mixé les modes de transport en choisissant l’avion pour les chemins de plus longue distance, c’est-à-dire entre Paris et les villes du sud, et le train pour les autres. Sur la figure 12, l’icône de la fonction à optimiser a changé. De plus, la contrainte de la tournée unique a été supprimée.

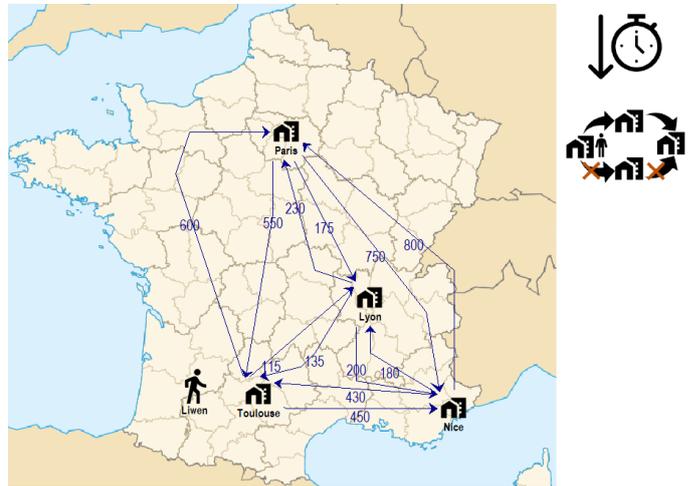


Figure 12- Spécification d’un TSP différent

Le résultat est présenté sur la figure 13. Nous avons maintenant deux tournées distinctes. Une se fait en avion sur l’aller-retour Paris Toulouse, l’autre se fait en train sur l’aller-retour Lyon-Nice.

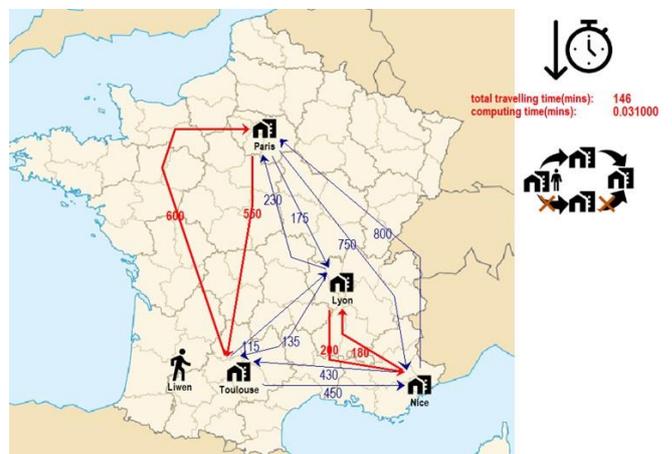


Figure 13- Spécification d’un TSP différent

Les deux modèles intermédiaires GORM et SMM ont été contrôlés et ont permis de valider notre mise en œuvre des règles de transformation.

6 CONCLUSIONS ET PERSPECTIVES

Nous avons présenté deux niveaux d'architecture de notre environnement d'aide à la prise de décision. La couche fonctionnelle est inspirée du processus cognitif de l'expert en recherche opérationnelle, décrit en en quatre phases. La couche logique implémente ces fonctions attendues et a permis de décrire une suite des modules de traitement de trois modèles dans le respect des principes du MDA (CIM, PIM et PSM).

Le développement du prototype sous ADOXX a été réalisé dans le contexte d'un langage spécifique à un problème de routage et ordonnancement. La preuve de concept est obtenue à partir d'une variation du problème du voyageur de commerce. L'implémentation de cette idée d'une « usine » de la connaissance sur le problème du voyageur de commerce (TSP) prouve la faisabilité d'une telle chaîne IDM en pratique.

Rappelons que notre but était d'explorer cette voie et de défricher un terrain vierge, celui de l'emploi de l'IDM dans les problèmes de prise de décision. Nous avons fait de nombreuses hypothèses dans cette prospection que nous avons soulignées dans cette présentation (DSL, décision opérationnelle, bibliothèque de formules, solveur). Nous nous attacherons à les lever progressivement pour donner plus de capacité à cet environnement de prise de décision, et c'est la perspective des chantiers à venir.

Le test de sensibilité des modèles à une modification du problème à traiter reste une piste à explorer afin d'estimer l'impact sur les différents stades du processus de la figure 2 et donc la capacité de l'IDM à faire face à la demande de changement du problème à résoudre. Ne perdons pas le cap.

Fort de cette première avancée, la suite de notre travail s'inscrira aussi dans un cadre plus difficile à traiter. Il s'agit du problème de coordination opérationnelle d'une prise en charge à domicile de patients. Ce problème est connu et intitulé « *Home HealthCare Routing and Scheduling Problem* » (HHCRSP) dans la communauté de recherche opérationnelle. Il est lui aussi sujet à de nombreuses variations liées à la différence entre les organisations en charge de coordination, leurs pratiques, leur territoire de chalandise, leur réseau d'intervenants, les moyens de transport, les horaires de travail, etc. Nous avons choisi d'étendre les compétences de notre prototype sur ce nouveau cas d'études. Nous avons produit un modèle mathématique d'aide à la décision qui est en fait un TSP clustérisé multi-voyageurs. Pour cette nouvelle étude, une partie de ce que nous venons de présenter est *de facto* réutilisable.

7 REFERENCES

- Acretoiaie, V., Störrle, H., & Strüber, D. (2018). VMTL: A language for end-user model transformation. *Software & Systems Modeling*, 17(4), 1139–1167. <https://doi.org/10.1007/s10270-016-0546-9>
- Agrawal, A., Karsai, G., Neema, S., Shi, F., & Vizhanyo, A. (2006). The design of a language for model transformations. *Software & Systems Modeling*, 5(3), 261–288. <https://doi.org/10.1007/s10270-006-0027-7>
- Allweyer, T. (2016). *BPMN 2.0: Introduction to the standard for business process modeling*. BoD—Books on Demand.
- Assouroko, I., & Denno, P. O. (2016). *A metamodel for optimization problems* (NIST IR 8096; p. NIST IR 8096). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.IR.8096>
- Benaben, F. (2012). *Conception de Système d'Information de Médiation pour la prise en charge de l'Interopérabilité dans les Collaborations d'Organisations* [Thesis, Institut National Polytechnique de Toulouse]. <https://hal-mines-albi.archives-ouvertes.fr/tel-01206234/document>
- Biard, T., Le Mauff, A., Bigand, M., & Bourey, J.-P. (2015). Separation of decision modeling from business process modeling using new “Decision Model and Notation” (DMN) for automating operational decision-making. *Working Conference on Virtual Enterprises*, 489–496.
- Burgueño, L., Cabot, J., & Gérard, S. (2019). The Future of Model Transformation Languages: An Open Community Discussion. *The Journal of Object Technology*, 18(3), 7:1. <https://doi.org/10.5381/jot.2019.18.3.a7>
- Chapron, J. (2006). *L'urbanisme organisationnel: Méthode et aides à la décision pour piloter l'évolution du système d'information de l'entreprise* [Phdthesis, Ecole Nationale Supérieure des Mines de Saint-Etienne]. <https://tel.archives-ouvertes.fr/tel-00796061/document>
- Chen, D., Vallespir, B., & Doumeingts, G. (1997). GRAI integrated methodology and its mapping onto generic enterprise reference architecture and methodology. *Computers in Industry*, 33(2–3), 387–394.
- Cuadrado, J. S., Guerra, E., & de Lara, J. (2014). A Component Model for Model Transformations. *IEEE Transactions on Software Engineering*, 40(11), 1042–1060. <https://doi.org/10.1109/TSE.2014.2339852>
- Czarnecki, K., & Helsen, S. (2003). Classification of model transformation approaches. *Proceedings of the 2nd OOPSLA Workshop on Generative Techniques in the Context of the Model Driven Architecture*, 45, 1–17.
- Fill, H.-G., & Karagiannis, D. (2013). On the Conceptualisation of Modelling Methods Using the ADOxx Meta Modelling Platform. *Enterprise Modelling and Information Systems Architectures (EMISAJ)*, 8(1), 4–25. <https://doi.org/10.18417/emisa.8.1.1>
- Gómez, F. J., Vanfretti, L., & Olsen, S. H. (2018). CIM-Compliant Power System Dynamic Model-to-Model Transformation and Modelica Simulation. *IEEE Transactions on Industrial Informatics*, 14(9), 3989–3996. <https://doi.org/10.1109/TII.2017.2785439>
- Gutin, G., & Punnen, A. P. (2006). *The traveling salesman problem and its variations* (Vol. 12). Springer Science & Business Media.
- James, W., Athansios, Z., Nicholas, M., Louis, R., Dimitios, K., Richard, P., & Fiona, P. (2013). What do metamodels really look like. *Frontiers of Computer Science*.
- Jouault, F., Allilaire, F., Bézivin, J., & Kurtev, I. (2008). ATL: A model transformation tool. *Science of Computer Programming*, 72(1–2), 31–39. <https://doi.org/10.1016/j.scico.2007.08.002>
- Karagiannis, D., Mayr, H. C., & Mylopoulos, J. (2016). *Domain-Specific Conceptual Modeling: Concepts, Methods and Tools*. Springer International Publishing. <http://public.eblib.com/choice/publicfullrecord.aspx?p=4587053>
- Mellor, S. J., Scott, K., Uhl, A., & Weise, D. (2004). *MDA distilled: Principles of model-driven architecture*. Addison-Wesley Professional.
- Mu, W., Benaben, F., Boissel-Dallier, N., & Pingaud, H. (2017). Collaborative Knowledge Framework for Mediation Information System Engineering. *Scientific Programming*, 2017, 1–18. <https://doi.org/10.1155/2017/9026387>
- Wang, T., Montarnal, A., Truptil, S., Benaben, F., Lauras, M., & Lamothe, J. (2018). A Semantic-checking based Model-driven Approach to Serve Multi-organization Collaboration. *Procedia Computer Science*, 126, 136–145. <https://doi.org/10.1016/j.procs.2018.07.217>
- Zhang, L., Fontanili, F., Lamine, E., Bortolaso, C., Derras, M., & Pingaud, H. (2020). A Systematic Model to Model Transformation for Knowledge-Based Planning Generation Problems. In H. Fujita, P. Fournier-Viger, M. Ali, & J. Sasaki (Eds.), *Trends in Artificial Intelligence Theory and Applications. Artificial Intelligence Practices* (pp. 140–152). Springer International Publishing. https://doi.org/10.1007/978-3-030-55789-8_13

Modélisation de CPPS : proposition d'un modèle générique

ADAMA ARAMA¹, LAURA LAGUNA SALVADO¹, ÉRIC VILLENEUVE¹, CHRISTOPHE MERLO²

¹ UNIV. BORDEAUX, ESTIA INSTITUTE OF TECHNOLOGY
F-64210 Bidart, France

a.arama@estia.fr, e.villeneuve@estia.fr, l.lagunasalvado@estia.fr

² UNIV. BORDEAUX, ESTIA INSTITUTE OF TECHNOLOGY, IMS, UMR 5218
F-64210 Bidart, France

c.merlo@estia.fr

Résumé – Depuis quelques années, l'introduction des nouvelles technologies, en particulier des systèmes cyber-physiques (CPS), dans la production industrielle est au cœur de nombreuses perspectives de recherche et de croissance économiques dans le monde. Ces évolutions sont souvent nommées Usines Intelligentes ou, Usines du Futur, et sont caractérisées entre autres, par l'interopérabilité entre ses sous-systèmes endogènes, l'aide dans la prise de décision, l'interaction entre l'humain et la machine, etc. Dans ce contexte de mutation industrielle, la modélisation paraît complexe et cruciale pour l'implémentation des CPS. Cet article, en s'appuyant sur des travaux de recherches antérieures menés dans le domaine de la modélisation des systèmes complexes, vise à contribuer en proposant un modèle générique des CPS. Ce modèle générique caractérise les principales composantes d'un CPS, et nous présentons un exemple d'instanciation de ce modèle dans le contexte de la production chimique. L'originalité réside dans l'intégration de l'humain comme élément au centre des préoccupations pour la modélisation de CPS, ce qui, à notre connaissance, a rarement été fait dans la littérature. Ainsi, ce modèle générique constitue un guide pour la modélisation de système de production à base de CPS avec notamment la prise en compte de l'humain dans la boucle et le pilotage du système par l'aide à la décision.

Abstract – In recent years, the introduction of new technologies, in particular cyber-physical systems (CPS), into industrial production has been at the heart of many prospects for economic research and growth around the world. These developments are often called Intelligent Factories or, Factories of the Future, and are characterized, among other things, by interoperability between its endogenous subsystems, assistance in decision making, human-machine interaction, etc. In this context of industrial change, modeling seems complex and crucial for the implementation of CPS. This article, building on previous research work in the field of complex systems modeling, aims to contribute by proposing a generic model of CPS. This generic model characterizes the main components of a CPS, and we present an example of instantiation of this model in the context of chemical production. The originality lies in the integration of the human as an element at the center of concerns for the modeling of CPS, which, to our knowledge, has rarely been done in the literature. Thus, this generic model constitutes a guide for modeling a CPS-based production system, in particular considering humans in the loop and controlling the system by decision support.

Mots clés – Aide à la décision, Système cyber-physique de production, Intégration humain – système, Gestion des connaissances.

Keywords – Decision support system, Cyber-physical production system, Human system integration, Knowledge management.

1 INTRODUCTION

À l'ère de la révolution industrielle actuelle appelée « Industrie 4.0 », l'accès aux données est favorisé par la digitalisation des processus industriels. Les systèmes cyber-physiques (ou Cyber-Physical Systems – CPS) représentent la technologie principale de cette digitalisation. Ils ont la capacité de fournir aux industriels de nouvelles connaissances sur leurs procédés de production mais aussi de les aider dans leurs prises de décisions quant à l'évolution et l'optimisation de ceux-ci. De ce fait, les CPS ont fait l'objet de nombreux travaux de recherches depuis quelques années et il existe à cet effet plusieurs définitions de CPS dans la littérature. Cependant, nous avons retenu celle de Lee [Lee, 2006] qui nous semble

plus en adéquation avec nos objectifs de modélisation de CPS. Il définit les CPS comme étant des « systèmes qui résultent de l'intégration de composants informatiques à des systèmes physiques. Ces composants informatiques, disposant de capacités de communication et de calcul, collaborent pour surveiller et contrôler les processus physiques via des boucles de rétroactions ».

De nos jours, les CPS sont au cœur des enjeux de domaines très variés notamment l'industrie, l'énergie, le transport, le militaire, etc., [Atat et al., 2018]. De plus, leurs applications ne cessent de croître car leur potentiel semble très important et ouvre de nouvelles perspectives sur les usages et les capacités en exploitation. C'est ainsi que dans le cadre de cette étude, nous nous intéresserons à l'utilisation des CPS dans le

domaine de la production industrielle donc aux CPPS (Cyber-Physical Production System). D'une part, les enjeux économiques liés à l'utilisation des CPS dans le secteur industriel sont considérables (création de croissance économique et d'emplois, etc.). D'autre part, leur utilisation permettrait une modernisation du secteur et un maintien de la compétitivité des entreprises [Zaratiegui, 2019]. Pour ces raisons, le projet HyperCOG, financé par le programme Horizon 2020 (H2020) de l'Union Européenne, est né. Nos travaux de recherche s'inscrivent dans le cadre de ce projet. Étant donné son contexte, l'objectif principal du projet est de démontrer le potentiel d'innovation des CPS et de l'analyse des données pour transformer l'industrie de production classique vers une industrie de production beaucoup plus autonome et intelligente. Cette évolution doit permettre d'améliorer la performance de la production par une assistance au pilotage du système de production grâce notamment à des mécanismes d'aide à la décision.

C'est dans cette perspective et dans une démarche globale visant à implémenter le système de production à base de CPS que nous proposons un modèle générique de CPPS.

En effet, la modélisation constitue une étape importante voire même indispensable pour une bonne implémentation des CPPS. Elle permet non seulement de mettre au clair le fonctionnement global du système de production mais aussi de faciliter son implémentation, composante par composante.

La suite de cet article s'articule donc en cinq parties. La section suivante est consacrée à l'étude de la littérature scientifique autour des modélisations des CPS. Ensuite la section 3 est dédiée à la proposition de notre modèle générique de CPS. La section 4 est quant à elle basée sur l'instanciation de notre modèle générique à un cas d'étude issue d'une industrie française du secteur chimique. La section 5 discute les résultats obtenus avant une conclusion générale.

2 ETAT DE L'ART

2.1 Définitions

L'évolution récente de l'industrie 4.0 a été un facteur clé au développement des CPS et à l'accroissement du nombre des travaux de recherche associés. Pour autant, ces travaux de recherche liés aux CPS sont complexes à cause de la diversité de leurs domaines d'applications et des technologies endogènes.

Dans cet article, nous nous intéresserons donc aux CPS dans le domaine de production. Les Systèmes de Production Cyber-Physiques (ou Cyber-Physical Production Systems - CPPS) sont des sous-concepts du concept CPS qui s'occupe du volet de production (ou Manufacturing) des CPS.

A partir de la définition de CPS, Monostori [Monostori, 2014] spécifie par rapport au contexte de la production qu'un « CPPS consiste en entités et sous-systèmes autonomes et coopérants, connectés au travers d'une relation contextualisée, au sein et au travers de tous les niveaux de la production, du process aux réseaux logistiques ».

La mutation croissante des industries classiques vers l'industrie 4.0, donc vers l'adoption des technologies CPS, est aussi surtout due aux avantages liés à l'utilisation de celles-ci, mais aussi à ces caractéristiques particulières [Al-Mhiquani et al., 2017]. Dans [Rudtsch et al., 2014], les auteurs ont recensé quelques bénéfices liés à l'utilisation des CPPS tels que : l'optimisation des processus de production, l'optimisation de la personnalisation des produits, production efficiente, etc. Partant des travaux réalisés dans [Nunes et al., 2015], nous pouvons affirmer que la prise en compte de l'humain dans la

boucle de production industrielle demeure une nécessité dans les systèmes de production (Figure 1).

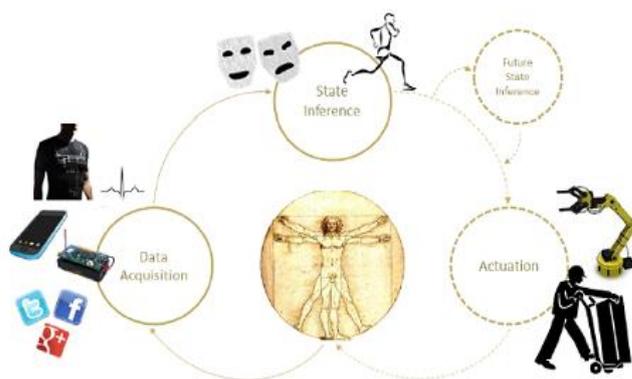


Figure 1. Intégration de l'humain dans le système de production [Nunes et al., 2015].

Ainsi, l'évolution des travaux de recherche tend désormais vers l'interaction entre l'Humain et les CPPS, [Cimini et al., 2020].

Cette tendance a conduit à la définition de la notion de systèmes cyber-physiques et humains (Cyber-Physical and Human Systems – CPHS) où l'Humain est considéré comme faisant partie intégrante du système à la fois comme utilisateur mais aussi comme source de données et/ou comme actionneur, [Villeneuve et al., 2019].

De tout ce qui précède, nous pouvons affirmer que le CPS peut avoir différentes définitions selon le fonctionnement attendu du système, et donner naissance à des sous-concepts tels que : CPHS, CPPS, etc. Cependant, en nous appuyant sur l'ensemble de ces travaux, nous définissons donc un CPPS comme « un système composé de sous-systèmes socio-techniques en interaction avec son environnement, et qui, par l'échange et le traitement de données et/ou d'informations, assure comme mission principale la réalisation d'un processus de production et son pilotage au travers d'une boucle « outillée » de rétroaction portant sur ce processus de production ».

L'interaction entre le CPPS et l'humain peut se faire à plusieurs niveaux du système (Figure 1) cependant, dans cette étude, nous nous focaliserons, sur la décision humaine en association avec le CPS. Dans [Dworschak et Zaiser, 2014] les auteurs ont défini quatre niveaux dans l'automatisation de la prise de décision entre l'humain et le CPPS qui sont :

- Full (Tout) : l'humain n'a qu'un rôle de supervision du CPS, qui prend la totalité de ses décisions sans l'intervention de l'humain ;
- Automation (Automatique) : le CPS guide l'humain dans sa tâche en prenant la majeure partie des décisions et en laissant les fonctions d'adaptation à l'humain ;
- Tool (Outil) : l'humain guide le CPS en prenant la majeure partie des décisions ;
- Manual (Manuel) : le CPS n'a qu'un rôle d'informateur de l'humain qui va prendre la totalité des décisions.

Nous retenons de cette partie que le choix du niveau d'automatisation des CPS, et par extension des CPPS est non seulement fonction des objectifs de l'entreprise mais aussi des moyens dont dispose celle-ci. Autrement dit, plus une entreprise voudra rendre autonome ses installations, plus elle

devra déployer de systèmes de mesure, de supervision, de pilotage, etc.

Cependant, dans le cadre des travaux présentés dans cet article, nous nous intéresserons principalement au niveau Manual, Tool et Automation dans l'aide à la prise de décision de l'humain dans le CPPS. Précisément, le modèle générique de CPPS qui sera proposé dans cette étude comportera un système d'aide à la décision qui évoluera en fonction du type de décision prise par l'Humain et donc du degré d'automatisation du CPS.

L'outil d'aide à la décision qui sera proposé aidera non seulement l'humain dans sa prise de décision, mais aura aussi une certaine autonomie pour prendre certaines décisions. Ainsi, pour un CPPS de niveau Manual, le décideur fera une analyse des différentes propositions du système d'aide à la décision (Decision Support System – DSS) et fera un choix parmi ces propositions. La notion d'interactivité entre le DSS et l'Humain est également considérée pour permettre à l'Humain de donner de l'information au DSS.

2.2 Modélisation de CPS et/ou de CPPS

Dans cette partie, nous proposons quelques approches pour modéliser le CPS et/ou CPPS extraits de la littérature.

2.2.1 Modèle conceptuel de CPS

Ce modèle (Figure 2) a été proposé dans [Griffor et al., 2017], il représente un cadre général ou une vue simplifiée d'un modèle CPS.

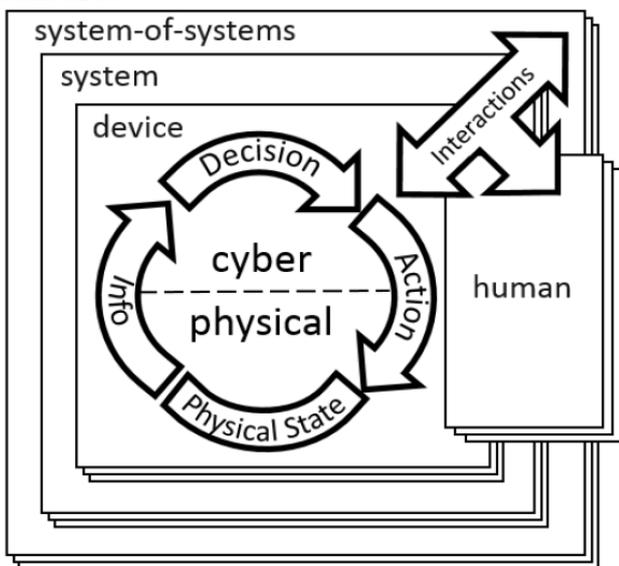


Figure 2. Modèle conceptuel de CPS [Griffor et al., 2017]

Nous retenons de ce modèle qu'un CPS est un système des systèmes (ou System-of-systems) c'est-à-dire un système composé d'autres sous-systèmes. D'autre part, l'humain peut interagir avec le CPS à tous les niveaux (Physique, Information ou données, Décision et Action). Il manque toutefois la précision du niveau d'interaction entre l'humain et le CPS et, le niveau d'aide à la décision.

2.2.2 Modèle conceptuel de CPPS

La Figure 3 représente un modèle de CPPS proposé dans [Cimini et al., 2020], conçu dans un contexte de production (Manufacturing) et de système multi-agents.

Via ce modèle, nous retenons que le CPPS peut être fait à base d'agents qui peuvent représenter toute la ligne de production. C'est un modèle qui permet de faire la simulation du système de production. De plus, la particularité de ce modèle est le fait de considérer l'humain comme un élément physique du CPPS,

ce qui est rarement fait dans la littérature. Un jumeau numérique [Tao et al., 2019] est associé au CPPS afin de faire la représentation de tous les éléments physiques qui entre dans la ligne de production par des agents.

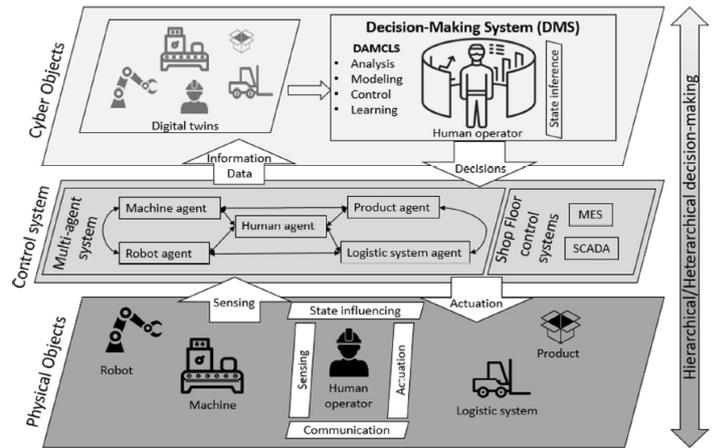


Figure 3. Modèle conceptuel de CPPS selon [Cimini et al., 2020]

2.2.3 Architecture de CPPS avec l'humain dans la boucle

Ce modèle (Figure 4) de CPPS ci-dessous a été proposé dans [Gnidenko et al., 2019]. A la différence du modèle précédent, l'humain n'est pas partie intégrante du CPPS (mais est modélisé comme un élément extérieur). En plus, le centre de décision n'est pas lisible sur ce modèle et cela nous permet d'affirmer que ce CPS est plus basé sur un mode de pilotage manuel. Cependant, le jumeau numérique est aussi associé à ce modèle.

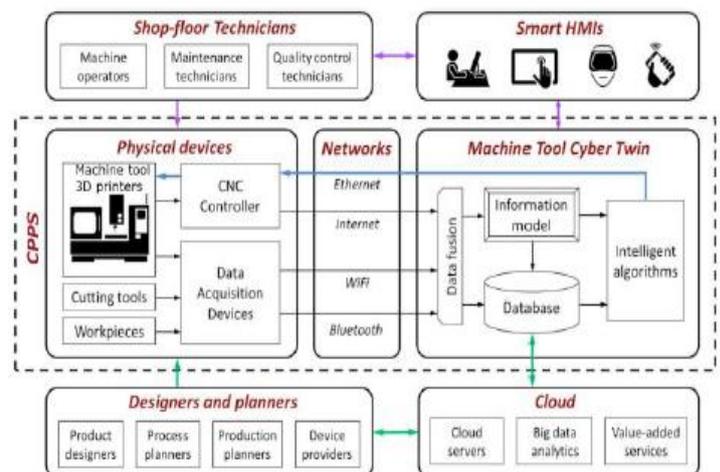


Figure 4. Modèle conceptuel de CPPS selon [Gnidenko et al., 2019]

2.3 CPPS et l'aide à la décision

L'aide à la décision constitue l'un des plus importants volets des CPPS, car elle permet d'assister le pilotage de CPS [Dworschak et Zaiser, 2014]. D'une façon générale, la démarche de l'aide à la décision consiste à choisir une solution en comparant plusieurs possibilités. Kadri [Kadri, 2014] affirme que : « la décision, dans un contexte donné, constitue la réponse d'un décideur qui effectue un choix entre plusieurs solutions susceptibles de résoudre le problème et la situation auxquels il est confronté ». Ainsi, selon Simon [Simon, 1960], une décision suit un processus dont les étapes sont :

- La perception de la nécessité d'une décision (renseignement ou information) ;
- L'inventaire et l'analyse des choix possibles (conception ou modélisation) ;
- La sélection (choix) ;
- La mise en œuvre (action) ;
- L'évaluation d'une des possibilités sélectionnées (évaluation).

L'étape de l'évaluation par le retour d'expérience, et est souvent négligée dans les modèles de CPPS. Cela peut s'expliquer par la faible prise en compte de l'humain le système physique de CPPS et donc l'absence de retour d'expérience pour l'évaluation humaine des actions entreprises. Ce processus décisionnel nécessite la mise à disposition un ensemble de solutions (ou alternatives) pour la résolution d'un problème donné en tenant compte de contraintes diverses (ou critères). La qualité de ce processus réside dans la capacité du décideur à appréhender différents paramètres : la nécessité d'une décision, le besoin de comprendre le problème qui se pose, et la maîtrise des critères et des leviers d'action [Merlo & Girard, 2004] pour proposer des solutions adaptées. Pour cela, nous nous appuyons sur le mécanisme de décision proposé dans [Endsley & Jones, 2001] illustré dans la **Figure 5**. Selon les auteurs, la conscience de la situation se divise en trois niveaux qui permettent :

- Perception : par le décideur des informations sur le système et sur son environnement ;
- Compréhension : la signification de ces informations ;
- Projection (ou anticipation) : de leur évolution dans le temps.

En plus de ce mécanisme, l'ajout du retour d'expérience dans la chaîne de décision avec notamment l'intégration de l'humain dans le système physique semble pertinent. Dans ce contexte, l'objectif d'un DSS est de réduire l'écart entre la conscience de la situation du décideur et la réalité de l'état du système pour permettre une prise de décision performante.

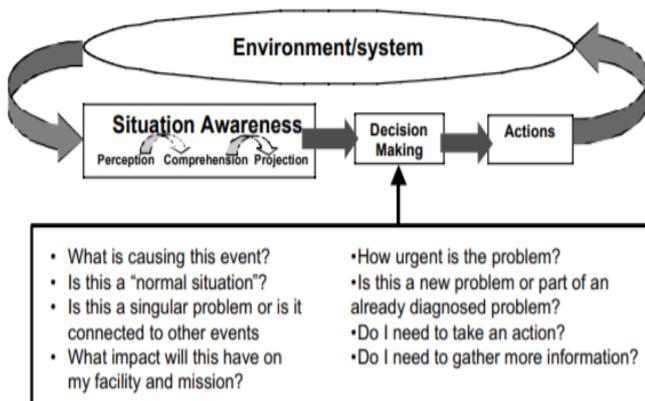


Figure 5. Modèle de DSS [Endsley & Jones, 2001]

En synthèse, il existe de nombreux travaux combinant la modélisation de CPS (ou CPPS) et le processus d'aide à la décision. Pour autant, les modèles existants dans la littérature sont généralement basés sur des domaines spécifiques et donc très spécialisés. De surcroît, l'humain n'est souvent pas considéré comme partie intégrante du système physique du CPPS. De plus, le pilotage par l'aide à la décision est rarement mis en valeur.

Ainsi, le modèle générique proposé dans la section suivante

vise à intégrer l'humain à la modélisation en tant que « composant » du système physique avec des rôles définis, et le pilotage du CPPS par l'aide à la décision. L'humain est donc également considéré dans le système décisionnel.

3 PROPOSITION D'UN MODELE GENERIQUE DE CPPS

Cette section a pour objectif de proposer un modèle générique de CPPS adapté à la conduite d'un système de production. Ce modèle générique (Figure 6) est basé sur l'approche GRAI [Doumeingts, 1984]. Cette approche propose de modéliser un système de production en trois sous-systèmes en vue d'en assurer la conduite. Cette approche correspond à notre objectif de conduite d'un CPPS au sein duquel nous étudions un système d'aide à la décision.

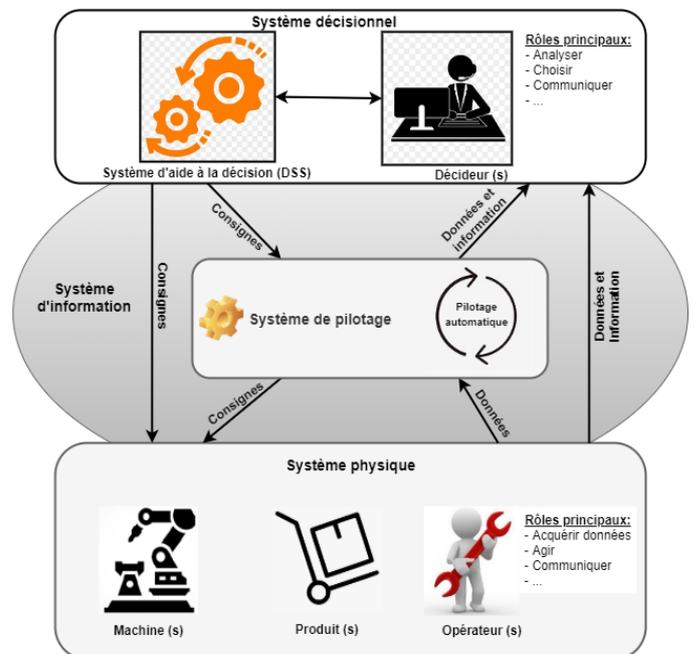


Figure 6. Modèle générique de CPPS proposé.

Comme illustré dans la Figure 6, notre modèle générique de CPPS (le système de production) est donc caractérisé par trois sous-systèmes :

- **Le système décisionnel (SD) :** il est composé de deux grandes parties qui sont : le (i) système d'aide à la décision et le (ii) le décideur (l'humain). Il assure le pilotage du système de production (SP). Il est basé sur les données et/ou informations collectées depuis l'environnement (technique et social) de production via les capteurs ou autres moyens d'acquisition de données et/ou informations.

Ces données sont basiquement de deux types : les (i) données dites technologiques c'est-à-dire les données qui proviennent des caméras, des capteurs, etc. et les (ii) données expertes en provenance des experts humains (opérateurs / techniciens / décideurs). Après la fusion de ces deux sources de données, des techniques de visualisations ou des algorithmes de machine learning, permettent d'aider l'humain (le décideur) dans sa prise de décision. Son rôle est donc d'analyser les informations à sa disposition, de choisir parmi les alternatives possibles et de donner des consignes.

- **Le système d'information** : joue le rôle de passerelle entre le système de décision et le système physique. C'est à travers lui que le système de production est piloté et contrôlé. Il reçoit les données via les capteurs et les caméras installés sur les lignes de production et transmet les données et informations pertinentes au système de décision. D'autre part, il reçoit des consignes depuis le système de décision qu'il va dispatcher vers le système physique (actionneurs, automates, etc.).
- **Le système physique** : représente l'ensemble des éléments opérationnels physiques du système de production (machines, produits, ...) y compris l'humain (l'opérateur). Celui-ci joue un rôle primordial dans le fonctionnement du système : il agit, il capte ou recueille des données et informations, et il en communique certaines pour le décideur.

Depuis l'environnement physique de production, l'opérateur peut acquérir des données et/ou informations et éventuellement les communiquer aux autres systèmes (d'information et de prise de décision). Il peut aussi agir selon les décisions émanant du SD

En synthèse, les principaux apports de ce modèle générique sont :

- L'intégration de l'humain dans le système physique et la définition de son rôle ;
- Le pilotage par l'humain assisté par le système d'aide à la décision et donc la considération de l'humain comme partie intégrante du SD ainsi que la définition de son rôle.

La section suivante vise à montrer que ce modèle est instanciable sur un cas concret issu de l'industrie de procédé chimique.

4 INSTANCIATION

Cette section 4 vise à démontrer l'applicabilité du modèle générique proposé sur un exemple concret de système de production. Pour se faire, le processus d'une des entreprises partenaires du projet HyperCOG (que nous nommerons : « l'entreprise » pour des raisons de confidentialité) a été simplifié et anonymisé afin de pouvoir utiliser les informations issues de l'analyse de ses processus de production tout en préservant la confidentialité de son activité.

Quatre visites sur le terrain ont été effectuées pour mener des entretiens et des observations afin de comprendre le fonctionnement de l'usine et de récolter des connaissances expertes. Cette méthodologie permet d'avoir une image du fonctionnement de l'entreprise et de pouvoir modéliser le CPPS associé. De plus, elle nous permet de recueillir des connaissances expertes qui seront exploitées dans les prochaines étapes notamment dans la mise en place du DSS.

En application du modèle générique, et de ses principales composantes, nous avons formalisé deux modèles (Figure 7 et Figure 8) représentant le fonctionnement de l'entreprise étudiée. L'instanciation porte sur trois grandes parties du modèle générique qui sont : le système décisionnel, le système d'information et le système physique.

D'une part, nous avons utilisé le formalisme SADT (en anglais *Structured Analysis and Design Technique*) qui permet de décrire les processus métiers pour modéliser le système physique [Ahmed et al., 2014]. D'autre part, nous avons utilisé le formalisme BPMN (*Business Process Model and Notation*) pour modéliser le processus global de prise de décision du système décisionnel [White, 2004]. Quant au système

d'information, nous avons pour l'instant identifié les principales informations qu'il véhicule car notre travail sur le terrain n'est pas encore terminé. Le choix porté sur les formalismes de modélisation SADT et BPMN sont ici justifiés par leur simplicité à représenter respectivement le processus des systèmes complexes et le processus du système décisionnel.

a. Système physique

La Figure 7 représente la modélisation du système physique et notamment de son processus de production.

A l'image du modèle générique, cette partie est constituée de tous les éléments physiques du système de production y compris l'opérateur (l'humain). Cependant, nous avons deux éléments physiques dans ce cas, à savoir les batteries et les opérateurs.

Le système de production est basé sur deux principales fonctions. La première est de « Transformer matière première (MP) en liquide ». Pour se faire les MP sont introduites dans les machines. Cela étant, les opérateurs agissent sur les machines afin de parvenir à la transformation de la matière première en liquide. La deuxième fonction est de « Séparer les matériaux » issus de la première fonction.

Pour se faire, ces matériaux liquides sont introduits dans les batteries. Pour parvenir maintenant à leur séparation, les opérateurs agissent sur ces batteries.

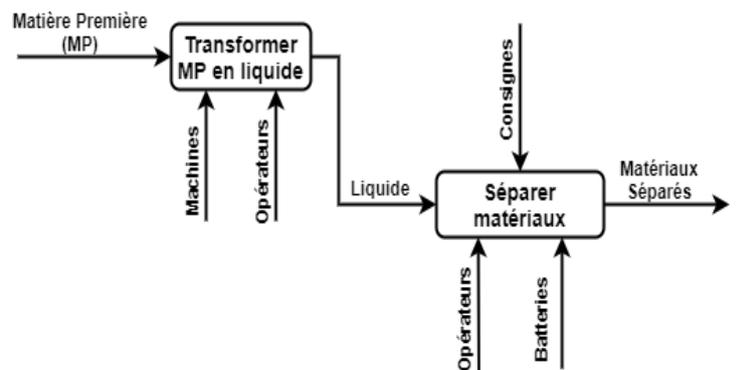


Figure 7. Modèle du système physique

Alors la production dans ce contexte consiste à séparer des matériaux qui sont mélangés sous forme de solide initialement. Ces matériaux arrivent à l'entreprise à l'état. Après une première transformation dans un premier atelier (première fonction), les matériaux deviennent alors de liquides de différents types (ou natures) mélangés. Ensuite dans le second atelier (deuxième fonction), ces matériaux liquides issus du premier atelier font l'objet d'une séparation soit par groupe de matériaux (supérieur à deux types) ou soit par type de matériaux (un seul type).

b. Système d'information

Pour le système d'information, nous allons identifier une liste des informations pertinentes stockées et surtout transmises.

i. Système de pilotage

Il représente la plus importante des composantes du système d'information. Il est alimenté par des données qu'il reçoit des capteurs, des caméras, etc. donc des données technologiques. Ensuite, il transmet les données et/ou informations au DSS pour traitement. Après analyse de ces données et/ou informations dans le SD, il reçoit alors des « consignes » du SD qu'il exécute via les actionneurs sur le système physique. Ces consignes peuvent avoir été formulées par le décideur (l'humain) lors de sa prise de décision. Notre objectif est également d'exploiter données et informations au regard de

règles métiers par le système d'aide à la décision (DSS) pour formuler sans intervention humaine certaines décisions reposant sur des contextes ou problèmes bien identifiés. Un dispositif de « pilotage » automatique est en charge de mettre en œuvre ce « pilotage automatique » du système physique.

ii. Consignes

Composé du DSS et du décideur, le système décisionnel sert à piloter l'ensemble du système physique de production. Pour se faire, le décideur a besoin de prendre des décisions tout en s'appuyant sur le DSS. La décision à la sortie du DSS peut être simplement une consigne (prélever d'un échantillon, vérifier de séparation des matériaux, etc.) destinée à un opérateur ou un groupe d'opérateurs. Le décideur peut, par exemple demander aux opérateurs de regarder le niveau de la décantation dans la salle des batteries, etc.

iii. Ensemble d'actionneurs

La décision peut également être destinée directement au système physique via un ensemble d'actionneurs, sans passer par l'opérateur comme par exemple : baisser la température, augmenter le régime des batteries, etc.

iv. Ensemble de capteurs

Dans l'entreprise étudiée, quelques capteurs sont installés dans les lignes de production afin de récolter les données. A ce jour, ces capteurs servent principalement à recueillir les données des températures des batteries, du niveau de décantation des matériaux dans les batteries, etc.

v. Données et informations

Les opérateurs qui surveillent la production depuis la salle de batteries peuvent remonter des informations directement au décideur via des talkies-walkies. Cela signifie que toutes les informations de la production ne lui remontent pas via les capteurs, ou via un média numérique.

c. Système de décision

La Figure 8 représente la modélisation du système décisionnel de l'entreprise.

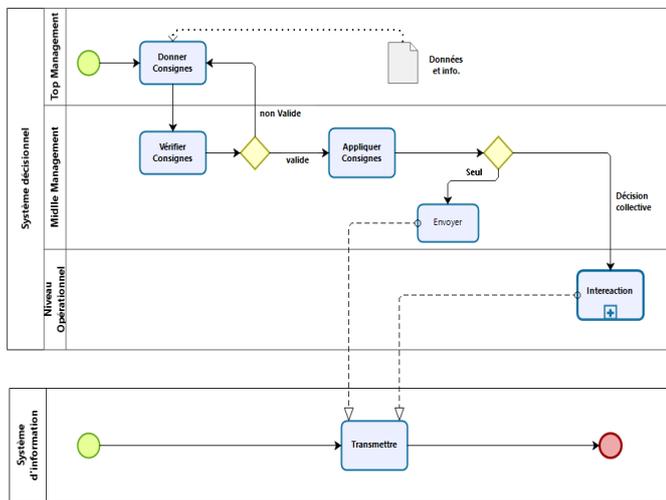


Figure 8. Modèle du système décisionnel

Partant du principe de la modélisation avec le formalisme BPMN, nous avons trois acteurs principaux dans le processus décisionnel de l'entreprise qui sont : le « Niveau opérationnel », le « Middle Management » et le « Top Management ». Dans le processus décisionnel commence par le « Top Management » qui, sur la base de certaines données et informations telles que le stock de matière première, ou la demande des clients, donne des consignes de production au « Middle Management ». Ce dernier à son tour, vérifie l'applicabilité des consignes reçues. Si les consignes sont

applicables, il a deux options. Soit il agit directement sur le système de production soit il passe par le niveau opérationnel qui agira alors sur le système de production. Sinon, il retourne les consignes au top « Management » pour être réévaluées.

5 DISCUSSION

L'instanciation du modèle générique sur le système de production de l'entreprise étudiée nous permet d'affirmer que le système de production de l'entreprise est un CPPS. Cette affirmation est basée sur deux aspects importants. Cette instanciation, nous permet, d'une part, de retrouver ce que dit la définition (proposée dans la section 2) concernant entre autres, les sous-systèmes (capteurs, actionneur, etc.), la rétroaction, et la présence de l'humain dans le système physique. D'autre part, nous pouvons affirmer que le système de production de l'entreprise est un CPPS sur des critères tirés dans [J. Lee et al., 2015]. Selon les auteurs, le CPPS est fonction de cinq grands niveaux (C1 à C5) de maturité ou de complexité :

- **C1** : Au niveau Connexion, le CPS opère sur un réseau Plug & Play et utilise des données envoyées par un réseau de capteurs ;
- **C2** : Au niveau Conversion, le CPS sait traiter les données et les retranscrire en informations de plus haut niveau ;
- **C3** : Au niveau Cyber, le CPS a une connaissance des autres CPS de l'environnement et peut interagir avec eux pour enrichir son propre traitement d'information ;
- **C4** : Au niveau Cognition, le CPS est capable d'établir un diagnostic basé sur des simulations de son propre comportement et une analyse différentielle des données de capteurs ;
- **C5** : Au niveau Configuration, le CPS peut s'adapter seul en cas de défaillance, se reconfigurer ou ajuster de manière autonome ses paramètres afin de retourner à un comportement nominal. Selon Cardin[Cardin, 2019] ce niveau est rarement atteint.

Ainsi, sur la base de ces critères cités ci-dessus, nous pouvons affirmer que le système de production actuel de l'entreprise est un CPPS et qu'il correspond aux deux niveaux C1 et C2. De là, nous constatons que le CPPS de l'entreprise a un niveau de maturité et de complexité très bas.

L'aboutissement du projet HyperCOG a pour but de faire évoluer l'entreprise. D'ailleurs, c'est pour ce cadre précis que l'entreprise s'est engagée dans le projet HyperCOG, afin de parvenir à une industrie 4.0 avec un niveau de complexité et de maturité plus élevé. Pour se faire, le projet contribuera notamment par l'installation des capteurs supplémentaires, des caméras sur les lignes de production, la mise en place d'un système d'aide à la décision, ou encore la formation du personnel dans l'utilisation des futurs outils donc à faire à évoluer aussi les compétences.

6 CONCLUSION ET PERSPECTIVES

Dans cet article, nous constatons que la réussite du développement des systèmes de production cyber-physiques (CPPS) passe d'abord par une modélisation réussie et complète. Cela revient à relever plusieurs défis de types et de nature différents qui sont identifiés précédemment tels que, entre autres :

- La caractérisation des rôles des humains dans le système physique de production ;

- L'étude des informations et des mécanismes de prise de décision afin de proposer un système d'aide à la décision pertinent et pro-actif.

Le modèle générique proposé dans la section 3 permet de répondre à ces défis et donc contribue à combler le manque que nous avons identifié dans la littérature. Comme nous l'avons évoqué dans la section 2, l'humain est rarement intégré dans la modélisation des CPPS.

Aussi, ce modèle pourrait être adapté à n'importe quel CPPS avec évidemment des niveaux de complexités et de maturités différents.

Pour rappel, l'objectif du projet est de démontrer le potentiel du CPPS par les innovations qu'il introduit mais aussi à travers l'analyse des données. Cette modélisation fait donc partie de ce long processus. Dans cette perspective, en plus de l'entreprise étudiée, ce modèle générique (Figure 6) sera instancié sur d'autres exemples concrets issus des autres entreprises partenaires du projet. Ensuite, toutes les instanciations feront l'objet d'une implémentation afin d'atteindre les objectifs du projet.

Par ailleurs, il est intéressant de noter que la conduite d'un système de production de type CPS via un système d'aide à la décision (DSS) implique la prise en compte des niveaux de compétence et/ou de qualification non seulement des décideurs, mais également des opérateurs, ainsi que des « experts métiers » que nous identifions avec un rôle fondamental pour garantir le bon niveau de fonctionnement du DSS. Au-delà de la modélisation du CPPS, et de la caractérisation de modèles de connaissance sur lesquels construire le futur DSS, nous envisageons de structurer une démarche complète de conception et de réalisation du DSS qui permette d'intégrer ces différents rôles, les compétences et qualifications associées, et allant jusqu'à définir les conditions de leur évolution dans une phase de déploiement.

7 REMERCIEMENTS

Les travaux de recherche présentés dans cet article ont été financés dans le cadre du projet HyperCOG par le programme d'innovation et de recherche de l'Union Européenne Horizon 2020 par la convention de subvention n°869886. Nous remercions l'ensemble des partenaires du projet HyperCOG pour leur aide et leur implication.

8 REFERENCES

- Ahmed, F., Robinson, S., & Tako, A. A. (2014). Using the structured analysis and design technique (SADT) in simulation conceptual modeling. *Proceedings of the Winter Simulation Conference 2014*, 1038-1049.
- Al-Mhiqani, M. N., Ahmad, R., Abdulkareem, K. H., & Ali, N. S. (2017). investigation study of cyber-physical systems: characteristics, application domains, and security challenges. *12(22)*, 12.
- Atat, R., Liu, L., Wu, J., Li, G., Ye, C., & Yi, Y. (2018). Big Data Meet Cyber-Physical Systems: A Panoramic Survey. *ArXiv:1810.12399 [Cs, Stat]*. <http://arxiv.org/abs/1810.12399>
- Cardin, O. (2019). Classification of cyber-physical production systems applications: Proposition of an analysis framework. *Computers in Industry*, 104, 11-21.
- Cimini, C., Pirola, F., Pinto, R., & Cavalieri, S. (2020). A human-in-the-loop manufacturing control architecture for the next generation of production systems. *Journal of Manufacturing Systems*, 54, 258-271.
- Doumeings, G. (1984). *Méthode GRAI: méthode de conception des systèmes en productique*. Bordeaux 1.
- Dworschak, B., & Zaiser, H. (2014). Competences for Cyber-physical Systems in Manufacturing – First Findings and Scenarios. *Procedia CIRP*, 25, 345-350.
- Endsley, M. R., & Jones, D. G. (2001). Disruptions, Interruptions and Information Attack: Impact on Situation Awareness and Decision Making. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 45(2), 63-67.
- Gnidenko, A., Sobolevsky, V., Potriasaev, S., & Sokolov, B. (2019). Methodology and integrated modeling technologies for synthesis of cyber-physical production systems modernization programs and plans. *IFAC-PapersOnLine*, 52(13), 642-647.
- Griffor, E. R., Greer, C., Wollman, D. A., & Burns, M. J. (2017). Framework for cyber-physical systems: Volume 1, overview (NIST SP 1500-201; p. NIST SP 1500-201). National Institute of Standards and Technology.
- Kadri, F. (2014). Contribution à la conception d'un système d'aide à la décision pour la gestion de situations de tension au sein des systèmes hospitaliers. Application à un service d'urgence.
- Lee, E. A. (2006). Cyber-Physical Systems—Are Computing Foundations Adequate? Position Paper for NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap, 2, 9.
- Lee, J., Bagheri, B., & Kao, H.-A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18-23.
- Merlo, C., & Girard, Ph. (2004). Information system modelling for engineering design co-ordination. *Computers in Industry*, 55(3), 317-334.
- Monostori, L. (2014). Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. *Procedia CIRP*, 17, 9-13.
- Nunes, D. S., Zhang, P., & Sa Silva, J. (2015). A Survey on Human-in-the-Loop Applications Towards an Internet of All. *IEEE Communications Surveys & Tutorials*, 17(2), 944-965.
- Rudtsch, V., Gausemeier, J., Gesing, J., Mittag, T., & Peter, S. (2014). Pattern-based Business Model Development for Cyber-Physical Production Systems. *Procedia CIRP*, 25, 313-319.
- Simon, H. A. (1960). The new science of management decision.
- Tao, F., Qi, Q., Wang, L., & Nee, A. Y. C. (2019). Digital Twins and Cyber-Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering*, 5(4), 653-661.
- Villeneuve, É., Akle, A. A., Merlo, C., Masson, D., Terrasson, G., & Llaría, A. (2019). Decision support in precision sheep farming. *IFAC-PapersOnLine*, 51(34), 236-241.
- White, S. A. (2004). Introduction to BPMN. IBM Corporation.
- Zaratigui, I. (2019). Hyperconnected Architecture for High Cognitive Production Plants. 70.

CIGI 2021

Exploring the role of Lean to cope with crisis: COVID-19

HAJAR SAIDI HASSANI ALAOU¹, EL HASSAN IRHIRANE¹, BADR DAKKAK¹

¹ Laboratoire Ingénierie des Systèmes et Applications «LISA»
Marrakech, Morocco

alaouisaidi.hajar@gmail.com, e.irhirane@uca.ac.ma, b.dakkak@uca.ac.ma

Résumé –

Les organisations sont confrontées à une alarmante crise sanitaire qui peut modifier leur situation, voire affecter leur survie. En période d'incertitude, les professionnels ont du mal à prendre de bonnes décisions pour faire face à la nouvelle normalité. La gestion du changement et la capacité d'adaptation rapide sont des facteurs clés pour transformer cette menace en une opportunité de croissance. Après Covid-19, une crise économique se profile, entraînant : chômage, récession, incertitude et risques accrus, problèmes de sécurité et de santé, etc.

Comment les organisations peuvent-elles surmonter cette crise et comment peuvent-elles résister à ses perturbations les plus imprévisibles ? Comment assurer la transformation post-COVID-19 et comment accompagner ce changement ?

Pour répondre à ces questions, des travaux de recherche sur le rôle du Lean ont été menés. Les environnements Lean peuvent-ils aider à lutter contre cette pandémie et fournir des pratiques permettant aux entreprises de se redresser ?

L'objectif de cet article est d'examiner le rôle du Lean ainsi que d'apporter des améliorations à ses pratiques pour résister à l'impact de la pandémie. Cette étude est l'une des premières déployées pour explorer les avantages du Lean en temps de crise ou d'incertitude, offrant de nouvelles réflexions et perspectives au domaine

Abstract –

Organizations face a serious health crisis which can alter their situation or even affect their survival. In times of uncertainties, professionals are struggling to make good decisions to cope with the new normal. Change management and rapid adaptability are key factors in order to transform this threat to an opportunity to grow. After Covid-19 crisis, an economic crisis is shaping up bringing about: partial unemployment, unemployment, recession, uncertainty and increased risks, safety and health problems, etc.

How can organizations overcome this crisis of a completely different nature and how can they resist to its unpredictable shocks? How to transform the company and accompany the change?

To answer these questions, research work on the role of Lean system has been carried out. Can Lean environments help fight this pandemic, and provide practices for business to recover?

The aim of this paper is to examine the role of Lean to resist to the impact of COVID-19 pandemic on businesses and to provide improvement to cope with the change. This study is one of the initial efforts to explore the benefit of lean in crisis or uncertain times, offering new reflexions and insights to the field.

Mots clés - Lean, pandémie, post-COVID-19, compétences de gestion de crise.

Keywords – Lean, pandemic, post-COVID-19, crisis management competencies.

1 INTRODUCTION

The Covid-19 health crisis has led to many changes in the way we live and work. Beyond the changes linked to social distancing, quarantines and lockdowns, it has brought about many economic consequences and disruptions within organizations: decreased or increased demands, shipping and logistic issues, delays and problems with raw materials acquisition, financial consequences, remote or smart working, employee well-being, workforce capacity, and their health and safety (Ivanov & Dolgui, 2020; Tortorella et al., 2020; Zanni Tim, 2020; Zhang et al., 2020).

In these circumstances, at a time where the unique constant is to change constantly, change management, previously applied to business transformation and the implementation of new

initiatives, is a key requirement for adapting our systems to the new normal that covid-19 has generated (Kuzmanova & Ivanov, 2019; Mishra et al., 2020). This challenging times, emphasized the importance of actions and decision-making at the right time based on changes of orders, climate, and governments decisions. It's a chance for organizations to learn and develop their capabilities of adaptability and proactivity at short and long terms in order to be able to survive in hard times and open up new opportunities to improve their processes to meet new challenges (Sinapin, 2020).

Organizations have taken various measures to preserve employees' health and security without compromising the business results. Most companies moved some of their employees to smart working, while others who couldn't work

remotely stayed onsite. As a result, contagion risk prevention policies in the work environment, redistribution of work teams, workplace redesign and restructuring processes were implemented to ensure staff protection and proper reactivity to the fallout from the covid-19 crisis (Ivanov & Das, 2020).

Many of these organizations in different sectors have adopted the Lean philosophy years ago. Lean is notorious for its simultaneous integration of management and operational activities and for its systematic reduction of wastes in order to achieve operational excellence (Kamble et al., 2020). The successful implementation of Lean Manufacturing practices has proven to provide competitive advantages in terms of improving product quality, increasing productivity, enhancing workforce health and safety, and improving customer satisfaction (Jasti & Kodali, 2014; Netland, 2016). It is widely recognized by researchers and practitioners in the field that lean manufacturing has a positive impact on the performance of organizations, enabling more flexible and efficient operations and providing an independent, competent workforce driven by customers' requirements (Taj & Morosan, 2011; Womack & Jones, 1996). With the crisis of Covid-19, however, came new consequences and challenges for all organizations. Yet the fallouts remain unknown, which raises risks and ambiguities about how lean cultures and environments will adjust to the upheavals of the pandemic. Thus, based on these observations, the next research question emerges: What is the role of lean environments in providing competitive advantages to organizations during and after the COVID-19 crisis?

The purpose of this study is to review the role of Lean, its practices, tools, and the skills and competences developed within Lean organizations to help them overcome this crisis and its potential implications while providing support for the required changes. This study is based on a literature review of Lean practices and crisis management as well as discussions with Lean professionals from different countries and companies of different sizes operating in different fields in order to provide revisited best practices, necessary skills and business requirements to resiliently cope with the covid-19 crisis and persist in the post-crisis period.

The rest of our article is organized as follows. Section 2. presents the methodology of this research. Section 3. discusses the impacts of COVID-19 on economy and manufacturing. Section 4. reviews Lean and change management, Section 5. covers Lean, crisis management and resilience, and examines the reaction of Lean environments to crisis and the improvements needed to successfully withstand the implications of current and future crises, Section 6. Discusses lessons learned from the crisis as well as future considerations and changes. Finally, section 7. closes the paper by presenting the implications of our findings, along with the perspectives that may motivate future studies.

2 METHODOLOGY

Based on the nature of this study and in order to meet its objectives, the researchers used a qualitative analysis based on literature and discussions with lean professionals.

In addition to literature, data collection were conducted through virtual interviews as well as webinars organized by Lean companies that provide training to professionals (LeanFrontiers for example) (*Home | Lean Frontiers*, n.d.). Each guest and interviewee discussed and explained the role of lean culture/environments and how they could respond to these difficult times. We decided to use interviews due to their ability to gather a large amount of information from a small number of lean practitioners, and for their effectiveness in increasing our

understanding and exploration of the impact of Lean and its practices during the COVID-19 crisis as expressed by Lean practitioners at real time in the different phases of this crisis (before lockdown during lockdown and after lockdown).

The questions in this interview were open-ended questions that allowed us to collect in-depth information. Questions were divided into two parts as follows :

- Part 1: it consists of the company and interviewee information: the company size, its sector, manufacturing/service, its maturity in terms of lean, the role of the interviewee in the company, years of experience, etc.
- Part 2: The fallouts from the crisis and the role of lean to cope with this change: increase/ decrease of the demand after the COVID-19 crisis, changes introduced by the crisis, the measures and decisions carried out to respond to the crisis, information on the decision-making teams, the role of the developed Lean competencies, required trainings for lean professionals to deal with such crisis, the role of each of the managerial and technical practices of lean to overcome this crisis, Lean practices that proved to be unsuitable for such a situation, and how can we modify certain practices to better respond to these crises.

The next section will present the COVID-19 impacts on manufacturing.

3 THE FALLOUT OF COVID-19 CRISIS ON MANUFACTURING

At the apogee of globalization, we have witnessed a transformation of the economy, which considers itself increasingly open. This has given rise to international trade which, according to David Ricardo's theory of comparative advantage, will be mutually beneficial provided that each country specializes in its comparative advantage (Pflüger & Tabuchi, 2019). This openness has led to the increased dependence of global industries on their suppliers (van Ee, 2021).

With the Covid-19 crisis, the negative effect of this dependence has become more severe, forcing industries to adapt to potential changes.

Multinationals have long relied on offshoring, but the current crisis has shown that this is not necessarily the best decision to make; trade deficits, rising unemployment and dependence on suppliers are the downsides of offshoring. During this crisis, countries have quickly run out of the basic needs to cope with a crisis that continues to devastate our economy. Hence the need to implement stimulus packages for national industries, yet in the current recession, public finance investments have been put on hold due to flat performance this year, hindering growth and change.

At the beginning of the epidemic, Renault suspended production in most of its factories because of components that were not delivered by Chinese suppliers, which ended up with the permanent closure of a large part of these factories. There was a colossal loss for all industries in all sectors which developed an increasingly intense dependence on Chinese, Indian and other suppliers (Hoek, 2020).

The Chinese government had to close major production sites to control the spread of the coronavirus, as China is an industrial hub, resulting in a global supply chain disruption that affected almost all sectors, from pharmaceuticals to automotive (Maffioli, 2020). China has taken an incontrovertible place in global supply chains. Nowadays, it is impossible to do without China for production. The need to rethink traditional industry

and management strategies is imminent, and supply chains and company logistics are undergoing metamorphosis (Chen et al., 2019). A French company that had been sourcing high-tech parts from a Chinese supplier for years finally opened a highly automated factory to shorten supply chains and limit its dependence on its former supplier. This has a higher production cost (between 10% and 15%), but the quality of its products has increased considerably, and it is now easier to manage its stock. The disruption of the industry in China has had an impact on other countries, where prices have increased due to the shortage of raw materials from China. As Chinese industries had suspended production, world prices for raw materials such as metals declined everywhere, causing a major slowdown in the various metal-producing economies (Maffioli, 2020).

After the oil shocks (1973-1979), industries decided to opt for other energy sources to replace oil and limit dependency, so the R&D budget increased, and this led to major advances in renewable energies (Poulsen & Lema, 2017). Today, it is great time to move massively towards renewable energies as well as developed work organizations and practices.

Covid-19 has caused a major shock to manufacturing and services industries and also to global supply chains. Some sectors, such as food and consumer goods, or even the pharmaceutical sector experienced a recovery and even an increase in demand during the crisis, while others, including the aerospace and automotive industries and their affiliated suppliers, suffered a dramatic decline in demand (Fernandes, 2020).

All over the world, companies experienced a health impact that caused a drop in local and external demand, a breakdown in the production and supply chains (Hoek, 2020; Mena et al., 2020; S. Singh et al., 2020), serious financing problems, as well as administrative restrictions. Companies that rely on exports are the most affected during the crisis as well as small businesses who have not been able to cope with the crisis.

At a time when most businesses were completely closed due to government restrictions, the food sector is one of the sectors that must continue to operate throughout the supply chain in order to secure food. In these challenging times, it is essential to maintain the health and safety of employees while ensuring a high level of food safety and consumer confidence (Nakat & Bou-Mitri, 2021).

At the beginning, companies were challenged to adjust very quickly to deal with the emergency. The interconnected nature of the crises - health, economic, social and environmental - has changed the situation, and all stakeholders must consider the lasting impact of this new and intense experience. The lasting repercussions of the crisis on management have to be identified and analyzed.

Job security, financial consequences, remote work, worker well-being and career attitudes are essential shifts caused by the covid-19 crisis that organizations need to understand and react to in order to cope with the new normal in short and long-term (Howe et al., 2020).

To close this section, the crisis has enhanced the role of proximity and committed organizations to a higher humanization of management, with new forms of work organization and work life. Finally, Covid-19 enabled managers to review their global management of the company, to be more proactive than before, to understand the post-crisis period and to anticipate any other eventuality.

The next section will review lean, change and crisis management.

General "attitudes towards change usually refer to a person's cognitions towards change, emotional reactions to change, and a behavioural tendency to change" (Vakola & Nikolaou, 2005). Employee reactions to change can vary from positive attitudes: accepting and supporting change, to negative attitudes: resisting and opposing (van der Smissen et al., 2013). Through these positive and negative interactions, it is possible to create engagement, resistance and support for change (Dasborough et al., 2015). Support for change is not the same for everyone because individuals differ in their ability and willingness to adapt to change (Bovey & Hede, 2001).

Resistance to change is one of the challenges of change, but without change, an organization cannot move forward. Furthermore, Lean is a system based on continuous improvement, so it requires continuous change and subsequently a high level of employee commitment, willingness and readiness to change and to adapt systems and people to each change along with a change management to overcome resistance in order to ensure successful implementation of new actions (Asnan et al., 2015).

Change management refers to organizational change characterized as the process of continually renewing an organization's focus, structure and capabilities to meet the evolving requirements of external and internal customers (By, 2005).

Change management focuses on managing people to successfully cope with the shift through organizational tools including change model, training, communication and coaching (Asnan et al., 2015).

It was similarly believed and recommended to use communication, education, training, motivation, assurance, rewards and compensation to positively manage change (K. Singh et al., 2012). Training and communication are therefore key factors in change management as training is an activity that leads to the acquisition of information and skills and brings about a change in attitude or behaviour (Oláh et al., 2017).

In mature lean organizations, continuous change is normal, people are used to change and can adapt very fast due to their effective change management and respect for people (efficient leadership, adequate coaching, and adapted training programs). Yet, it is argued (Rosenbaum et al., 2018) that change management is simply a three steps process of what Lewin defines as unfreezing the present (unfreeze), moving at the new level (change), and freezing at the new level (refreeze) (Allaoui & Benmoussa, 2020; Cummings et al., 2016; Lewin, 1947).

The unfreeze phase has taken place under unpredicted circumstances originating from COVID-19, which has resulted in a shift to online training, communication, meetings, as well as new work organization because of the anticipated uncertainties in the continuation of the previous mode. Currently, it is quite impossible to conduct training or even work in regular mode in the context of the COVID-19 pandemic, where maintaining social distance is essential; therefore, the transition to online mode has definitely become a must that has brought the organization and the individual into an unfreeze phase. It was a time of motivation and preparation for the system and stakeholders. Step two includes changing the process to introduce in practice a new online/hybrid mode that suggests a new organization of work. Seeking a better applicable model that provides a balanced way of gaining the benefits of both normal and remote work remains the goal. Importantly, here change is not an isolated event rather it is a proactive and dynamic process that requires a perspective adjusted to time and a new mindset for the new connected mode at the individual and

organizational level in order to complete the transition phase in lean environment. Resilience needs to be built into Lean organizations: increasing innovation and reducing digital distance.

A paradigm shift occurred after months of experience of smart working and many other changes in supply chain management. Remote working has proved its benefits, leading companies to think about finding a balanced equation to include it even after the COVID-19 pandemic which will lead to refreezing. Refreezing can't be avoided to integrate technology into the work, learning and training processes that allows organizations to provide solutions and methods in which employees would not only feel comfortable and safe but can also meet the demands of technology/industry in this century.

In short, Lean organizations have experienced the three phases of change of Lewin during this pandemic: unfreeze, change and refreeze. Change and adaptation, which are normal concepts for Lean and its practitioners, have become a necessity today and could also affect lean practices in order to cope with such crises. Lean employees and leaders are still the most effective at dealing with change and instability, though.

The next section will discuss crisis management as a type of change management under different circumstances and conditions.

5 LEAN, CRISIS MANAGEMENT AND RESILIENCE

Crisis management stands for optimal planning and crisis control. It includes four key elements which are damage reduction, preparedness, emergency response and reconstruction. A well-developed crisis management system assesses potential risks and available resources, considers a balance between available resources and risks, and enables the crisis to be controlled using available resources (Fooladvand et al., 2013).

Crisis management is related to the response to, management and recovering from unpredictable situation. Crisis strategic management is a type of management. Principles of the effective crisis management are: the early crisis diagnosis in the company's financial activity, the rapidity of reaction to crisis phenomena, the relevance of the reactions to the actual threats to the financial resources and the implementation of the full potential to overcome the crisis. To recover from the impacts of the crisis, post-crisis assessment, determination of controlled and uncontrolled crisis elements, new strategies development, and strategy selection and implementation are all required (Groh, 2014). It is important to learn quickly from crisis events that actually occur. Organizational crisis management is a series of procedures used to sustain normal business operations, minimize stakeholder loss, and use learning to improve crisis management processes (Pearson & Clair, 1998).

In circumstances such as the current crisis, where there is an outbreak of massive change with high levels of uncertainty, complexity and disruptions. Organizations need to adopt quick reaction to customer needs as well as to the disruptive changes related to the environment.

Resilience is the ability to prepare, react and recover from unexpected disruption through maintaining business continuity at the desired level. According to Rapaccini et al, in such conditions, the following, are elements to guarantee resilience (Rapaccini et al., 2020):

- Agility reflects the ability of the organization to adapt rapidly to a changing environment. It is vital for business survival in such circumstances where one can't predict the future of offers and demands along with customers' future needs.

- Preparedness represents the capacity of decision makers based on previous experiences and lesson learned to rebuild their business through seeking new opportunities and implementing innovative ideas to cope with the disruptions (Muñoz et al., 2019; Portuguese Castro & Gómez Zermelo, 2020). Risk management which is concerned with the identification and evaluation of potential threats and the acceptability of the risk together with a contingency plan are essential to prepare for turbulent and volatile conditions.
- Elasticity which brought about the increasing share ability and exchangeability of interactions between people and things within a dynamic ecosystem.
- Redundancy which refers to resources (production facilities, material stocks, etc.) that can be quickly activated to rebuild the value stream (Linnenluecke, 2017).

In the same context, Noeline Sinapin stated that efficiency and resilience are key factors to agility in order to overcome the crisis and their unknown potential perturbations (Sinapin, 2020).

One of the benefits of Lean is its capacity to assist companies to cope with unforeseen and rapid changes while avoiding excessive and unnecessary pressure on their workforce. The Lean management system has the potential to adapt quickly to change through the robust adjustment mechanism that the Lean philosophy puts in place to address disturbances.

However, based on the elements of the crisis resilience, in order to overcome the repercussions of the crisis and to persist and grow in the post-crisis period, Lean environments must provide an agile approach that has proven its ability to adapt to change yet remains compatible with Lean as they have many similarities and complementarities.

5.1 Leadership competencies for crisis management compared to lean leadership competencies

5.1.1 Leadership competencies for crisis management

There is a huge relationship between crisis management and leadership. Effective communication, building trust, increasing productivity and protecting brand value are roles played by leaders throughout the time of crisis. Leadership competencies are determinants for the success of crisis management (Bhaduri, 2019; Lockwood, 2005). For instance, emotional intelligence are important leadership competencies to overcome crisis.

Effective leadership in crisis situations is essential to control the damage of the crisis (Sommer et al., 2016). Leadership plays a crucial role in the different phases of crises: pre-crisis, during the crisis, and post-crisis (Jin et al., 2019). The leadership vision can avoid or mitigate crises with proper planning. Resilient leadership in crisis circumstances can be effective in protecting the company's material and human resources. In a crisis, the resilience of employees can be strongly influenced by the behaviour of the organization's leaders. As a result, crisis Leaders need to have higher levels of leadership skills (Wisittigars & Siengthai, 2019). Crisis management and leadership are interconnected organizational competencies (Boin & 'T Hart, 2003). Crisis situations are known by high levels of uncertainty and pressure in decision-making (Cook et al., 2015). Leadership competencies include personality traits, skills, values, knowledge, abilities, and aptitudes to best perform their tasks (Boyatzis, 2009). Crisis leadership competencies ensure quick adaptability to changing circumstances based on effective communication and people management. They can be classified in three categories: Self-management competencies, business management competencies, and people management competencies.

Researchers have identified four critical leadership competencies related to crisis leadership: a sense of urgency; strong emotional intelligence; problem-solving skills; and sensitivity of communication (Betancourt et al., 2017; Tubin, 2017). Recently, Wisittigars and Siengthai stated that having a management plan that deals with crisis prevention, crisis management and post-crisis recovery is critical to be able to minimize the impact of crisis in actual times along with developing employees' crisis-related competencies to be sure they can manage crisis situations ((Wisittigars & Siengthai, 2019). Table 1 represent the Leadership competencies required to resist and overcome a crisis (Table 1).

Table 1. Leadership competencies required to resist and overcome crisis

	Crisis Leadership Competencies	Authors
Emergency preparedness	Emergency planning and business continuity, Law and regulatory issues, Risk management, Emergency preparedness, sense making, perspective taking, Critical situation handling skills, Knowledge of safety and security, Safety awareness.	(Stankeviciute & Savaneviciene, 2014; Wisittigars & Siengthai, 2019)
Crisis communication	Public speaking, Collaboration, Press management when crisis, Communication, Delegating, Relationship with stakeholders.	(Bhaduri, 2019; Stankeviciute & Savaneviciene, 2014; Wisittigars & Siengthai, 2019)
Emotional intelligence	Self-confidence, Motivation, Stress management, Confront issues, Courage, Conflict management, Self-awareness, teamwork skills, relationships management, empathy	(Holt & Wood, 2017; Wisittigars & Siengthai, 2019)
Leadership skills	Making decision, Leading team, Leadership skills, Management skills, Negotiation skills.	(Bhaduri, 2019; Fearn-Banks, 2016; Wisittigars & Siengthai, 2019)
Problem solving	Problem solving, Situation analytical skills, Problem analysis and investigation, Simplify complexity.	(Anderson & Sun, 2017; Wisittigars & Siengthai, 2019)

5.1.2 Lean leadership competencies

Many studies have identified effective leadership as a critical success factor for lean implementation (Alagaraja, 2014; Emiliani, 2003; Marodin & Saurin, 2015). A recent research study delivered a set of leadership competencies that were

refined with the support of lean experts. Based on the fourteen principles of lean, these competencies were developed into the following 16 competencies: identifying what adds value to internal and external customers, identifying and solving problems with teams using the PDCA principle and coaching, continuously using lean practices and principles, managing with a focus on the flow of value rather than isolated operations, problem-solving based on facts and data (Gemba), leading by example and inspiration, stabilize the process, provide value-added information in a clear and objective manner, place the interests of the group above individual interests, develop and implement guidelines, plans and policies aimed at the development of people (Seidel et al., 2017). Building trust, express confidence that team can attain objectives, providing support and encouragement to team, listen to people's concerns, fostering continuous learning and motivation, delegation, visual management, and assessing performance were identified as critical leadership behaviours to lean initiatives' success.

We can therefore conclude that the competencies set of Lean leaders covers the competencies of crisis management leaders, however, for a better response to crisis situations, the competencies of Lean leaders need to be revised to include emergency preparedness related competencies.

As the importance of crisis management competencies has been stated, human resources (HR) managers should develop recruitment and selection processes, training, performance management and compensation systems to meet the needs of crisis management. Concerning recruitment processes, they should set appropriate criteria for recruiting candidates who fulfil the required competencies.

5.2 Lean as a culture of continuous learning

The culture of continuous learning developed by the lean philosophy that provides continuous training and development programs has several advantages. Firstly, individual and collective competencies are strengthened. In addition, training can help impart organizational knowledge and capabilities that contribute to the company's strategic objectives. It can also serve multiple purposes: It can increase job satisfaction and boost employee motivation, develop employees' ability to adapt to change, increase profitability, foster innovation and creativity, reduce staff turnover by strengthening the employer's brand image and disseminate the company's values (Latukha, 2016).

As a result, lean companies have benefited during this health crisis from the advantage of having loyal and motivated workforce with high levels of ownership due to their continuous learning and continuous improvement philosophy, employees adapt quickly to change, and are involved in problem detection and solving despite the physical and psychological stress caused by the crisis because of the fear of losing their jobs.

5.3 Lean wastes during the COVID-19 crisis

Lean Manufacturing calls for activating the value stream at the customer's request (Just in Time JIT) with the focus on eliminating waste and non-value-added tasks that the customer is not willing to pay for (Gupta & Jain, 2013). Waste can come in the form of transportation, inventory, movement, waiting, over-processing, over-production, and defects (Gupta & Jain, 2013).

Within the COVID-19 crisis, one of the pillars of lean management is being re-examined. The heavy reliance on just-in-time delivery systems has demonstrated its weaknesses (Sarkis et al., 2020). Although used to improve production

efficiency and flexibility, JIT is subject to high operational stress caused by such a crisis (Bryce et al., 2020).

Researchers and practitioners have argued that the JIT has not really helped production neither logistics in this pandemic period (Bryce et al., 2020; Sarkis et al., 2020).

In such a situation, there are two options:

- The first option is that overproduction and stock should no longer be considered as types of wastes when a crisis is detected, rather, work with a well-dimensioned stock to compensate market fluctuations would be safer, this option requires designing an emergency plan to deal with crises.
- The second option is to consider local sourcing, in sourcing, and consider different suppliers instead of being dependent on one or two suppliers (Chinese and Indian). Yet, this option is considered to respect lean practices since it offers a decrease in transportation and an increase of quality as types of wastes.

The positive impacts on the planet, the people and the profit of the second option are huge as it offers a green and socially responsible solution.

5.4 *JIT in crisis times*

The COVID-19 pandemic has revealed an important issue related to the way companies manage their inventories. Companies need to rethink a hybrid system that is not totally based on the just-in-time, but which takes into consideration the reduction of the risk of breakdowns and then increases the availability and reliability of the supply chain (Tissir et al., 2020).

5.5 *Lean offers a powerful and skilled workforce*

Although lean may seem vulnerable to crises, because of its lower inventories as cited previously, its power remains in daily training and practice of its workforce to solve small problems which prepare them for such unprecedented crisis (Puvanavarman et al., 2008). Their behaviours and practices to react and learn from reaction, the developed relationships and teamwork, the independence, autonomous, and engaged workforce have it all to better cope with the crisis.

5.6 *Lean and respect for people*

One of the lean pillars is respect for people (Coetzee et al., 2019), Lean managers in this crisis times have put psychological and physical safety of the employees and customers first, the value second, and the cost third as an equation to make a balance in these unprecedented times.

5.7 *Lean CI culture, change culture, and capabilities development during the crisis*

The other pillar of lean is the continuous improvement (Stimec & Grima, 2019). This approach is based on the generation of creative improvements ideas from the frontline workers to solve the problems that they encounter while the top and middle management ensure the communication of objectives, directions and coaching as an essential element in creating the CI culture which has proven to attenuate the stress level and achieving or surpassing the company's strategic objectives (Stimec & Grima, 2019).

Lean has invested years ago in building frontline's problem-solving capabilities and trust as a sustainable capital to the organisation. It is based on the capacity of trained employees to change and adapt constantly seeking perfection through incremental small improvements. The culture of change and adapt is normal in mature lean organizations. As a result, the frontline employees deal with the changes of customers' needs

and government decisions automatically just like they used to. They didn't wait for top management to provide solutions like traditional top-down management, instead they focused on supporting and enabling teams as well as working on safety issues. Here are some recommended adopted changes in lean environments to gain the buy-in of people and to manage within these constant changing conditions:

- Smart working and reducing the number of employees on-site;
- Changing and redesigning workplaces to be safer: reorganize the workspaces (Plexiglas), the tools (talking to machines so as not to touch them anymore), the ways of working (smart working), environment (relocating production against the dependency linked to globalization);
- Using PDCA to cope with changes;
- Evaluate the financial health of your key suppliers to assess the potential impact to your supply chain;
- Dealing with employees fear and anxiety;
- Training and coaching employees;
- Ensuring communication even virtually;
- Creating virtual locations for daily meeting and daily problem-solving to stay focused;
- Create a safe work environment;
- Using visual management for communicating new security instructions, encouraging expression and feeling communication;
- Having good relationships with leaders and teamwork helped building trust and feeling confident towards leaders' decisions;
- Sharing value-added information and explaining the measures taken;
- Explain to people that we are learning with time and reacting accordingly: the process that was adopted is Plan, Execute, and Adjust;
- Identifying customers' needs;
- Prepare plans on how to proceed with reduced number of employees caused by contamination or other needs;
- Ensuring communication between frontline and leaders;
- Production planning change every day;
- Supporting employees in a daily manner, send them videos of motivation and new safety standards and instructions;
- Ensuring Gemba walks while keeping social distance
- Maintaining the standard and the consistent quality is the objective;
- Programming Shifts and employees on-site;
- After lockdowns, motivate people to get back to workplace and provide coaching for them;
- Leadership teams focus on taking care about customers, employees, business and suppliers;
- Having quick feedbacks and reactions;
- Making plans to help people dealing with stress and ensure check-in on employees' well-being;
- Encouraging and listening to employees' ideas and give them feedback;
- Being always flexible and adaptable;
- Continuously communication with suppliers and customers to adapt our planning.

6 LESSONS LEARNED FROM THE COVID-19 CRISIS AND FUTURE CONSIDERATIONS

During those times when Covid-19 was a catalyst of transformations, Lean leaders need to continue to learn from the crisis to provide feedback on what worked well and what was more difficult to implement in companies to highlight good practices that can be sustained, and assessing the performance of employees and their productivity within the new implemented strategies.

6.1 Remote work

Face-to-face contact seems to be essential for the efficiency and creativity of work teams, but the remote work will be an important element of the new normality. Therefore, there is an urgent need to find a new balance between smart and on-site work. This poses a challenge in human resources management, management practices and control systems.

Hybrid work requires rethinking the organization of work in an agile way, with, on the one hand, management modes adapted to the performance of activities on site, within the framework of regular meetings and project activities, and, on the other hand, specific ways of managing remote work. Workspaces need also to be evolved both within the company to make them even more relevant in terms of interactive practices, and at home to guarantee an ideal work space. Three working hours will be recognized: the time spent on the office space; the non-localized time of online availability; and the non-programmable time of creativity.

The lessons that most companies have learned are that good training, good meetings, and good relationships remain possible even virtually and that the lean thinking, effective change management, mindset of continuous improvement, and autonomous workforce are even more valuable in crisis. As a result of their openness to improvement and their flexibility, some lean organizations have concluded that it was more relevant to work remotely. Investment in new online platforms is needed to facilitate remote leadership. They need to use their power in waste elimination to build new strategies based on adopting hybrid working (smart and normal) in order to ensure green and responsible outcomes without compromising the business results. To do so, measurement of the possible increase in productivity fostered by smart working, and determination of how it leads to new ways of collaboration is needed along with its effect on employees' well-being and health and safety. On the one hand, with remote working, people will have more autonomy, but other points need to be highlighted. It was set up as a matter of urgency, further checking of the conformity of the work equipment based on the standards expected in an office is needed. In this respect, Lean management is inconceivable without particular attention to the ergonomics and the layout of workstations, but also to the autonomy of teams, either on company or at home. The need to reverse the habits of decision making was also emphasized, giving priority to collective intelligence rather than few managers.

6.2 Space reduction

Since the presence on site will be limited to certain days of the week, office space will also be reduced and huge financial savings will be made. However, for long-term success, this generalization relies on two essential conditions: the quality of the technical connections and a space dedicated at home with an adapted family organization.

This raises two essential questions for the company: the risk of weakening social ties and the feeling of belonging to the company, and the implementation of management by objectives

and results, with different management methods and skills to be acquired by employees.

6.3 Digitalisation

The crisis has also been an unexpected booster for the digitalization of organizations. The distance has brought the teams closer together. Whether it is the need to exchange in the face of the isolation of the organization or the urgent need to decide together and quickly, this crisis has seen an increase in the number of virtual meetings. In addition to meetings, training, recruitment and social dialogue have been held in distance. The conviviality in the company has developed virtually via the Web.

At a time when companies will be faced with savings and the need to reduce their CO2 emission, the efficiency will be critical and reducing travel based on remote working will be an irresistible opportunity. It is therefore a good bet that the digitalization undergone during the crisis will lead to the desired modernization of working methods.

6.4 Supply chain management

Post-Covid-19 management was a good incentive to focus on proximity. COVID-19 may accelerate companies' quest for ways to reduce supply chain risk and the likelihood of future disruption. Building agility and speed into the supply chain by creating micro-supply chains; decentralized, agile, with flexible suppliers and relationships with manufacturing closer to the point of purchase, and investigating strategies to purchase where you manufacture, and manufacture where you sell has become a necessity.

Lean managers need to assess opportunities to diversify the supplier base in order to maintain JIT or hybrid production without supply chain disruption. Identify geographically diverse suppliers to on-board in the event of emergency. Consideration should be given to dual-sourcing for critical components. Moreover, moving away from a short-term approach toward a long-term strategy of driving value is critical. Periodically test and revise your supply chain strategy taking into account organizational growth and environmental change.

6.5 Human capital competencies' and management

Among the winning functions in the COVID-19 crisis, HR has established itself as an effective lever for the implementation of transformation measures.

The humanization of the lean managerial practices, the exchange of information between hierarchies, the building of the feeling of trust between managers and employees, and the successful and long lasting interpersonal relationships have proven their fruits in mature lean environments through increased ownership and implication of their employees throughout this crisis. In these times of crisis, it is time to put people, their health, safety and well-being first like it was always recommended by Lean. The time has come to reinvent a social contract capable of sustaining creativity and creating a workforce with a dynamic capacity for resilience as well as conducting social justice through reconciling acting well and well-being. Finally, to ensure all the change previously cited, investment in intangible capital is necessary because it provides technological and soft skills required for the success of post-Covid-19 organizational transformation (digitalisation, supply chain management, remote work).

7 CONCLUSION

This paper attempts to examine the role of Lean in overcoming and resisting the changes caused by COVID-19 crisis which was

an indicative time of the level of maturity of organizations in the face of the challenges of the 21st century. This research describes how the Lean organizations have adapted, the difficulties encountered, the innovations that emerge and the management problems that arise with the crisis. This study was conducted by first analysing the fallouts of the crisis, then the leadership competencies required for crisis management in relation to the leadership competencies required for Lean. Analysis of the pillars, wastes and principles of Lean approach towards the consequences of this pandemic was also conducted, as well as improvements and changes to be made to Lean based on Lewin change process to deal with the fallouts of the current crisis and to resist to future crises.

Based on the consequences and lessons learned from the crisis, we suggested rethinking supply chain management, considering digitalization, combining lean and agile, rethinking a hybrid production model (JIT and normal), considering a hybrid work system based on smart and normal work, and finally reviewing the needs in terms of human resources to ensure the successful implementation of these actions aimed at improving the lean philosophy. The advantage of these improvements lies in the fact that they have a positive impact on the environment as well as on the competitiveness of organizations without compromising the profitability of companies.

All in all, Lean companies have shown a more effective response to this crisis thanks to their ability to change and adapt to customer requirements. Yet, efforts have to be made to change their management strategies to better manage the new normal and prevent future crises. It is now necessary to opt for strategies that take into account the following points: shorter supply chains and consideration of insourcing/local sourcing, best practices in internal and external communication, automation, new human resources practices (recruitment), better adaptability, and so on. These points represent avenues for improving lean that will be the subject of future research.

8 REFERENCES

- Alagaraja, M. (2014). A Conceptual Model of Organizations as Learning-Performance Systems. *Human Resource Development Review*, 13(2), 207–233. <https://doi.org/10.1177/1534484313495852>
- Allaoui, A., & Benmoussa, R. (2020). Employees' attitudes toward change with Lean Higher Education in Moroccan public universities. *Journal of Organizational Change Management*, 33(2), 253–288. <https://doi.org/10.1108/JOCM-08-2018-0232>
- Anderson, M. H., & Sun, P. Y. T. (2017). Reviewing Leadership Styles: Overlaps and the Need for a New 'Full-Range' Theory. *International Journal of Management Reviews*, 19(1), 76–96. <https://doi.org/10.1111/ijmr.12082>
- Asnan, R., Nordin, N., & Othman, S. N. (2015). Managing Change on Lean Implementation in Service Sector. *Procedia - Social and Behavioral Sciences*, 211, 313–319. <https://doi.org/10.1016/j.sbspro.2015.11.040>
- Betancourt, J. R., Tan-McGrory, A., Kenst, K. S., Phan, T. H., & Lopez, L. (2017). Analysis & commentary: Organizational change management for health equity: Perspectives from the disparities leadership program. *Health Affairs*, 36(6). <https://doi.org/10.1377/hlthaff.2017.0022>
- Bhaduri, R. M. (2019). Leveraging culture and leadership in crisis management. *European Journal of Training and Development*, 43(5–6), 554–569. <https://doi.org/10.1108/EJTD-10-2018-0109>
- Boin, A., & 'T Hart, P. (2003). Public Leadership in Times of Crisis: Mission Impossible? In *Public Administration Review* (Vol. 63, Issue 5). <https://doi.org/10.1111/1540-6210.00318>
- Bovey, W. H., & Hede, A. (2001). Resistance to organisational change: The role of defence mechanisms. *Journal of Managerial Psychology*, 16(7). <https://doi.org/10.1108/EUM000000006166>
- Boyatzis, R. E. (2009). Competencies as a behavioral approach to emotional intelligence. *Journal of Management Development*, 28(9), 749–770. <https://doi.org/10.1108/02621710910987647>
- Bryce, C., Ring, P., Ashby, S., & Wardman, J. K. (2020). Resilience in the face of uncertainty: early lessons from the COVID-19 pandemic. *Journal of Risk Research*. <https://doi.org/10.1080/13669877.2020.1756379>
- By, R. T. (2005). Organisational change management: A critical review. *Journal of Change Management*, 5(4), 369–380. <https://doi.org/10.1080/14697010500359250>
- Chen, H. Y., Das, A., & Ivanov, D. (2019). Building resilience and managing post-disruption supply chain recovery: Lessons from the information and communication technology industry. *International Journal of Information Management*, 49. <https://doi.org/10.1016/j.ijinfomgt.2019.06.002>
- Coetzee, R., van Dyk, L., & van der Merwe, K. R. (2019). Towards addressing respect for people during lean implementation. *International Journal of Lean Six Sigma*, 10(3), 830–854. <https://doi.org/10.1108/IJLSS-07-2017-0081>
- Cook, M. D., Holt, M. K., Cook, M. D., & Holt, M. K. (2015). Crunch Time: When Leadership Matters Most. *Journal of Commerce and Trade*, 10(1), 41–47. <https://econpapers.repec.org/RePEc:jct:journl:v:10:y:2015:i:1:p:41-47>
- Cummings, S., Bridgman, T., & Brown, K. G. (2016). Unfreezing change as three steps: Rethinking Kurt Lewin's legacy for change management. *Human Relations*, 69(1), 33–60. <https://doi.org/10.1177/0018726715577707>
- Dasborough, M., Lamb, P., & Suseno, Y. (2015). Understanding emotions in higher education change management. *Journal of Organizational Change Management*, 28(4). <https://doi.org/10.1108/JOCM-11-2013-0235>
- Emiliani, M. L. (2003). Linking leaders' beliefs to their behaviors and competencies. *Management Decision*, 41(9), 893–910. <https://doi.org/10.1108/00251740310497430>
- Fearn-Banks. (2016). *Crisis Communications: A Casebook Approach - Kathleen Fearn-Banks - Google Livres*. <https://books.google.co.ma/books?id=aOPLDAAQBAJ&printsec=frontcover&dq=Fearn+banks+2016&hl=fr&sa=X&ved=2ahUKEwiJ4Or9xOrtAhV5SxUIHbVeC8oQ6AEwAHoECAAQAg#v=onepage&q=Fearn+banks+2016&f=false>
- Fernandes, N. (2020). *Economic effects of coronavirus outbreak (COVID-19) on the world economy*. <https://ssrn.com/abstract=3557504>
- Fooladvand, M., Haghshenas, A., Yarmohammadian, M., Alavi, A., & Atighechian, G. (2013). Crisis management competencies and their relation to self-efficiency of managers in the organizations related to disasters. *International Journal of Health System and Disaster Management*, 1(2), 64. <https://doi.org/10.4103/2347->

- 9019.128113
- Groh. (2014). STRATEGIC MANAGEMENT IN TIMES OF CRISIS. *American Journal of Economics and Business Administration*, 6(2).
<https://doi.org/10.3844/ajebasp.2014.49.57>
- Gupta, S., & Jain, S. K. (2013). A literature review of lean manufacturing. *International Journal of Management Science and Engineering Management*, 8(4).
<https://doi.org/10.1080/17509653.2013.825074>
- Hoek, R. van. (2020). Responding to COVID-19 Supply Chain Risks—Insights from Supply Chain Change Management, Total Cost of Ownership and Supplier Segmentation Theory. *Logistics*, 4(4), 23.
<https://doi.org/10.3390/logistics4040023>
- Holt, S. S., & Wood, A. (2017). Leadership and emotional intelligence. In D. S. Marques J. (Ed.), *Leadership Today* (Issue 5, pp. 111–138). Springer Texts in Business and Economics. Springer, Cham.
https://doi.org/10.1007/978-3-319-31036-7_7
- Home | *Lean Frontiers*. (n.d.). Retrieved March 6, 2021, from <https://leanfrontiers.com/>
- Howe, D. C., Chauhan, R. S., Soderberg, A. T., & Buckley, M. R. (2020). Paradigm shifts caused by the COVID-19 pandemic. *Organizational Dynamics*, 2019, 100804.
<https://doi.org/10.1016/j.orgdyn.2020.100804>
- Ivanov, D., & Das, A. (2020). Coronavirus (COVID-19/SARS-CoV-2) and supply chain resilience: A research note. *International Journal of Integrated Supply Management*, 13(1). <https://doi.org/10.1504/IJISM.2020.107780>
- Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research*, 58(10).
<https://doi.org/10.1080/00207543.2020.1750727>
- Jasti, N. V. K., & Kodali, R. (2014). Lean production: Literature review and trends. *International Journal of Production Research*, 53(3), 867–885.
<https://doi.org/10.1080/00207543.2014.937508>
- Jin, X. (Cathy), Qu, M., & Bao, J. (2019). Impact of crisis events on Chinese outbound tourist flow: A framework for post-events growth. *Tourism Management*, 74, 334–344. <https://doi.org/10.1016/j.tourman.2019.04.011>
- Kamble, S., Gunasekaran, A., & Dhong, N. C. (2020). Industry 4.0 and lean manufacturing practices for sustainable organisational performance in Indian manufacturing companies. *International Journal of Production Research*, 58(5), 1319–1337.
<https://doi.org/10.1080/00207543.2019.1630772>
- Kuzmanova, M., & Ivanov, I. (2019). Relation Between Change Management and Crisis Management: Survey Evidence. *International Conference KNOWLEDGE-BASED ORGANIZATION*, 25(1), 255–260.
<https://doi.org/10.2478/kbo-2019-0042>
- Latukha, M. (2016). Talent Management in Emerging Market Firms. In *Talent Management in Emerging Market Firms*. <https://doi.org/10.1057/978-1-137-50606-1>
- Lewin, K. (1947). Frontiers in Group Dynamics. *Human Relations*, 1(1), 5–41.
<https://doi.org/10.1177/001872674700100103>
- Linnenluecke, M. K. (2017). Resilience in Business and Management Research: A Review of Influential Publications and a Research Agenda. *International Journal of Management Reviews*, 19(1), 4–30.
<https://doi.org/10.1111/ijmr.12076>
- Lockwood, N. R. (2005). Crisis Management in Today's Business Environment: HR's Strategic Role. In *SHRM®Research Quarterly*. www.cmiatl.com.
- Maffioli, E. M. (2020). Perspective piece: How is the world responding to the novel coronavirus disease (COVID-19) compared with the 2014 west African Ebola Epidemic? The importance of China as a player in the Global Economy. In *American Journal of Tropical Medicine and Hygiene* (Vol. 102, Issue 5).
<https://doi.org/10.4269/ajtmh.20-0135>
- Marodin, G. A., & Saurin, T. A. (2015). Classification and relationships between risks that affect lean production implementation: A study in southern Brazil. *Journal of Manufacturing Technology Management*, 26(1), 57–79.
<https://doi.org/10.1108/JMTM-12-2012-0113>
- Mena, C., Melnyk, S. A., Baghersad, M., & Zobel, C. W. (2020). Sourcing Decisions under Conditions of Risk and Resilience: A Behavioral Study. *Decision Sciences*, 51(4), 985–1014. <https://doi.org/10.1111/deci.12403>
- Mishra, L., Gupta, T., & Shree, A. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *International Journal of Educational Research Open*, 1, 100012.
<https://doi.org/10.1016/j.ijedro.2020.100012>
- Muñoz, P. ;, Kimmitt, J. ;, Kibler, E. ;, & Farny, S. (2019). *Living on the slopes: entrepreneurial preparedness in a context under continuous threat*.
<https://doi.org/10.1080/08985626.2018.1541591>
- Nakat, Z., & Bou-Mitri, C. (2021). COVID-19 and the food industry: Readiness assessment. *Food Control*, 121, 107661. <https://doi.org/10.1016/j.foodcont.2020.107661>
- Netland, T. H. (2016). Critical success factors for implementing lean production: The effect of contingencies. *International Journal of Production Research*, 54(8), 2433–2448.
<https://doi.org/10.1080/00207543.2015.1096976>
- Oláh, J., Szolnok, A., Nagy, G., Lengyel, P., & Popp, J. (2017). The Impact of Lean Thinking on Workforce Motivation: A Success Factor at LEGO Manufacturing Ltd. *Journal of Competitiveness*, 9(2), 93–109.
<https://doi.org/10.7441/joc.2017.02.07>
- Pearson, C. M., & Clair, J. A. (1998). Reframing crisis management. *Academy of Management Review*, 23(1).
<https://doi.org/10.5465/AMR.1998.192960>
- Pflüger, M., & Tabuchi, T. (2019). Comparative advantage, agglomeration economies and trade costs. *Journal of Urban Economics*, 109.
<https://doi.org/10.1016/j.jue.2018.11.003>
- Portuguez Castro, M., & Gómez Zermeño, M. G. (2020). Being an entrepreneur post-COVID-19 – resilience in times of crisis: a systematic literature review. In *Journal of Entrepreneurship in Emerging Economies*. Emerald Group Holdings Ltd. <https://doi.org/10.1108/JEEE-07-2020-0246>
- Poulsen, T., & Lema, R. (2017). Is the supply chain ready for the green transformation? The case of offshore wind logistics. In *Renewable and Sustainable Energy Reviews* (Vol. 73). <https://doi.org/10.1016/j.rser.2017.01.181>
- Puvanasvaran, A. P., Megat, M. H. M. A., Tang, S. H., Muhamad, M. R., & Hamouda, A. M. S. (2008). A review of problem solving capabilities in lean process management. In *American Journal of Applied Sciences* (Vol. 5, Issue 5, pp. 504–511). Science Publications.
<https://doi.org/10.3844/ajassp.2008.504.511>
- Rapaccini, M., Saccani, N., Kowalkowski, C., Paiola, M., &

- Adrodegari, F. (2020). Navigating disruptive crises through service-led growth: The impact of COVID-19 on Italian manufacturing firms. *Industrial Marketing Management*, 88(May), 225–237. <https://doi.org/10.1016/j.indmarman.2020.05.017>
- Rosenbaum, D., More, E., & Steane, P. (2018). Planned organisational change management: Forward to the past? An exploratory literature review. In *Journal of Organizational Change Management* (Vol. 31, Issue 2). <https://doi.org/10.1108/JOCM-06-2015-0089>
- Sarkis, J., Cohen, M. J., Dewick, P., & Schröder, P. (2020). A brave new world: Lessons from the COVID-19 pandemic for transitioning to sustainable supply and production. In *Resources, Conservation and Recycling* (Vol. 159). <https://doi.org/10.1016/j.resconrec.2020.104894>
- Seidel, A., Saurin, T. A., Marodin, G. A., & Ribeiro, J. L. D. (2017). Lean leadership competencies: a multi-method study. *Management Decision*, 55(10), 2163–2180. <https://doi.org/10.1108/MD-01-2017-0045>
- Sinapin, M. N. (2020). *Les entreprises et l'après crise de Covid-19 : un nouveau modèle d'organisation agile entre efficience et résilience*.
- Singh, K., Saeed, M., & Bertsch, A. (2012). Key Factors Influencing Employee Response Toward Change : A Test in the Telecom Industry in India. *Journal of Management Policy and Practice*, 13(3).
- Singh, S., Kumar, R., Panchal, R., & Tiwari, M. K. (2020). Impact of COVID-19 on logistics systems and disruptions in food supply chain. *International Journal of Production Research*, 1–16. <https://doi.org/10.1080/00207543.2020.1792000>
- Sommer, S. A., Howell, J. M., & Hadley, C. N. (2016). Keeping Positive and Building Strength: The Role of Affect and Team Leadership in Developing Resilience During an Organizational Crisis. *Group and Organization Management*, 41(2). <https://doi.org/10.1177/1059601115578027>
- Stankeviciute, Z., & Savaneviciene, A. (2014). SUSTAINABILITY AS A CONCEPT FOR HUMAN RESOURCE MANAGEMENT. *ECONOMICS AND MANAGEMENT*, 18(4). <https://doi.org/10.5755/j01.em.18.4.5631>
- Stimec, A., & Grima, F. (2019). The impact of implementing continuous improvement upon stress within a Lean production framework. *International Journal of Production Research*, 57(5), 1590–1605. <https://doi.org/10.1080/00207543.2018.1494391>
- Taj, S., & Morosan, C. (2011). The impact of lean operations on the Chinese manufacturing performance. *Journal of Manufacturing Technology Management*, 22(2), 223–240. <https://doi.org/10.1108/17410381111102234>
- Tissir, S., El Fezazi, S., & Cherrafi, A. (2020). Lean management and industry 4.0 impact in covid19 pandemic era. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, August, 3123–3129.
- Tortorella, G., Narayanamurthy, G., Godinho Filho, M., Portioli Staudacher, A., & Mac Cawley, A. F. (2020). Pandemic's effect on the relationship between lean implementation and service performance. *Journal of Service Theory and Practice*. <https://doi.org/10.1108/JSTP-07-2020-0182>
- Tubin, D. (2017). Every teacher carries a leadership wand. *International Journal of Leadership in Education*, 20(2). <https://doi.org/10.1080/13603124.2015.1028463>
- Vakola, M., & Nikolaou, I. (2005). Attitudes towards organizational change: What is the role of employees' stress and commitment? *Employee Relations*, 27(2). <https://doi.org/10.1108/01425450510572685>
- van der Smissen, S., Schalk, R., & Freese, C. (2013). Organizational change and the psychological contract. *Journal of Organizational Change Management*, 26(6). <https://doi.org/10.1108/jocm-08-2012-0118>
- van Ee, M. (2021). Approximability of the dispersed p→-neighbor k-supplier problem. *Discrete Applied Mathematics*, 289. <https://doi.org/10.1016/j.dam.2020.10.007>
- Wisittigars, B., & Siengthai, S. (2019). Crisis leadership competencies: the facility management sector in Thailand. *Facilities*, 37(13–14), 881–896. <https://doi.org/10.1108/F-10-2017-0100>
- Womack, J. P., & Jones, D. T. (1996). From lean production to the lean enterprise. *IEEE Engineering Management Review*.
- Zanni Tim. (2020, April). *Tech COVID-19 supply chain disruption - KPMG Global*. <https://home.kpmg/xx/en/blogs/home/posts/2020/04/technology-supply-chain-disruption.html>
- Zhang, S. X., Wang, Y., Rauch, A., & Wei, F. (2020). Unprecedented disruption of lives and work: Health, distress and life satisfaction of working adults in China one month into the COVID-19 outbreak. *Psychiatry Research*, 288. <https://doi.org/10.1016/j.psychres.2020.112958>

Preserving resilience while managing changes, an assessment framework

ERIC RIGAUD

CENTRE OF RESEARCH ON RISKS AND CRISIS MANAGEMENT (CRC),
MINES PARISTECH, PARIS-SCIENCES-ET-LETTRES,
RUE CLAUDE DAUNESSE 1,
06904 SOPHIA ANTIPOLIS CEDEX, FRANCE
ERIC.RIGAUD@MINES-PARISTECH.FR

Résumé – Les changements technologiques ou managériaux visent à améliorer les performances, la productivité générale, la rentabilité ou la qualité des systèmes sociotechniques. Dans certaines circonstances, les changements peuvent affecter, comme effet secondaire, les propriétés de sécurité du système et faciliter la survenue d'accidents ou bien diminuer la capacité du système à prévenir et atténuer les effets d'un accident. L'anticipation des potentielles conséquences de changement sur la résilience des systèmes sociotechniques nécessite des connaissances, des méthodes et des outils prenant en considération leur complexité. Cet article présente un cadre méthodologique fondé sur l'Ingénierie de la Résilience visant à identifier les potentiels impact de changements sur les comportements individuels et collectifs dans une perspective de gestion de sécurité.

Abstract – Technological or managerial changes aim to improve socio-technical systems' performance, general productivity, profitability, or quality. In some circumstances, changes can affect, as a side effect, system safety properties and facilitate the occurrence of crises or by decreasing system ability to mitigate crises. Anticipating potential consequences of change on socio-technical systems' safety properties requires knowledge, method, and tools covering socio-technical' complexity. This paper presents a Resilience Engineering-based methodological framework aiming to support the identification of the impact of a change on individual and collective behaviours within the perspective of safety management.

Mots clés – Changements, Impacts, Sécurité, Résilience, Hommes

Keywords – Change, Impacts, Safety, Resilience, Humans

1 INTRODUCTION

Movements, adaptations, changes are organizations' new routines (Alter 2003). Evolutions in their cultural, political, social, technological, economic, physical, or legal environment impose organizations to initiate change management projects to preserve their competitive advantage and survive (Porter 2008, Hatch 2012).

Changes in a complex system can lead to different types of consequences (Merton 1936, Morin 1990): positive, unexpected benefits usually referred to as serendipity or a windfall; adverse effects, which occur in addition to the desired effect of the change; Perverse effects, i.e., the unexpected adverse effect is greater than the expected beneficial effect; the futility of innovation, e.g., the more things change, the more they stay the same; threat of achievements, i.e., the consequences are the opposite of the initial intentions.

Consideration of change is one of the prerequisites of a proactive safety management system. The anticipation of opportunities and threats induced by change or innovation is one of the cornerstones of organizational resilience (Hollnagel et al., 2006). ISO 31000 monitoring and review requirements include detecting changes in the external and internal context, requiring revision of risk management processes, and identifying emerging risks (NF ISO 31000).

Managing change is generally associated with efforts to overcome resistance induced by a change by identifying potential sources of resistance and applying the appropriate

change management strategy. Such approaches aim to make the change effective. However, several examples of accidents and disasters illustrate that successful changes can contribute to unwanted consequences, and research needs to be conducted that complements change management approaches with safety concerns.

This paper presents a methodological framework aiming to define a strategy of assessment guiding experiments to prevent the failure of change and anticipate the occurrence of unwanted impacts on the resilience of sociotechnical systems. The first section presents a synthesis of the theoretical background integrated within the framework. Then, the second describes the framework, and the last one illustrates its application.

2 THEORETICAL BACKGROUND

Leading a change must overcome various obstacles that can negate the efforts and investments made. Societal, organizational, and technological constraints can be at the origin of the rejection of a seemingly relevant change (Owens 2012). Existing frameworks dedicated to change management in maritime, nuclear, or railway used risk assessment, generally technical risk assessment, and cost benefits. The proposed framework aims to extend the scope of potential impacts considered when conducting change management by considering targets related to the different risk management periods: impacts on the risk of technical failure, human and organizational factors, and system resilience (cf. figure 1).

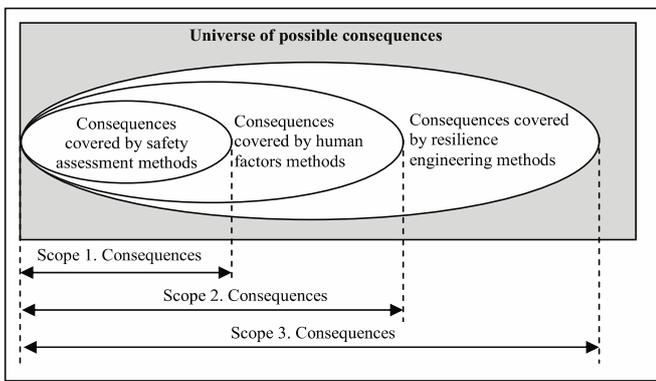


Figure 1. Type of potential impact considered

2.1 Impacts on the risks of technical failures

The need for safe and reliable technologies has led to the development of risk assessment methods to prevent failures and the associated consequences. Based on the practices of the chemical, aeronautics, or nuclear industries, the field of Operational Safety (Villemeur, 1988; Lannoy, 2008) structures a set of fundamental concepts such as reliability (ability of an entity to perform a required function under given conditions and for a given duration), availability (the ability of an entity to perform a required function under given conditions and at a given time) or safety (the ability of an entity to avoid the occurrence, under given conditions, of critical or catastrophic events). Tools like the frequency-severity matrix make it possible to classify risks according to their probability of occurrence and the severity of their potential consequences. The preliminary risk analysis (RPA), the analysis of failure modes, their effects, and their criticality (FMECA), or the Hazard and Operability Analysis (HAZOP) make it possible to qualify, quantify and describe the risk scenarios.

2.2 Impacts on the risks of human errors

Human error was first associated with the deviation from a standard, a procedure, or a frame of reference. Then, it considers different behavior modes and associated errors: inattention in a routine situation, applying an inappropriate rule in a regulated situation, and inability to adopt an acceptable behavior in a situation for which no rule has been formalized (Rasmussen and Jensen, 1974). Reason (1993) proposes a typology of human error by distinguishing between faults, failures, selection, sequence, time, or quality errors. Work is carried out on information processing modes that distinguish different reasoning modes depending on the situation, particularly the SRK model (Rasmussen, 1983), which distinguishes behavior guided by skills, rules, or experience. The first methods of analyzing human reliability were developed, such as the THERP method, Technique for Human Error Rate Prediction (Swain and Guttman, 1983), which makes it possible to calculate a probability of error in the execution of a task in combining a decomposition of the task using a tree representation and error probability tables. The occurrence of major technological disasters in the mid-1980s, such as the Bhopal, Chernobyl, or Challenger accidents, questioned the theories, models, and practices of risk management and led to the consideration of the organizational characteristics of the systems.

2.3 Impacts on the risks of organizational failure

The development, in sociology, of theories about industrial accidents, offers new perspectives on accidents and risk

management needs by integrating organizational topics (hierarchical system, culture, or technology).

The idea that a period in which visible symptoms are not perceived and not considered by the system preceded accident or disaster is first stated (Turner, 1978). Vaughan enunciates the normalization of deviance theory (1996) and corroborates this idea. Based on the study of the evolution of NASA preceding the shuttle Challenger explosion, this theory shows that the organization accepts disruptions and deviations from the rules and gradually generates a routinization of deviations, creating the conditions for the occurrence of a disaster.

The normal accident (Perrow, 1984), and the High-Reliability Organizations (Laporte and Consolini, 1991) theories are developed following the Three Miles Island nuclear accident. The first theory states that the complexity of organizations (centralization, decentralization, or redundancy.) is such that it is almost impossible to represent them correctly and consequently to identify and prevent failure scenarios. Also, it states that risk management barriers contribute to the complexity of the systems and thus increase the systems' vulnerability. Therefore, the accident is inevitable because it is not the result of failures but of the normal functioning of systems whose complexity is greater than the ability to control risk management systems.

The second theory enunciates that organizational mechanisms (culture and priority of leaders, organizational and technical redundancy, adequate decision-making and communication structure, organizational culture, continuous training of agents, organizational learning) promote prevent failures and disasters. New methods such as CREAM, Cognitive Reliability, and Error Analysis Method (Hollnagel, 1998) or TRIPOD (Groeneweg et al., 1998) are developed. Their ambition is to consider organizational factors when performing risk analysis and accidental investigation tasks.

2.4 Impacts on system resilience

Resilience Engineering aims to foster sociotechnical systems with the requisite imagination and the resources to respond and overcome the diversity and the complexity of possible situations (Adamski et Westrum 2003, Woods et Hollnagel 2006). While risk management provides basements to prevent failures identified when analyzing accidents, Resilience Engineering aims to prevent systems from losing the potential of adaptation to their changing environment. Among many definitions, the ones proposed by Woods (2006) and Hollnagel (2011) influence theoretical and practical researches within Resilience Engineering. Woods (2006) defines resilience as the "ability to recognize and adapt to handle unanticipated perturbations that call into question the model of competence, and demand a shift of processes, strategies, and coordination." Hollnagel (2011) suggests that resilience refers to "the intrinsic ability of a system to adjust its functioning before, during or following changes and disturbances so that it can sustain required operations under both expected and unexpected conditions." Hollnagel (2013) considers two approaches for safety management. "Safety I" focus on managing failures and "Safety II" on the management of the system capacity responding and overcoming disturbances.

3 A FRAMEWORK FOR PRESERVING RESILIENCE DURING CHANGES

The framework aims to define a strategy of assessment guiding experiments to prevent the failure of change and anticipate the occurrence of unwanted impacts on sociotechnical systems'

resilience. Two processes constitute the framework. The first objective is to create a culture of resilience and formalize a system resilience model. When a change occurs, the second process aims to study it to prevent unexpected consequences. Safety managers can use the framework to enhance their understanding of the system's complexity and structure change management activities to improve their security. They can apply it at a different scale (technological system, process, unit, or plant).

3.1 Roles and responsibilities

A set of essential roles supports the distribution of responsibilities when applying the framework, considering that one person can assume different roles.

The “evaluation owner” is the person who is mainly responsible for the system to be assessed. This critical role encompasses the following responsibilities: defining the goal and scope of the evaluation process, supporting the assessment team in providing access to the agents of the system, and to documents and resources needed by the assessment (room, material, etc.).

The “evaluation coordinator” is the person who is mainly responsible for the evaluation process. The evaluation coordinator should cover the following responsibilities: defining the target, the scope, and the objective of the evaluation process with the “evaluation owner,” planning the different steps of the assessment, monitoring the realization of the different steps, managing issues when performing the different steps.

The “stakeholder coordinator” is the person who is mainly responsible for the coordination with the various agents involved in the assessment. The stakeholder coordinator should cover the following responsibilities: identifying the agents, invite the agents to workshops, provide feedbacks of the assessment to the agents.

The “technical coordinator” is the person who is mainly responsible for the realization of the assessment task. The evaluation coordinator should cover the following responsibilities: organizing and animating workshops, writing deliverables.

3.2 Defining resilience performance.

For defining the resilience performance of the studied system, the team firstly organized workshops for describing the system studied, the diversity of events it has to respond to, and its capacity to respond. The team designs an assessment methodology and associated supportive material (diagnostic schedule, interviews and observation guidelines, and assessment grid). The team conducts individual and collective interviews and observations for collecting qualitative and quantitative data about the system structure and dynamic in regular times and when disturbances occur with considering the different actors of the system (operational, managers, and directors).

3.3 Promoting a culture of resilience preservation while conducting changes.

For integrating into the organization's safety culture, the relevance of studying the consequences of resilience performance changes: The team share lessons learned about situations where change contributes to unwanted consequences. They formulate a methodology of managing the impact of changes' consequences on resilience. They develop related expertise within the organization. They control that consequences of change on resilience management are carried

out with sufficient time so that change managers can take into consideration potential adverse impacts. Finally, they diffuse lessons learned within the organization.

3.4 Studying changes for preventing unwanted consequences

For conducting investigations on the potential consequences of a change on the system's resilience, the team achieves the following tasks. Firstly, the team provided a model of the change. This model is used to define the assessment strategy and design the different data collection and analysis modules. Then, the different modules are conducted, and a list of consequences is deduced. The team deduces from the consequences a list of risks and opportunities associated with the change. This list is used to provide recommendations.

4 LESSONS FROM THE APPLICATION OF THE FRAMEWORK

The framework's ongoing development progresses with the application of instances of the framework within different industrial contexts.

4.1 Application within the gas transportation system

The first instance of the framework was applied on a case of change within the gas transportation system (Zarea et al. 2010, 2013a, 2013b).

The change was an innovative workflow aiming to detect, identify and localize threats susceptible to affect pipeline integrity: 1) Acquisition of image data with geographic reference via a lightweight autonomously operating UAV; 2) Automatic image pre-processing: image mosaicking and geo-referencing; 3) Automatic change detection: automatic detection of changes (threat candidates) using object-based image analysis techniques; 4) Visual threat candidate validation: Software guided validation of threat candidates.

The team designs a strategy aiming at assessing: 1) the relevance of using an automated aerial surveillance system to monitor a pipeline section; 2) the legality to use an automated aerial surveillance system within an area; 3) the feasibility to use an automated aerial surveillance system within an area; 4) the pipeline stakeholders' acceptability to use automated aerial surveillance.

The experiment's campaign firstly consisted in the realization of experimental flight in France, Spain, and Germany. Secondly, surveys, interviews, and role games to identify issues related to acceptance of civil aviation, acceptance of safety regulators, general public acceptance, and acceptance of traditional service providers doing pipeline surveillance.

The studies' results were promising; nevertheless, due to the various constraints, a set of requirements were proposed to support change management with minimizing risks of failure, efficiency, and reliability of the workflow, conformity with changing regulation, and acceptability by both operators and society.

4.2 Application within the maritime context

A refined instance of the framework was applied to study the potential impact of using the 3D technologies by ship pilot (Rigaud et al. 2012a).

The team designs a strategy aiming at assessing: 1) the relevance of using a 3D chart when piloting ship 2) maritime stakeholders acceptability constraints.

The experiment's campaign firstly consisted of the simulation of search and rescue operations using bridge simulators. The team compares pilot behaviors with situations where they used

a traditional radar and used a 3D chart. Secondly, focus groups on identifying maritime stakeholders' perception of the 3D chart and potential obstacles to overcome to be used on bridges.

The studies' results demonstrate the ship bridges simulator's relevance to operationalize human factors risk assessment methods such as situation awareness, stress, or communication risks assessment. They demonstrate the potential of the change but the existence of regulatory obstacles.

A dedicated team used lessons from this first case study to study the potential added value and constraints of using a visual analytic environment when monitoring maritime areas' safety and security (Rigaud et al. 2014, Vatin et al. 2014). A focus on the impact of data quality and its potential consequences on monitoring performance was analyzed.

4.3 Application within the air traffic control context

The framework was also applied to study the potential impact of a technological change within the air traffic control network (Rigaud et al. 2012b, Martinie et al. 2012).

The team provides a model of the air traffic control system's resilience and develops an integrated model to support technological failure propagation and escalation within a complex network.

The model was used to study the potential consequences of using a fully automated decision support system in tower control under various conditions and the potential of systemic failure.

4.4 Application within the railway context

An extended version of the framework was provided following its application within the railway context (Rigaud et al. 2013, Cote et al., 2014a, 2014b, Rigaud 2017, Rigaud et al., 2017, 2018).

A dedicated module aimed at supporting the definition and assessing the resilience of a sociotechnical system was elaborated. More particularly, a set of indicators related to the system's capacity to respond to the regular, irregular, and unexampled threats was developed so as a dedicated data collection and analysis method.

An extended version of the framework was provided following its application within the railway context (Rigaud et al. 2013, Cote et al., 2014a, 2014b, Rigaud 2017, Rigaud et al., 2017, 2018).

A dedicated module aimed at supporting the definition and assessing the resilience of a sociotechnical system was elaborated. More particularly, a set of indicators related to the system's capacity 1) to respond and overcome the diversity of situations that may arise; 2) to monitor that which changes or may change in the near term that it will require a response; 3) to learn from both positive and negative experience of the past; 4) to anticipate development, threats, and opportunities further into the future.

The team formalize seven situations of resilience to be considered when collecting data and assessing the resilience capacity:

1. The situation is normal, considered by procedure or good practices, and the context (time, knowledge, competencies, and information) necessary to respond is available. Agents can recognize the situation, define their future behaviour by using their experience or with adapting a known and regularly applied procedure, and apply it in conformity with all the dimensions of performance of the activity.
2. The situation is normal, considered by procedure or good practices. However, the context (time, knowledge,

competencies, information) necessary to respond is not available. Agents can recognize the situation, define their future behaviour by using their experience or adapting a known and regularly applied procedure and apply it with creativity to conform with all dimensions of the activity's performance despite the lack of one kind of resource.

3. The situation is normal and not considered by procedure or good practices. Agents can recognize the situation and that neither procedure nor good practices support them to define the behaviour to adopt, they are creative to define their future behaviour and apply it in conformity with all dimensions of the activity's performance.
4. The situation is abnormal (perturbation, crisis, etc.), considered by procedure or good practices, and the context (time, knowledge, competencies, and information) necessary to respond is available. Agents can recognize the situation and the necessity to adopt a non-routine behaviour, they define their future behaviour with using their experience or with adapting a known procedure or find one in a guideline, they apply it in conformity with all dimensions of the activity's performance in contributing to the continuity of the activity of the system.
5. The situation is abnormal, considered by procedure or good practices, but the context (time, knowledge, competencies, information) necessary to respond is not available. Agents can recognize the situation and the necessity to adopt a non-routine behaviour; they define their future behaviour by using their experience or with adapting a known procedure or find one in a guideline, and apply it with creativity in order to conform with all dimensions of the activity's performance despite the lack of one kind of resources.
6. The situation is abnormal and not considered by procedure or good practices. Agents can recognize the situation and the necessity to adopt a non-routine behaviour. Neither procedure nor good practices support them to define the behaviour to adopt. They are creative in defining their future behaviour and applying it in conformity with all the activity's performance dimensions.
7. The situation is unexampled. Agents are creative to respond and to contribute to the continuity of activity of the system.

These situations structure the change management module where a guidelines to organize interviews and focus groups on identifying how a change can positively or negatively affect one or several situations is targeted.

4.5 Lessons from applications

The application and the refinement of the framework within different context allow the identification of a set of lessons:

1. The definition of resilience performance is time-consuming and requires the operator to understand and consider specific concepts and contextualize them to their context.
2. Modeling changes and their potential impacts on the system requires integrating several models and a specific simulation framework.
3. Agents are willing to discuss how they adapt when disturbances happen. Nevertheless, talking about actions at the limit or outside the procedural context is difficult with the hierarchy's presence.
4. Budget optimization policies make it difficult to realize changes aiming to consider the analysis results.
5. Considering the impact of change on the risk of failure is complicated, even within the organization where

anticipating unwanted consequences of changes is present. Succeed in applying the change is generally considered as the priority.

6. Technological simulator provides insights for identifying unimagined impacts.

5 CONCLUSION

This article presented a framework whose purpose is to identify the potential impact of changes on systems' security. Based on the theoretical framework of Resilience Engineering, it aims to extend the potential for preventing the unintended consequences of a change in individual and collective behaviour, allowing control of the diversity of situations that may arise. Applying the method requires time and resources to collect and organize the knowledge necessary to model and assess resilience. Although the results provide recognized added value, taking them into account in operational and strategic decisions requires an evolution of organizations' safety culture.

The framework's perspective is to adapt it in order to be able to apply it for studying the impact of trends, changes, and crises on urban resilience.

6 RÉFÉRENCES

- Alter, N. (2003). Mouvement et dyschronies dans les organisations. *L'Année sociologique*, vol. 53(2), 489-514.
- Côte T., Rigaud E., Garbolino E. (2014a). Anticipating impacts of change on individual behaviors within a perspective of safety management. *24th European Safety and Reliability Conference - ESREL 2014*, Wroclaw, Poland. pp.1011-1018 - ISBN 9781138026810.
- Côte T., Rigaud E., Garbolino. E. (2014b). Diversité et complexité de la mise en œuvre du principe « globalement au moins aussi équivalent » (GAME). *Congrès λ 19 (Lambda Mu 19) - 19e Congrès de Maîtrise des Risques et Sécurité de Fonctionnement – IMDR*, Dijon, France.
- Grossetti M. (2004) *Sociologie de l'imprévisible. Dynamiques de l'activité et des formes sociales*. Presses Universitaires de France, 225p.
- Groeneweg Jop et Roggeveen Victor, « Tripod: Controlling the human error components in accident », *Safety and Reliability*, Lydersen, Hansen and Sandtorv, 1998, p. 809-816.
- Hatch M.J. & Cunliffe A. L. (2013). *Organization theory*. Oxford university press.
- Hollnagel Erik, *Cognitive Reliability and Error Analysis Method*, Oxford, Elsevier Science and Technology, 1998, 302 p.
- Hollnagel, E. (2011). RAG – The resilience analysis grid. In: E. Hollnagel, J. Pariès, D. D. Woods & J. Wreathall (Eds) *Resilience Engineering in Practice. A Guidebook*.
- Hollnagel E. (2013). Making health care resilient: From safety-I to safety-II. In Hollnagel E., Braithwaite J. and Wears R. L. (eds.) *Resilience healthcare*. Farnham, UK: Ashgate.
- Hollnagel E. & Woods D. D. (2006). "Epilogue: Resilience Engineering Precepts." In *Resilience Engineering: Concepts and Precepts*. Ashgate.
- Laporte Todd R. et Consolini Paula M., « Working in Practice but not in Theory: Theoretical Challenges of "High-Reliability Organizations" », *Journal of Public Administration Research and Theory*, vol. 1, n° 1, janv. 1991, p. 19-48.
- Lannoy André, *Maîtrise des risques et sûreté de fonctionnement. Repères historiques et méthodologiques*, Paris, Lavoisier, 2008, 128 p.
- Martinie C., Palanque P., Pasquini A., Ragosta M., Rigaud E. et al. (2012a). Using Complementary Models-Based Approaches for Representing and Analysing ATM Systems' Variability. *2nd International Conference on Application and Theory of Automation in Command and Control Systems - ATACCS '12*, May 2012, London, United Kingdom. pp.146-157 - ISBN 978-2-917490-20-4.
- McDonald N. (2006). Organizational Resilience and Industrial Risk", in *Resilience Engineering: concepts and precepts*, Hollnagel E., Woods D. D. and Leveson N. (eds), pages 143–168. Ashgate.
- Merton R. (1936). The Unanticipated Consequences of Purposive Social Action, in « *American Sociological Review* », vol. 1.
- Morin, E. (1990). *Introduction à la pensée complexe*. Paris : Edition du seuil.
- NF ISO 31000 Risk Management – Principles and guidelines
- Owens D.A. « Creative People must be stopped. Six ways we kill innovation (without even trying) », Jossey-Bass 2012
- Perrow Charles, *Normal Accidents. Living with High-Risk Technologies*, New York, Basic Books, 1984, 456 p.
- Porter, M. (2008). *On competition*. Harvard business school publishing corporation.
- Rasmussen Jens et Jensen Alice Rolandini, « Mental procedures in real-life tasks: a case study in electronic troubleshooting », *Ergonomics*, vol. 17 n° 3, 1974, p. 293-307.
- Rasmussen Jens « Skills, rules, and knowledge: signals, signs and symbols and other distinctions in human performance models », *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13, n° 3, mai-juin 1983, p. 257-267.
- Reason James, *L'erreur humaine*, Paris, PUF, coll. « Le Travail humain », trad. J. M. Hoc, *Human Error*, Cambridge, Cambridge University Press, 1990.
- Rigaud. E. (2017). Lessons from the design of a Resilience Engineering based socio-technical system diagnostic method. *7th REA Symposium, 'Poised to Adapt: Enacting resilience potential through design, governance and organization.'* Liège, France.
- Rigaud E., Lützhöft M., Kircher A., Schröder-Hinrichs J. W., Baldauf M. (2012a) Impact: More Than Maritime Risk Assessment. *Procedia - Social and Behavioral Sciences*, Elsevier, 2012, 48, pp.1848-1854
- Rigaud E., Hollnagel E., Martinie C., Palanque P., Pasquini A., et al. (2012b) A framework for modeling the consequences of the propagation of automation degradation: application to air traffic control systems. *The Second SESAR Innovation Days*, Nov 2012, Braunschweig, Germany. 8 p. - ISBN 978-2-87497-068-9.
- Rigaud E., Neveu C., Duvenci-Langa S., Obrist M. N., Rigaud. S. (2013). Proposition of an organisational resilience assessment framework dedicated to railway

- traffic management. Edited by Nastaran Dadashi, Anita Scott, John R Wilson, Ann Mills. *Rail Human Factors: Supporting Reliability, Safety and Cost Reduction*, Taylor & Francis.
- Rigaud E., Côte T., Vatin G., Napoli A. (2014). La conduite du changement digital au défi de l'acceptabilité et de la sécurité. *Questions de Management*, 2014, 7 (3), pp.143 -153.
- Rigaud E., Neveu C., Duvenci-Langa S. (2017) Lessons from the design of a Resilience Engineering based train management in station system diagnostic method. *Sixth International Human Factors Rail Conference*, London, United Kingdom.
- Rigaud E., Neveu C., Duvenci-Langa S. (2018). Lessons from the application of a resilience engineering based assessment method to evaluate the resilience of a train departure and arrival management system. *Safety and Reliability – Safe Societies in a Changing World, ESREL 2018*, Jun 2018, Trondheim, Norway.
- Snowden D. J. and Boone M. E. (2007). A leader's framework for decision making. *Harvard Business Review*, 85(11) :68–76.
- Turner Barry A. et Pidgeon Nick F., *Man-Made Disasters*, Boston, Butterworth-Heinemann, 1978, 250 p.
- Vatin G., Rigaud E., Napoli A. (2014). Geovisual Analytics Environment for Supporting the Resilience of Maritime Surveillance System. *CESUN 2014, 4th International Engineering Systems Symposium*, Jun 2014, Hoboken, NJ, United States. 7 p.
- Vaughan Diane, *The Challenger launch decision. Risky technology, culture and deviance at NASA*, Chicago, University of Chicago Press, 1996, 592 p.
- Villemeur Alain, *Sûreté de fonctionnement des systèmes industriels*, direction des études et recherches d'Électricité de France, Paris, Eyrolles, 1988, 795 p
- Weick K. E. (1993). Collapse of sensemaking in organizations: the Mann Gulch disaster". *Administrative Science Quarterly*, 38(4) :628–652.
- Westrum R. (2006). A typology of resilience situations, in *Resilience Engineering: concepts and precepts*, Hollnagel E., Woods D. D. and Leveson, N. (eds), Ashgate.
- Zarea M., Pognonec G., Lecchhi M., Lana J., Sohlbach M., Rigaud E. (2013a) Anticipating Risks of Innovative Safety Barriers: Findings on an Automated Aerial Surveillance System. *5th IntegRisk Conference "Horizon 2020: From Integf-Risk Radar to the European Emerging Risk Radar*, May 2013, Stuttgart, Germany.
- Zarea M., Pognonec G., Schmidt c., Schnur T., Lana J., Rigaud E. et. al. (2010). First steps in developing an automated aerial surveillance approach. 2. *iNTeg-Risk Conference 2010 "New Technologies & Emerging Risk: Dealing with multiple and interconnected emerging risks"*, Jun 2010, Stuttgart, Germany. pp.223-236.
- Zaréa M., Pognonec G., Schmidt C., Schnur T., Lana J., Rigaud E. et. Al. (2013b). First steps in developing an automated aerial surveillance approach. *Journal of Risk Research*, Taylor & Francis (Routledge), 2013, 16 p. 407-420.

Session 1.3 : Design et Pilotage des systèmes de production de biens et de services

Chairman : Vincent Giard - Khaled Hadj-Hamou

Vers une approche globale pour soutenir le processus Adaptive Sales and Operations Planning

JEAN-BAPTISTE VIDAL¹, RAPHAEL OGER¹, MATTHIEU LAURAS¹, JACQUES LAMOTHE¹, ROMAIN MICLO²

¹ IMT Mines Albi, Centre Génie Industriel
IMT Mines Albi, France
jean.vidal@mines-albi.fr
jacques.lamothe@mines-albi.fr
matthieu.lauras@mines-albi.fr
raphael.oger@mines-albi.fr

²AGILEA
Toulouse, France
romain.miclo@agilea-group.com

Résumé –

Dans un contexte de plus en plus incertain, les entreprises sont toujours à la recherche de l'optimisation de la planification de leur supply chain. Depuis 2011, des auteurs proposent une nouvelle approche appelée DDAE/DDMRP censée permettre de réagir de manière agile dans un contexte complexe et volatile, le graal est-il là ? Cette approche rencontre un succès grandissant au niveau des industriels et des scientifiques mais ce système de planification hiérarchisée multi-niveaux (opérationnel, tactique, stratégique) comporte de nombreuses zones d'ombres qui restent à éclaircir. Cette étude se porte sur le niveau stratégique de ce modèle : le processus AS&OP. A date, AS&OP n'étant pas suffisamment précisément défini et outillé, nous proposons de développer un système d'aide à la décision dédié. Cette étude apporte deux contributions distinctes : une modélisation du processus métier AS&OP ainsi qu'une approche logicielle associée permettant de soutenir l'aide à la décision. Nous étudions le potentiel de l'utilisation des arbres de décisions et de la gestion des risques pour développer une planification scénarisée dans cette optique. Cette plateforme d'expérimentation appelé AS&OP-DSS est testée sur un cas illustratif. L'article conclut sur la faisabilité de cette approche et ouvre sur un ensemble de perspectives de recherche.

Abstract –

In an increasingly uncertain context, companies are always looking for relevant means able to optimize their supply chain planning. Since 2011, some authors propose a new approach called DDAE/DDMRP which is supposed to allow reacting in an agile way in a complex and volatile context. Is the grail there? This approach is meeting with growing success among practitioners and scholars, but this multi-level hierarchical planning system (operational, tactical, strategical) has many grey areas that remain to be cleared up. This study focuses on the strategic level of this model: the AS&OP process. For now, AS&OP is not accurately defined and equipped. Consequently, we propose to develop a dedicated decision support system. This study brings two distinct contributions: a modeling of the AS&OP business processes and an associated software approach to support the decision mechanism. We study the potential of using decision trees and risk management approaches to develop a scenario-oriented planning with this perspective. This experimentation platform called AS&OP-DSS is tested on a theoretical case. The article concludes on the feasibility of this approach and opens avenues for future research.

Mots clés – S&OP, AS&OP, DDMRP, système d'aide à la décision, gestion des risques, planification scénarisée

Keywords – S&OP, AS&OP, DDMRP, Decision Support System, Risk Management, scenario planning

1 INTRODUCTION

La mondialisation des marchés et la concurrence accrue ont conduit les entreprises à maximiser leur rentabilité tout en tenant mieux compte des incertitudes et des variations de leur contexte. En matière de *Supply Chain Management* (SCM), les praticiens et les chercheurs ont développé de multiples approches pour gérer cette situation : Juste à temps (JIT), Théorie des contraintes (TOC), LEAN ou encore *Material*

Requirements Planning (MRP2). L'objectif étant de pouvoir continuer à engendrer des bénéfices tout en optimisant l'utilisation des ressources de l'entreprise. Mais comme le démontre (Miclo et al., 2019), la plupart de ces méthodes ne parviennent pas à traiter correctement les incertitudes et les variabilités croissantes auxquelles les entreprises doivent faire face aujourd'hui. Dans un contexte où les ruptures de chaîne d'approvisionnement deviennent la norme, la récente crise liée à la COVID-19 en étant un bel exemple, une innovation récente se distingue : le *Demand Driven Adaptive Enterprise* ou DDAE.

Depuis 2011, le DDAE tente de faire une synthèse des pratiques des précédentes méthodes de planification de production, et y ajoute quelques innovations (Smith & Smith, 2014). (Miclo et al., 2019) démontrent que cette approche permet d'être plus agile et robuste dans un contexte perturbé que les méthodes Kanban et MRP2. Cependant, malgré l'intérêt croissant des praticiens et des universitaires pour le DDAE, celle-ci n'en est encore qu'à ses débuts et nécessite des contributions supplémentaires pour consolider à la fois la méthode elle-même ainsi que les systèmes d'aide à la décision associés (Miclo et al., 2019). Le *Demand Driven Institute* (DDI (DDI, (s. d.))) est une association de praticiens internationaux du SCM, du LEAN et de la TOC qui a consacré beaucoup d'efforts au cours de la dernière décennie à décrire par le détail le niveau opérationnel du DDAE dit DDOM (*Demand Driven Operation Model*). Malheureusement, le DDI demeure beaucoup plus vague sur les niveaux tactique et stratégique de la méthode. Ainsi, le niveau stratégique, appelé AS&OP (*Adaptive Sales and Operations Planning*) a rarement été étudié et gagnerait à faire l'objet d'études supplémentaires pour soutenir des mises en œuvre concrètes (J.-B. Vidal et al., 2018). Parmi tous les axes d'amélioration possibles, une question clef concerne la faisabilité de l'utilisation d'une approche *what-if analysis* ou planification scénarisée pour soutenir l'aide à la décision dans un environnement incertain. En effet, pour gagner en robustesse et en agilité dans un environnement DDAE, (C. Ptak & Smith, 2019) préconisent que les équipes de direction prennent leurs décisions non pas sur la base d'un unique scénario d'étude, ni même de quelques-uns, mais bien sur la base d'une multitude intégrant des hypothèses de risques et d'opportunités (fournisseur qui fait faillite, nouveau marché, variation des ventes, capacité, paramètres du modèle opératoire, etc.). Bien que les praticiens valident le principe de cette orientation, il n'existe pas aujourd'hui d'approche globale structurée, ni d'outil d'aide à la décision capable de générer et traiter ces scénarios dans une optique AS&OP prenant en compte les spécificités de l'approche DDAE.

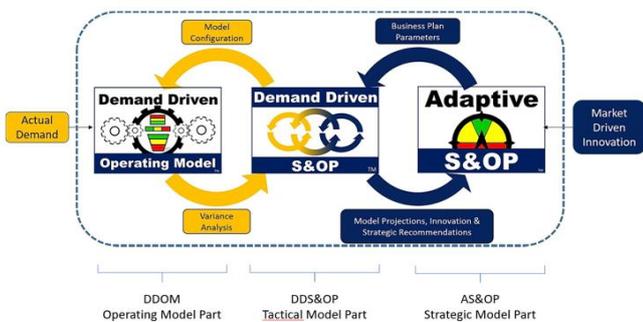


Figure 1. Modèle DDAE (C. Ptak & Smith, 2019)

Ce travail de recherche se concentre finalement sur cet énoncé du problème et tente (i) de démontrer la faisabilité de l'utilisation d'une approche de planification scénarisée pour soutenir le processus AS&OP, et (ii) de discuter des avantages et des limites d'une telle approche.

La suite du document est composée de 4 sections principales. Premièrement, une analyse de la littérature est développée afin de positionner les questions de recherche et la contribution attendue. Deuxièmement, un modèle intégrant une approche planification scénarisée est proposé pour soutenir la prise de décision. Troisièmement, un cas illustratif est présenté afin de valider la contribution et discuter de ses avantages et de ses limites. Quatrièmement, une conclusion rapide est donnée et des pistes de recherche futures sont proposées.

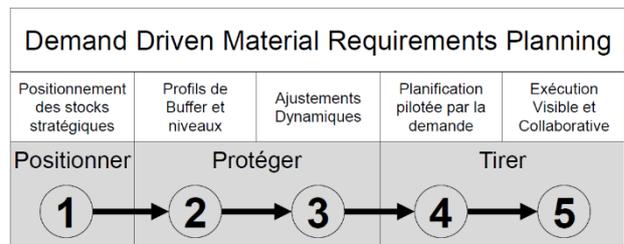
La question de recherche suivante peut être formulée : *Comment développer un système d'aide à la décision soutenant efficacement le processus AS&OP ?*

2 REVUE DE LITTÉRATURE

2.1 Fondamentaux du DDAE et du DDMRP

Depuis 2011, le *Demand Driven Institute* (DDI) a conçu, développé et popularisé le modèle DDAE par des formations, des conférences, des mises en œuvre et des publications (C. Ptak & Smith, 2019; Martin et al., 2018; Smith & Smith, 2014; Demand Driven Institut, (s.d.)). Fondamentalement, le modèle DDAE suit une structure hiérarchique de planification classique (Figure 1.), comme pour le MRP2 (C. Ptak & Smith, 2019). En général, il comprend trois niveaux de planification consacrés respectivement aux questions opérationnelles (*Demand Driven Operation Model DDOM*), tactiques (*Demand Driven Sales and Operations Planning DDS&OP*) et stratégiques (*Adaptive Sales and Operations Planning AS&OP*).

Le DDMRP fait partie intégrante du DDOM (Bahu et al., 2019). Le DDMRP est un système de gestion des flux présenté dans (C. A. Ptak & Smith, 2011) qui consiste à mixer en une même approche les avantages d'un système à flux poussé et d'un système à flux tiré. La mise en œuvre d'un système DDMRP est présenté dans (C. Ptak & Smith, 2019), elle se fait en plusieurs étapes : (i) positionnement stratégique des buffers



de stock, de temps ou de capacité, (ii) dimensionnement des buffers, (iii) ajustement dynamique de ces buffers (iv) planification pilotée par la demande et (v) exécution des ordres de planification.

Figure 2. Cinq composante de DDMRP (C. Ptak & Smith, 2019)

L'innovation DDMRP réside principalement dans l'intégration de buffer de stock qui est géré par une loi de re-complètement spécifique au comportement dynamique. Chaque buffer de stock est constitué de plusieurs zones (figure 3.) dont les tailles changent au fil du temps en fonction de différents paramètres : la consommation moyenne journalière (*Average Daily Usage ADU*), le délai découplé (*Decoupled Lead Time DLT*), le Facteur de délai (*Lead Time Factor LTF*), le facteur de variabilité (*Variability Factor VF*) et la quantité économique (*Minimum Order Quantity MOQ*). (C. A. Ptak & Smith, 2011) propose des règles de calcul pour chaque zone à partir de ces principaux paramètres. A noter que la zone verte peut être calculée de trois manières différentes en fonction de la prise en compte par exemple d'un MOQ ou alors d'un cycle d'approvisionnement imposé. La façon de dimensionner cette zone verte est choisie par l'entreprise pour la façon la plus pratique de le faire : en général, c'est la valeur maximale de (1), (2) ou (3) (figure 3.). En fonction de la position d'un indicateur d'état du flux (calculé via l'équation du flux disponible (C. Ptak & Smith, 2019)), un ordre de re-complètement est alors généré. Comme ces zones évoluent en fonction de la demande, ceci ajuste automatiquement le comportement de la loi de pilotage du stock géré par buffer, ce qui lui donne bien la caractéristique dynamique attendue.

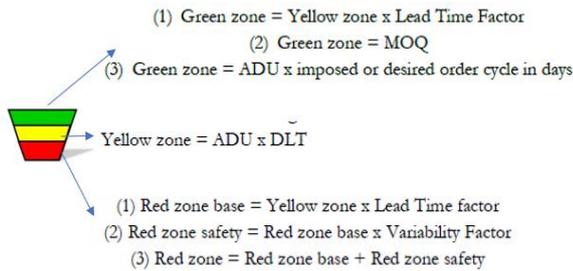


Figure 3. Principe d'un buffer de stock DDMRP (Miclo, 2016)

Pour plus de détail sur le mécanisme du DDMRP dans son ensemble, le lecteur peut se référer à (C. Ptak & Smith, 2019; Bahu et al., 2019; Smith & Smith, 2014; Demand Driven Institute, (s.d.)).

2.2 Récents travaux de recherche autour du DDAE

Plus récemment, les scientifiques ont commencé à étudier le système DDAE (Dessevre et al., 2020; Bayard et al., 2020; Ducrot & Ahmed, 2019; Meinzel, 2019; Miclo et al., 2019; Acosta et al., 2019; Miclo, 2016) notamment en se concentrant sur les spécificités de l'étage opérationnel DDOM. Ils ont notamment démontré l'agilité du DDMRP, en particulier dans des environnements perturbés et incertains, pour maintenir les performances (coûts, délais et stocks). Le comportement du modèle DDMRP dans différents environnement a également été étudié afin de confirmer ou d'infirmer les principales hypothèses proposées par le DDI. Il est à noter qu'à ce jour, la communauté scientifique n'a étudié que très peu l'ensemble du système DDAE, et s'est principalement intéressée aux détails du modèle DDMRP du DDOM. Une explication peut être que le terme DDMRP initialement utilisé par le DDI pour désigner à la fois l'ensemble DDAE en co-construction et son modèle opératoire DDOM est visiblement associé dans la communauté scientifique uniquement à l'élément DDMRP du DDOM. Néanmoins, (Bozutti & Esposto, 2019; Martin et al., 2018) ont introduit chacun une proposition de modélisation formelle des processus DDS&OP et AS&OP, respectivement pour les étages tactique et stratégique du modèle DDAE.

2.3 Processus AS&OP, S&OP et planification scénarisée

(Pittman & Atwater, 2016) définissent Sales and Operations Planning S&OP comme « *Processus de développement des plans tactiques qui donne à la direction les moyens de piloter l'activité pour obtenir un avantage concurrentiel de manière continue, en intégrant dans la chaîne logistique les plans commerciaux des produits actuels ou à venir. Le processus intègre tous les plans de l'entreprise (vente, marketing, développement, production, approvisionnement et financement) dans un ensemble unique* ». Plusieurs articles récents (Pereira et al., 2020; Nabil et al., 2018; Kristensen & Jonsson, 2018; Tavares Thomé et al., 2012; Tuomikangas & Kaipia, 2014) montrent l'intérêt de poursuivre les recherches sur ce processus, soulignant que l'implantation du S&OP reste difficile pour les praticiens et que le sujet n'a pas encore été traité complètement d'un point de vue scientifique. Ils préconisent d'étudier (i) le potentiel de la planification scénarisée via S&OP, (ii) l'intégration de la gestion des risques via S&OP (iii) la considération poussée de l'incertitude. Or (Martin et al., 2018) ont souligné que les caractéristiques de l'AS&OP sont très proches de celles du S&OP mais ont quand même mis en évidence certaines spécificités : (i) positionnement des buffers

au niveau stratégique et (ii) capacité à traiter des opportunités stratégiques.

Concernant le premier point, au niveau du DDOM, la position et la nature des buffers dans le flux sont des éléments dimensionnant importants. Changer ces éléments est un bras de levier possible pour répondre à des opportunités de marché mais leurs effets doivent être estimés et validés par le niveau stratégique pour garantir la capacité à exécuter la nouvelle dynamique qui en résultera au niveau opérationnel. C'est une facette de DDAE qui ne semble pas encore bien étudiée.

Sur le second point, (Oger et al., 2020) proposent un système d'aide à la décision explorant le potentiel de la scénarisation dans un environnement incertain au niveau de la gestion stratégique de la chaîne de logistique globale. Ils démontrent le potentiel de celle-ci et concluent qu'il serait intéressant d'étendre cette piste à des niveaux de granularité plus fin comme celui du S&OP. (Tiberius et al., 2020; Lindgren & Bandhold, 2002) explicitent que la scénarisation reste sous-exploitée pour le traitement des sujets stratégiques. Par extension, il est intéressant d'étudier toutes ces possibilités au travers de l'étude de AS&OP car les praticiens via le DDI reprennent fréquemment ces idées sans pour autant proposer d'outillage pour le faire.

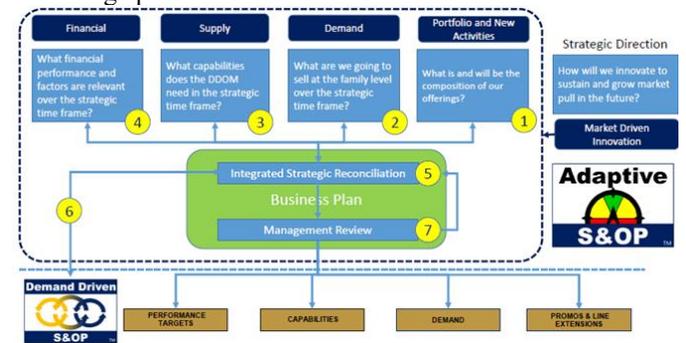


Figure 4. Processus AS&OP (C. Ptak & Smith, 2019)

(J.-B. Vidal et al., 2018) concluent que le processus d'AS&OP (Figure 4.) proposé par le DDI (C. Ptak & Smith, 2019) restent flou quant à l'orchestration des étapes.

(J. Vidal et al., 2020) ont notamment fait remarquer qu'une ambiguïté demeure quant au niveau de granularité à prendre en compte pour soutenir le processus AS&OP ainsi qu'une absence de modèle proposée par le DDI pour calculer ces indicateurs. Nous intégrerons dans notre étude le modèle purement agrégé proposé par (J. Vidal et al., 2020) qui permet un calcul des indicateurs de performance de type charge/capacité et Besoin en Fonds de Roulement (BFR) sur la base de la donnée pertinente de prévisions de vente établies au niveau de la famille de produit.

3 PROPOSITION

Nous considérons qu'un système d'aide à la décision inclut une vision processus métier et un système d'information dédié associé (*Decision Support System DSS*). Les paragraphes suivants présentent nos propositions sur ces deux aspects : (i) processus métier suggéré et (ii) les caractéristiques du DSS associé.

3.1 Processus métier AS&OP

La Figure 5. présente notre proposition au format *Business process model and notation BPMN* d'un processus métier AS&OP intégrant une orchestration des principales étapes vis-à-vis des services de l'équipe AS&OP : direction, finance, opérations industrielles (approvisionnement, production,

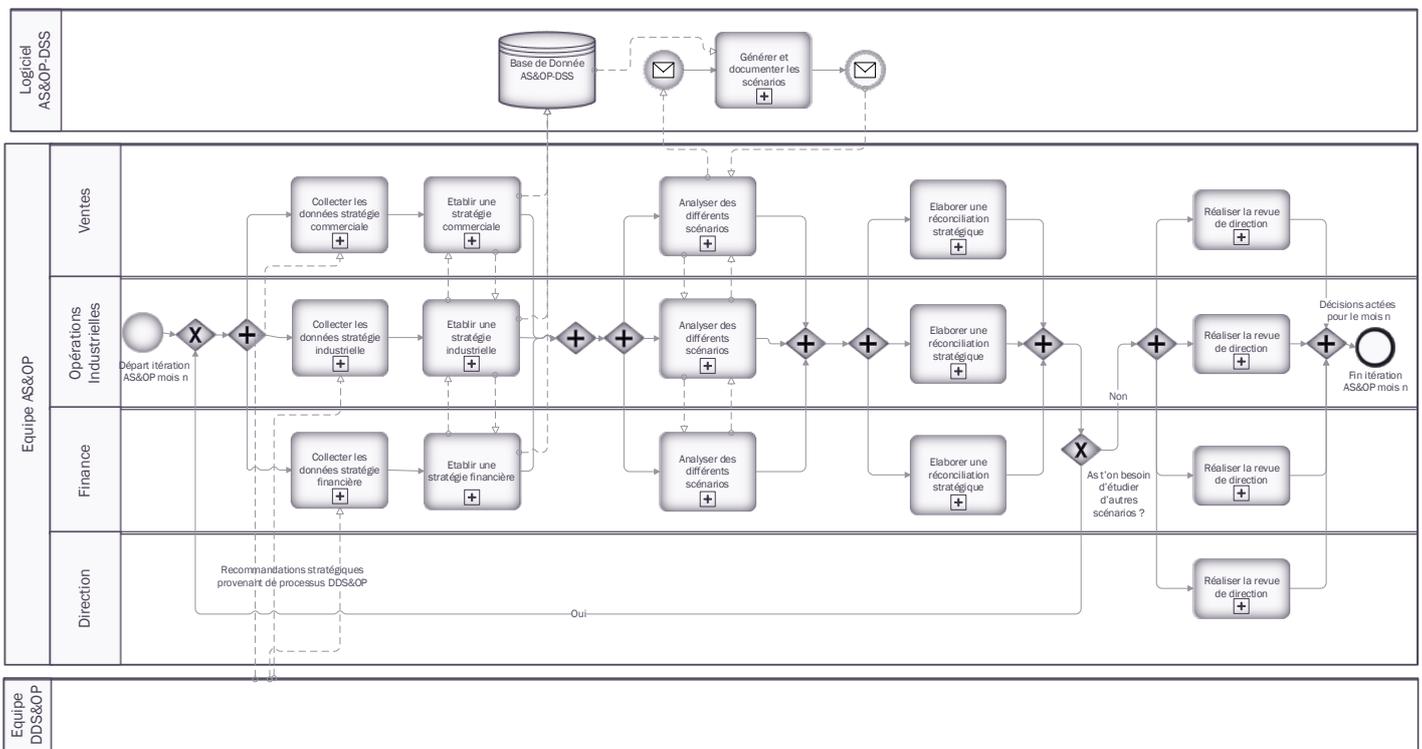


Figure 5. Proposition Processus Métier AS&OP

distribution) et ventes. L'étape « établir une stratégie » intègre l'identification des sources de des risques et des opportunités recommandées par le DDI. Ces données sont utilisées par notre système d'aide à la décision nommé AS&OP-DSS (*Adaptive Sales and Operations Planning Decision Support System*) pour apporter une solution technique à l'étape « Analyser les différents scénarios ». Enfin, l'étape « réaliser la revue de direction » incorpore la préconisation du DDI d'archiver les hypothèses réalisées pour acter les décisions de l'itération du mois n, afin d'en tirer des apprentissages lors des itérations futures du processus AS&OP.

3.2 Caractéristiques du DSS

La Figure 6. donne la vision d'ensemble du processus de fonctionnement de notre proposition AS&OP-DSS dont nous détaillons ci-dessous les principaux éléments.

L'AS&OP-DSS utilise les données suivantes en entrée : (i) les données techniques du scénario nominal (ii) les risques et (iii) les opportunités. Les données techniques incluent toutes les informations habituellement utilisées lors d'un processus S&OP classique (prévisions de vente, prix de vente, niveau des stocks, caractéristiques du système de production...). Les risques et opportunités représentent des événements possibles qui peuvent impacter fortement la performance globale de l'entreprise. Nous définissons une opportunité comme une décision à étudier pouvant apporter un effet bénéfique à la performance de l'entreprise. Le Tableau 1. présente les principales opportunités qui sont adressées au niveau AS&OP. Ces opportunités y sont classées en différents domaines de décision que les praticiens utilisent (Slack & Lewis, 2002). Un risque est un événement pouvant impacter de manière négative cette même performance pouvant être issue d'une des quatre sources suivantes : (i) variabilité de la demande, (ii) variabilité opérationnelle, (iii) variabilité d'approvisionnement ou (iv) variabilité du management.

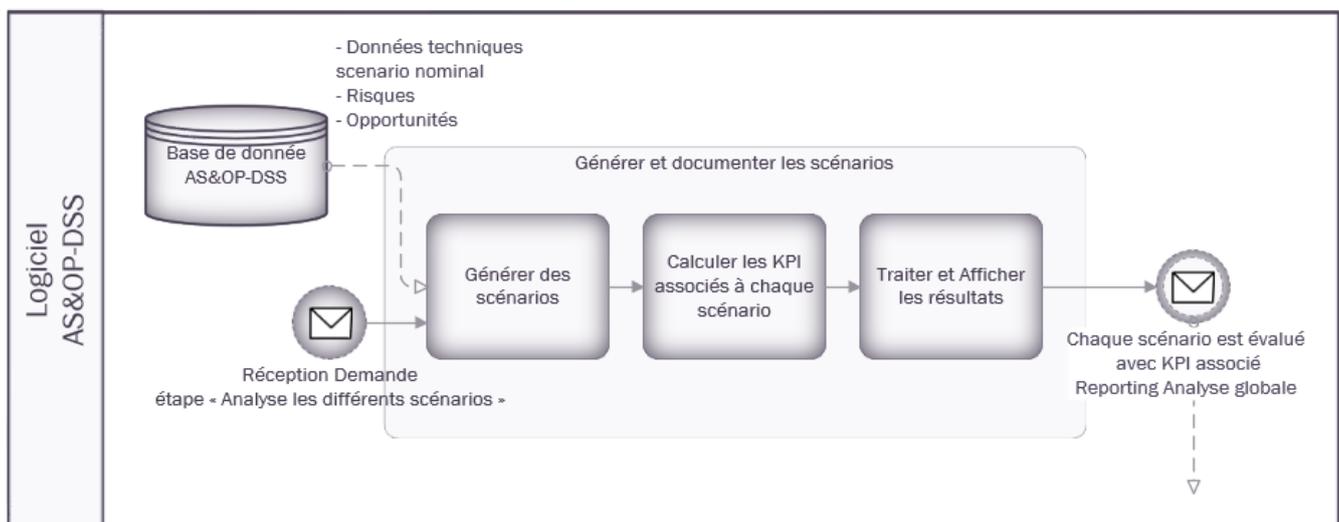


Figure 6. Processus détaillé AS&OP-DSS

A partir de ces données, trois étapes se succèdent : (i) Générer des scénarios, (ii) Calculer les indicateurs de performance associés à chaque scénario et (iii) Traiter et afficher les résultats. Les décideurs de l'équipe AS&OP peuvent alors jouer leurs rôles et acter ou tester de nouveau leurs décisions.

Ce sous-processus Figure 6. constitue le cadre pour structurer nos recherches, nos propositions à réaliser et les discussions associées. Les paragraphes 3.2.1 à 3.2.4 présentent le détail de chacune des étapes de ce modèle AS&OP-DSS

Tableau 1. Exemple d'opportunités adressées durant le processus AS&OP inspiré par (Slack & Lewis, 2002)

	DOMAINES DE DECISION	EXEMPLE
STRATEGIE INDUSTRIELLE	CAPACITAIRE Se préoccupe de la manière dont la capacité et les installations en général devraient être configurées	Capacité d'adaptation/changement
	RÉSEAU D'APPROVISIONNEMENT Concerner la manière dont les opérations sont liées à travers la chaîne logistique globale Capacité d'adaptation/changement	Fabriquer ou acheter un produit ? Changer de fournisseur
	TECHNOLOGIE DES PROCESS Concerner les process utilisés pour transformer les matières premières en produits finis.	Intégrer une nouvelle technologie
	AMELIORATION DE L'ORGANISATION Concerner la manière de convertir l'organisation en une meilleure	Réduire les temps de changement de série,
	GESTION DE PRODUCTION Concerner la manière dont les opérations gèrent la gestion des flux de production	Modification des paramètres du modèle DDMRP : position buffer, type de buffer, dimensionnement des buffers
STRATEGIE VENTE	PORTEFOLIO, NOUVELLES ACTIVITÉS et DEMANDE Se préoccupe de la façon dont le marché devrait être configuré	Se positionner sur un nouveau marché
STRATEGIE FINANCIERE	FINANCE Concerner la gestion des objectifs financiers	Modifier / adapter le niveau des objectifs financiers à atteindre sur une période donnée

3.2.1 Étape « Générer des scénarios »

Notre proposition consiste à structurer la génération des scénarios par le biais d'un arbre de décision. Les éléments de segmentation pris en compte peuvent être classés en trois groupes d'événements issues de l'interprétation des données d'entrées de l'étape : (i) stratégie commerciale, (ii) stratégie industrielle et (iii) risques (Figure 7.) Le groupe « stratégie commerciale » permet d'intégrer les données techniques nominales et les opportunités liées aux domaines de décision des stratégies vente et financière (Tableau 1.). De même pour le groupe « stratégie industrielle » avec les domaines de décisions associés (Tableau 1.). Enfin le groupe « risques » génère des scénarios sur la base de valeurs différentes qu'une donnée pourra prendre dans un intervalle connu suivant une règle de balayage à définir ou sur les variables de modèles ou les données impactées.

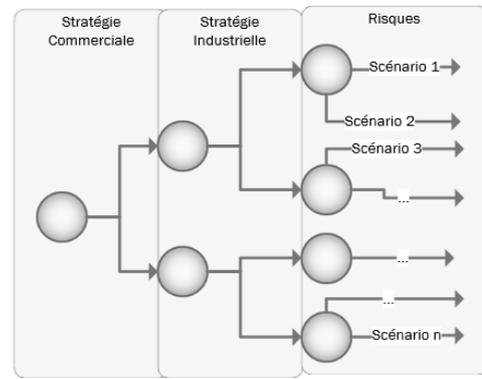


Figure 7. Principe de génération de scénarios

3.2.2 Étape « Calculer les indicateurs de performance associés à chaque scénario »

Pour cette étape, nous prenons en compte les apports d'une précédente étude qui propose un modèle de projection du comportement du DDMRP adapté au niveau stratégique (J. Vidal et al., 2020). Les indicateurs de performance recommandés par (Smith & Smith, 2014) y sont instanciés pour la perspective financière d'une part par le BFR et pour la perspective de capacité d'autre part par le taux de charge sur ressource critique.

3.2.3 Étape « Traiter et afficher les résultats »

Après la précédente étape, nous disposons d'un nombre de scénarios évalués important. Afin de faciliter l'aide à la décision de l'équipe AS&OP, notre proposition est de permettre un affichage sous forme de tableaux de bords flexibles et pouvant être modifiables à loisir de manière manuel par les décideurs, en agrégeant si nécessaires les données de manière cohérente. Cette agrégation est notamment guidée par l'utilisation d'un arbre de décision ciblant les nœuds générés par une option décisionnelle que les décideurs souhaitent voir illustrer et nécessitant de ce fait, le traitement adapté des scénarios associés (Figure 8.).

Aujourd'hui de nombreux logiciels de visualisation de données permettent de réaliser ceci. Nous optons pour l'utilisation du logiciel Tableau Software® (Tableau Software, (s.d.)) en s'inspirant de (Oger, 2019). Le traitement se fera donc de manière manuelle à travers les possibilités de ce logiciel.

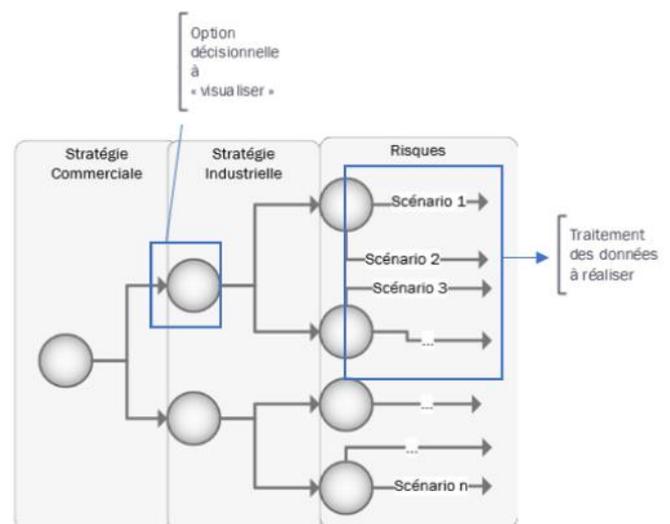


Figure 8. Principe d'utilisation de l'arbre de décision pour traiter et afficher les résultats

4 EXPERIMENTATION ET DISCUSSION

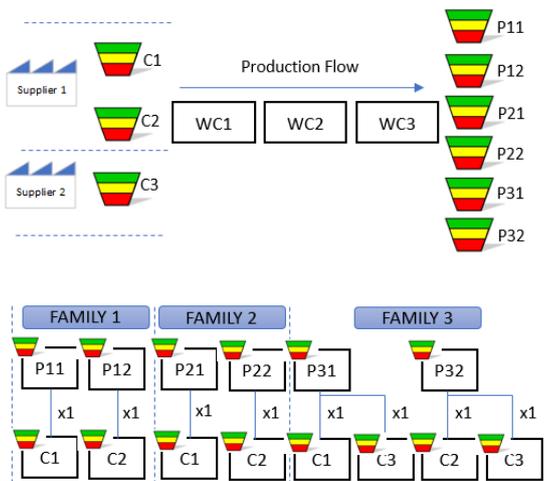
Afin de valider nos propositions, nous réalisons une expérience sur un cas illustratif dans un contexte incluant un scénario nominal, une opportunité et deux risques à travers notre programme AS&OP-DSS. Nous nous positionnons à l'étape « Analyser les différents scénarios » de notre processus AS&OP.

4.1 Plateforme d'expérimentation

Le moteur de la plateforme a été développé en langage Python afin de supporter les étapes « Générer des scénarios » et « Calculer les indicateurs associés à chaque scénario » en intégrant les éléments clés présentés dans la section 3. Un lien a ensuite été établi entre ce moteur et le logiciel de visualisation de données Tableau Software® (Tableau Software, (s.d.)). L'ensemble constitue notre plateforme d'expérimentation AS&OP-DSS. La suite présente le détail d'une expérimentation représentative.

4.2 Présentation du cas d'étude

Ce cas théorique est basé sur l'étude un atelier de production en *flow shop* composé de centres de charge WC_r avec $r \in \{1,3\}$, de produits P_{ij} divisés en familles F_i avec $i \in \{1,3\}$ et $j \in \{1,2\}$ ainsi que des composants achetés C_k avec $k \in \{1,3\}$. Une ressource critique est le centre de charge WC2. Chaque produit final P_{ij} et chaque composant C_k est géré individuellement par un buffer de stock physique. La Figure 9 décrit les principales données utilisées qui correspondent aux données techniques du scénario nominal. Le Tableau 2. rassemble les données concernant les risques et opportunités sur notre cas théorique.



Demand		Resource Planning			Finance		Operating Model				
Family	Product	WC1 (min/p)	WC2 (min/p)	WC3 (min/p)	Cube Utilization (m3)	Price (Dollars)	Margin Contribution (Dollars)	DLT	LTF	FV	MOQ
F1	P11	0,3	0,8	0	0,01	30	15	60	0,4	0,5	50
	P12	0,3	0,7	0	0,01	30	15	60	0,4	0,5	50
F2	P21	0,3	1,2	0,1	0,01	30	20	20	0,2	0,75	50
	P22	0,3	1,3	0,1	0,01	30	20	20	0,2	0,75	50
F3	P31	0,3	2	0,5	0,01	100	60	20	0,2	0,8	10
	P32	0,3	2,5	0,5	0,01	100	60	20	0,2	0,8	10

		Cube Utilization (m3)	Price	DLT	FD	FV	MOQ
Supplier 1	C1	0,005	2	5	0,7	0,3	200
	C2	0,005	3	5	0,7	0,3	200
Supplier 2	C3	0,005	20	30	0,3	0,75	200

Capacity	min per day	m3
WC1	480	
WC2	480	
WC3	480	
WH		250

Figure 9. Données techniques du scénario nominal du cas d'étude

Tableau 2. Risques et opportunités de notre cas théorique

	Désignation
Risque	la capacité du centre critique peut chuter de 20% Les prévisions de vente sont connues sous forme d'intervalle par mois.
Opportunité	Un nouveau marché est à l'étude avec une augmentation de 20% sur la famille 1 de volume de vente à partir du mois prochain

4.3 Utilisation de notre suite AS&OP-DSS sur cas théorique

La Figure 10. Présente l'arbre de décision de notre étude de cas avec 184 scénarios à générer issue de la combinatoire de la situation nominale, de l'opportunité et des deux risques.

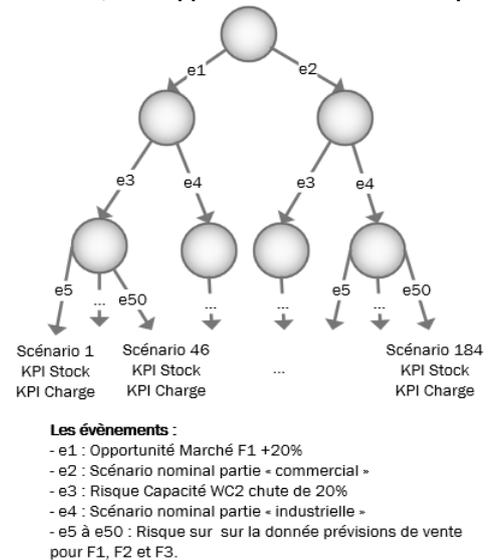


Figure 10. Arbre de décision de notre cas d'étude

Family	Product	Sales volume	Saisonnality	Market	Price	contribution margin	DLT	Variability
F1	P11	High	non	Market 1	low	low	long >1 month	low
	P12	High	yes	Market 1	low	low	long >1 month	low
F2	P21	Low	no	Market 1	low	middle	middle <1 month	middle
	P22	Low	no	Market 1	low	middle	middle <1 month	middle
F3	P31	Sporadic	no	Market 2	high	high	middle <1 month	high
	P32	Sporadic	no	Market 2	high	high	middle <1 month	high

Supplier	Component	DLT	Variability	Buying Cost
Supplier 1	C1	short	low	low
	C2	short	low	low
Supplier 2	C3	long	High	high

Work Center	Critical Resource
WC1	no
WC2	yes
WC3	no
WH	yes

L'exécution du programme permet d'obtenir en moins d'une minute un tableau de bord complet permettant d'observer simultanément les indicateurs intégrés comme le BFR, l'état de charge sur le centre critique ou encore l'évolution de chaque buffer stratégique. La Figure 11. par exemple montre la variation du BFR sur l'horizon de projection AS&OP (24 mois) indiquant les résultats de manière agrégé pour les 184 scénarii suivant l'axe des ordonnées et suivant l'axe des abscisses les différentes variantes de nominal, risque et opportunité du cas d'étude : (i) nominal, (ii) risque et opportunité et (iv) risque et opportunité. Elle montre aux décideurs que la perte de capacité a peu d'impact sur le

BFR. Par contre l'opportunité de marché induit une croissance de 20% du BFR L'incertitude sur la donnée de prévisions de vente engendre une variation de 5% du BFR sur chaque période considérée.

Nous avons ensuite testé d'autres alternatives décisionnelles en évaluant des variantes basées sur le Tableau 1. Toutes ces variantes ont été testées avec succès dans une durée de temps réelle compatible avec la durée moyenne accordée par les praticiens sur ce type de processus décisionnel. L'objectif étant de valider la faisabilité de la démarche, nous n'avons pas cherché à « optimiser » les décisions sur ce cas d'étude.

4.4 Discussions sur expérimentation

Premièrement, l'expérimentation nous a permis de tester la capacité de notre proposition de processus métier AS&OP, supporté par notre proposition d'AS&OP-DSS, à intégrer les recommandations décrites par le DDI : préparation des stratégies, intégration dans les supports informatiques dédiés, intégration de risques et opportunités pour avoir une vision réaliste du champs des possibles, génération des nombreux scénarios associés, visualisation des résultats et remédiation pour tester de nouvelles alternatives décisionnelles pour corriger les éléments nécessaires comme par exemple la limite capacitaire atteinte.

Deuxièmement, l'utilisation d'un arbre de décision structuré a permis de rapidement mettre en forme de manière intelligible le tableau de bord avec les nombreuses données générées en un minimum d'affichage.

Troisièmement, le délai d'exécution du script mesurée à chaque itération a été de 40 à 60 secondes ce qui permet une utilisation possible tant en direct pendant une réunion pour refaire tourner des scénarios/refaire des itérations que pendant la phase de préparation d'une réunion.

Quatrièmement, les décideurs doivent actuellement traiter manuellement les informations provenant du AS&OP-DSS, certes dans un logiciel flexible d'élaboration de tableaux de bords, mais avec de très nombreux scénarios à étudier, ce qui peut être contreproductif. De plus, tester manuellement les décisions pour valider un impact est possible mais reste sujet à la capacité des décideurs à percevoir ces possibilités. Automatiser la proposition de préconisations au niveau de l'étape « traiter et afficher les résultats », ou tout du moins la caractérisation de l'impact de certaines options décisionnelles face aux incertitudes du décideur, est possible en formalisant un arbre de décision. La structure de l'arbre formalise la hiérarchie des décisions que se posent les utilisateurs et les incertitudes qui leur font face. Sur chaque nœud, le ou les critères de décision peuvent être formalisés. Les feuilles constituent un ensemble de scénarios à évaluer. Le traitement des résultats, des feuilles jusqu'à la racine, permet ainsi de formaliser la ou les stratégies dominantes.

Cinquièmement, l'intégration du modèle agrégé proposé par (J. Vidal et al., 2020) a rendu possible notre proposition, cependant ce modèle n'a pas été caractérisé en profondeur en matière de sensibilité, de robustesse ni même de justesse par rapport à un modèle DDMRP non agrégé. Il semble donc important d'établir celle-ci et notre plateforme d'expérimentation peut nous permettre de le réaliser correctement via plusieurs plans d'expériences dédiés.

suivante: *Comment développer un système d'aide à la décision soutenant efficacement le processus AS&OP ?*

Nous avons notamment proposé un processus métier associé à AS&OP plus formel et détaillé que le schéma de principe proposée par les créateurs de la méthode. Nous avons également élaboré une approche structurée pour développer un système d'aide à la décision (DSS) intégrant les originalités du processus AS&OP. Nous avons également implanté avec succès cette approche dans un premier programme informatique opérationnelle. Nos propositions présentent un potentiel certain pour soutenir le processus AS&OP. Cependant, ce travail n'en est qu'à son début et permet d'ores et déjà de dégager plusieurs perspectives de recherche suivant deux volets : (i) processus métier AS&OP et (ii) support logiciel AS&OP-DSS.

Pour le premier volet (i), une expérimentation scénarisée avec des praticiens pourrait apporter un premier recul sur cette proposition. Dans notre revue de littérature nous avons également identifié des propositions de processus DDS&OP et DDMRP, mais aucune associant l'ensemble des niveaux en même temps de manière cohérente. Cette approche plus formelle aiderait également à une comparaison plus fine entre les différents travaux réalisés de modélisation du processus S&OP.

Pour le deuxième volet (ii), notre étude laisse le décideur faire le tri dans la multitude des scénarios manuellement, une amélioration pourrait être l'intégration d'une préconisation automatisée des décisions sur la base de l'exploitation de l'arbre des décisions. Nous pouvons également envisager étudier l'apport à notre AS&OP-DSS de l'utilisation de méthodes de type Aide Multicritères à la Décision (MCDA) en intégrant des critères de préférence sur les risques et opportunités testées. Le modèle agrégé du DDMRP utilisé dans le programme reste à étudier en profondeur en termes de validité, sensibilité et robustesse. La planification scénarisée mise en avant dans notre proposition pourrait enfin être étendue dans le cadre d'un processus S&OP plus traditionnel appartenant à la logique MRP2.

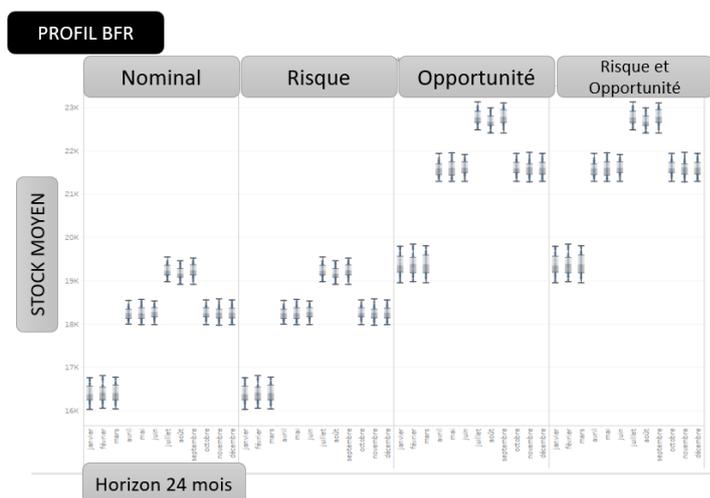


Figure 11. Extrait tableau de bord étude de cas, exemple : indicateur BFR

6 REFERENCES

- Acosta, A. P. V., Mascle, C., & Baptiste, P. (2019). Applicability of Demand-Driven MRP in a complex manufacturing environment. *International Journal of Production Research*, 0(0), 1-13. <https://doi.org/10.1080/00207543.2019.1650978>

- Bahu, B., Bironneau, L., & Hovelaque, V. (2019). Compréhension du DDMRP et de son adoption : Premiers éléments empiriques. *Logistique & Management*, 27(1), 20-32. <https://doi.org/10.1080/12507970.2018.1547130>
- Bayard, S., Grimaud, F., & Delorme, X. (2020). *Présentation des Etudes de cas de l'étude ESCAPADD* [Report]. Mines Saint-Etienne. <https://hal-emse.ccsd.cnrs.fr/emse-02884848>
- Bozutti, D. F., & Esposto, K. F. (2019). Sales and Operations Planning : A comparison between the demand-driven and traditional approaches. *International Journal of Production Management and Engineering*, 7(1), 23-38. <https://polipapers.upv.es/index.php/IJPM/article/view/9469>
- DDI. ((s. d.)). *The Demand Driven Institute*. Ddmain. <https://www.demanddriveninstitute.com>
- Demand Driven Institut. ((s.d.)). *Demand Driven Adaptive Enterprise (DDAE)*. The Demand Driven Institute - World Leader in Demand Driven Education. <https://www.demanddriveninstitute.com/demand-driven-adaptive-enterprise-m>
- Dessevre, G., Lamothe, J., Pomponne, V., Baptiste, P., Lauras, M., & Pellerin, R. (2020). A DDMRP implementation user feedbacks and stakes analysis. *ILS 2020 - 8th International Conference on Information Systems, Logistics and Supply Chain*, 204-211. <https://hal-mines-albi.archives-ouvertes.fr/hal-02940566>
- Ducrot, L., & Ahmed, E. (2019). *Investigation of potential added value of DDMRP in planning under uncertainty at finite capacity*. <https://dspace.mit.edu/handle/1721.1/121296>
- Kristensen, J., & Jonsson, P. (2018). Context-based sales and operations planning (S&OP) research : A literature review and future agenda. *International Journal of Physical Distribution & Logistics Management*, 48(1), 19-46. <https://doi.org/10.1108/IJPDLM-11-2017-0352>
- Lindgren, M., & Bandhold, H. (2002). *Scenario Planning : The Link Between Future and Strategy*. Springer.
- Martin, G., Baptiste, P., Lamothe, J., Miclo, R., & Lauras, M. (2018). *A process map for the demand driven adaptive enterprise model : Towards an explicit cartography*. <https://hal-mines-albi.archives-ouvertes.fr/hal-01883504>
- Meinzel, L. (2019). *DDMRP : Presentation of a new solution of stock management and master production scheduling*. <https://recercat.cat/handle/2072/360713>
- Miclo, R. (2016). *Challenging the « Demand Driven MRP » Promises : A Discrete Event Simulation Approach* [Ecole nationale des Mines d'Albi-Carmaux]. <http://www.theses.fr/2016EMAC0016>
- Miclo, R., Lauras, M., Fontanili, F., Lamothe, J., & Melnyk, S. A. (2019). Demand Driven MRP : Assessment of a new approach to materials management. *International Journal of Production Research*, 57(1), 166-181. <https://doi.org/10.1080/00207543.2018.1464230>
- Nabil, L., El Barkany, A., & El Khalfi, A. (2018). *Sales and Operations Planning (S&OP) Concepts and Models under Constraints : Literature Review*. International Journal of Engineering Research in Africa. <https://doi.org/10.4028/www.scientific.net/JERA.34.1>
- Oger, R. (2019). *A decision support system for long-term supply chain capacity planning : A model-driven engineering approach* [These de doctorat, Ecole nationale des Mines d'Albi-Carmaux]. <http://www.theses.fr/2019EMAC0013>
- Oger, R., Lauras, M., Montreuil, B., & Benaben, F. (2020). A decision support system for strategic supply chain capacity planning under uncertainty : Conceptual framework and experiment. *Enterprise Information Systems*, 0(0), 1-45. <https://doi.org/10.1080/17517575.2020.1793390>
- Pereira, D. F., Oliveira, J. F., & Carravilla, M. A. (2020). Tactical sales and operations planning : A holistic framework and a literature review of decision-making models. *International Journal of Production Economics*, 228, 107695. <https://doi.org/10.1016/j.ijpe.2020.107695>
- Pittman, P., & Atwater, J. B. (2016). *APICS Dictionary, 15th Edition*. APICS.
- Ptak, C. A., & Smith, C. J. (2011). *Orlicky's material requirements planning*. McGraw Hill. http://biblioteca.uazuay.edu.ec/opac_css/index.php?lvl=notice_display&id=78328
- Ptak, C., & Smith, C. (2019). *Demand Driven Material Requirements Planning (DDMRP), Version 3*. Industrial Press, Incorporated.
- Slack, N., & Lewis, M. (2002). *Operations Strategy*. Pearson Education.
- Smith, D., & Smith, C. (2014). *Demand Driven Performance*. McGraw Hill Professional.
- Tableau Software. ((s.d.)). *Tableau : Business Intelligence and Analytics Software*. Tableau. <https://www.tableau.com/home>
- Tavares Thomé, A. M., Scavarda, L. F., Fernandez, N. S., & Scavarda, A. J. (2012). Sales and operations planning : A research synthesis. *International Journal of Production Economics*, 138(1), 1-13. <https://doi.org/10.1016/j.ijpe.2011.11.027>
- Tiberius, V., Siglow, C., & Sendra-García, J. (2020). Scenarios in business and management : The current stock and research opportunities. *Journal of Business Research*, 121, 235-242. <https://doi.org/10.1016/j.jbusres.2020.08.037>
- Tuomikangas, N., & Kaipia, R. (2014). A coordination framework for sales and operations planning (S&OP) : Synthesis from the literature. *International Journal of Production Economics*, 154, 243-262. <https://doi.org/10.1016/j.ijpe.2014.04.026>
- Vidal, J., Lauras, M., Lamothe, J., & Miclo, R. (2020). Toward an Aggregate Approach for Supporting Adaptive Sales And Operations Planning. *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)*, 1031-1038. <https://doi.org/10.1109/ICIEA49774.2020.9102060>
- Vidal, J.-B., Lauras, M., Lamothe, J., & Miclo, R. (2018). Adaptive Sales & Operations Planning : Innovative Concept for Manufacturing Collaborative Decisions? In L. M. Camarinha-Matos, H. Afsarmanesh, & Y. Rezgui (Éds.), *Collaborative Networks of Cognitive Systems* (p. 362-374). Springer International Publishing. https://doi.org/10.1007/978-3-319-99127-6_31

Choix du niveau d'agrégation d'un processus S&OP dans un contexte de forte variété

YAHYA GHRAB^{1,2}, MUSTAPHA SALI², CLEMENT CHATRAS¹

¹ Renault Sas
Vélizy, FRANCE

yahya.ghrab@dauphine.eu
clement.chatras@renault.com

² Université Paris Dauphine
Paris, FRANCE
mustapha.Sali@dauphine.psl.eu

Résumé – Le Sales and Operations Planning (S&OP) est un processus de planification tactique multi-acteurs dont l'objectif est d'aligner, sur un horizon à moyen terme et à un niveau agrégé correspondant à des familles de produits, une demande commerciale et des capacités industrielles et logistiques. Dans un contexte de grande variété comme l'assemblage automobile, la notion de niveau d'agrégation correspond à un sous-ensemble de caractéristiques fonctionnelles dont les différents attributs sont utilisés pour décrire des groupes de produits homogènes. Au concept de famille de produit se substitue celui de niveau d'agrégation. Compte tenu de la diversité, il existe un grand nombre de niveaux d'agrégation possibles plus ou moins adaptés aux besoins de chaque acteur impliqué dans le S&OP. Afin d'identifier des niveaux d'agrégation équilibrés, un modèle d'optimisation multi-objectifs basé sur les critères de chaque acteur impliqué dans le S&OP est proposé. Les résultats d'une étude de cas réalisée à partir de données réelles fournies par un grand constructeur automobile sont présentés. Par rapport à la situation actuelle, les résultats mettent en évidence l'existence de niveaux d'agrégation alternatifs équilibrés offrant de meilleures garanties quant à l'alignement de la demande et des capacités.

Mots clés - S&OP, Personnalisation de masse, Niveau d'agrégation, Famille de produits, Alignement inter-fonctionnel, Optimisation multi-objectifs

Abstract – Sales and Operations Planning (S&OP) is a multi-stakeholder tactical planning process whose objective is to align, over a medium-term horizon and at an aggregated level corresponding to product families, sales demand and industrial and logistics capacities. In a context of high variety such as automotive assembly, the notion of aggregation level corresponds to a subset of functional characteristics whose different attributes are used to describe homogeneous product groups. The concept of aggregation level replaces the concept of product family. Given the diversity, there is a large number of aggregation levels more or less adapted to the needs of each stakeholder involved in the process. In order to identify balanced levels of aggregation, a multi-objective optimization model based on the criteria of each stakeholder involved in the S&OP is proposed. The results of a case study based on real data provided by a major car manufacturer are presented. Compared to the current situation, the results highlight the existence of alternative aggregation levels offering better guarantees in terms of demand and capacity alignment.

Keywords - S&OP, Mass customization, Aggregation level, Product family, Cross-functional alignment, Multi-objective optimization

1 INTRODUCTION

Les industriels impliqués dans certains secteurs d'activité sont confrontés à de nouvelles habitudes de consommation caractérisées par une quête de personnalisation et de délais de livraison de plus en plus courts [Marino et al., 2018]. Cette tendance se traduit par une augmentation de la diversité produit, qui modifie les logiques de pilotage des flux et de gestion des opérations [Um et al., 2017]. Beaucoup d'efforts ont été réalisés pour concilier volumes importants, grande variété, fournisseurs éloignés et délais de livraison courts dans des conditions économiques viables [Laurent Lim et al., 2014]. La modularité, la standardisation et la différenciation retardée [EIMaraghy et al., 2013] sont des approches largement diffusées dans l'industrie et particulièrement dans le secteur automobile. Malgré ces efforts, la question de la gestion efficace de la diversité reste d'actualité pour les constructeurs automobiles [Shao, 2020]. Du point de vue de la coordination de la chaîne logistique, l'une des principales difficultés posées par la diversité est la mise en place de processus de planification robustes qui permettent de concilier demande commerciale et

capacités industrielles à moyen terme. Cette conciliation suppose la possibilité de confronter deux logiques d'agrégation de produits correspondant chacune aux objectifs contradictoires des acteurs impliqués dans le processus de planification.

Dans le secteur automobile, le niveau d'agrégation de la demande correspond aux combinaisons d'attributs associés à un sous-ensemble de caractéristiques propres à un véhicule comme le moteur, la carrosserie et la boîte de vitesses par exemple [Chatras et al., 2016]. Par attributs, nous faisons référence aux modalités possibles d'une caractéristique. La caractéristique « Moteur » se décline par exemple en autant d'attributs qu'il y a de modalités de moteurs (Moteur 1, Moteur 2, ...). Pour simplifier, nous considérerons qu'un niveau d'agrégation est un sous-ensemble de caractéristiques.

La demande prévisionnelle est exprimée par les services commerciaux à un niveau d'agrégation prédéfini construit généralement à partir de caractéristiques perçues par le client. Elle est ensuite présentée aux équipes industrielles dont la mission est de vérifier la disponibilité des capacités de production et d'approvisionnement pour répondre à la demande.

Avec le niveau d'agrégation donné dans l'exemple précédent, cela signifie vérifier la capacité de produire les volumes demandés pour toutes les combinaisons d'attributs relatifs aux caractéristiques moteur, carrosserie et boîte de vitesse. La même logique de vérification capacitaire s'applique à l'approvisionnement des pièces nécessaires à la fabrication. Cependant, certaines contraintes industrielles critiques peuvent concerner des caractéristiques spécifiques qui ne sont pas représentées dans le niveau d'agrégation utilisé pour exprimer la demande. Par exemple, supposons qu'il existe une contrainte qui limite le volume de fabrication quotidien de véhicules équipés d'un toit ouvrant. Il est évidemment impossible de vérifier avec précision la compatibilité d'une telle contrainte étant donné la forme actuelle de la demande agrégée qui ne fournit pas de données prévisionnelles concernant le nombre de véhicules équipés de toit ouvrant. Pour vérifier que cette contrainte est compatible avec la demande commerciale, les équipes industrielles doivent effectuer une opération de désagrégation des données initiales, en se basant uniquement sur des taux historiques, pour inférer le nombre de véhicules dotés d'un toit ouvrant. L'utilisation de taux historiques sur des marchés volatils est très risquée et peut conduire à une sous-estimation ou une surestimation de la demande réelle. Ce risque est d'autant plus élevé que les équipes industrielles ne perçoivent pas directement les dynamiques de marché propres à certaines caractéristiques des produits. Deux cas peuvent alors se présenter après une opération de désagrégation qui valide théoriquement la compatibilité de la demande et des capacités. Dans le premier cas, la capacité peut s'avérer insuffisante pour répondre à la demande réelle et des problèmes à court terme peuvent survenir avec des conséquences sur les délais et les coûts de production [Laurent Lim et al., 2014]. Dans le second cas, la demande réelle peut-être très inférieure à une demande prévisionnelle ajustée à la baisse pour la rendre compatible avec la contrainte industrielle. Comme les caractéristiques sont dépendantes les unes des autres et que le volume global de production est stable, le risque est de générer des perturbations sur d'autres caractéristiques. Dans le premier cas, la sous-estimation de la demande donne la fausse impression de contraintes non bloquantes alors que dans le second, la surestimation de la demande donne la fausse impression de contraintes bloquantes qui déclenchent des ajustements inutiles susceptibles de rendre d'autres contraintes bloquantes.

Afin d'améliorer l'alignement des acteurs impliqués dans le processus S&OP et éviter les effets pervers des mécanismes de désagrégation basés sur de taux historiques, ce document propose d'identifier les sous-ensembles de caractéristiques qui doivent être utilisées pour exprimer la demande et faciliter le dialogue entre commerciaux d'un côté et industriels de l'autre. L'idée consiste à identifier les caractéristiques liées aux contraintes industrielles et à en tenir compte pour exprimer la demande prévisionnelle. Cela garantit que les contraintes de capacité sont prises en compte à un stade précoce du processus de planification et permet d'éviter les distorsions qu'impliquent les extrapolations basées sur l'historique.

Pour que cette proposition soit réaliste, le niveau d'agrégation doit intégrer le moins de caractéristiques possible. Il est en effet impensable que les équipes commerciales expriment une demande ventilée sur toutes les combinaisons d'attributs de caractéristiques concernées par l'ensemble des contraintes industrielles. Étant donné que la diversité des produits résulte d'une combinaison de nombreuses caractéristiques qui peuvent ou non être concernées par une contrainte industrielle, l'identification d'un niveau d'agrégation qui tienne compte des contraintes capacitaires tout en restant adapté à l'expression de

la demande à moyen terme est un problème combinatoire multi-objectifs. À notre connaissance, aucune étude antérieure n'a examiné en profondeur l'importance et les avantages de la prise en compte des contraintes industrielles pour définir le niveau d'agrégation de planification dans le contexte de la production fortement diversifiée.

Le présent document contribue à combler cette lacune en identifiant les propriétés d'un bon niveau d'agrégation et en proposant un modèle analytique qui vise à identifier des niveaux d'agrégation efficaces au regard de critères industriels (liés à la prise en compte des contraintes dans le niveau d'agrégation) et commerciaux (liés à la pertinence des caractéristiques du niveau d'agrégation pour l'expression de la demande). Le problème multi-objectif est résolu en utilisant une approche lexicographique où chaque fonction objective fait référence à un des deux critères mentionnés précédemment.

Une analyse numérique multi scénario a été menée sur des données réelles afin d'illustrer la démarche et fournir des éléments de prise de décision pour le partenaire de l'étude.

Le reste du document est organisé comme suit : la deuxième section est consacrée à une revue de littérature concernant les principaux concepts utilisés pour mener cette étude, plus particulièrement, ceux en lien avec le processus S&OP en contexte de forte diversité. La troisième section décrit le contexte industriel de l'étude en apportant des précisions concernant les logiques de représentation des produits dans cet environnement. Cette section introduit également les propriétés de ce qui fait un bon niveau d'agrégation pour la planification. La quatrième section présente le modèle d'optimisation multi-objectif et propose une traduction formelle des propriétés d'un bon niveau d'agrégation identifiées dans la section précédente. La cinquième section présente les résultats d'une étude de cas réalisée sur des données réelles. En conclusion nous ouvrons quelques perspectives pour de futures travaux.

2 LITTÉRATURE

La planification est un processus essentiel pour les chaînes logistiques largement évoquée dans la littérature [Lalami et al. 2017; Olhager 2003; Staebelin and Aoki 2015, etc.]. La planification peut être définie comme la conséquence d'une prise de décision hiérarchique dont le but est d'anticiper les besoins liés à une activité future envisagée [Gelders et al., 1982; Özdamar et al., 1998]. La planification consiste à définir des plans de production et d'utilisation de ressources (d'approvisionnement, de production et de distribution) par anticipation d'une demande prévisionnelle sur différents horizons temporels et à exprimer selon différentes mailles (du plus global à un niveau de planification stratégique au plus spécifique à un niveau d'ordonnancement quotidien en passant par une granularité intermédiaire à un niveau tactique) [Fontan et al., 2005; Gelders et al., 1982; Pereira et al., 2020]. Le niveau stratégique définit la ligne directrice à long terme en tenant compte des tendances du marché. Le dimensionnement de l'équipement industriel, les choix organisationnels ainsi que les investissements lourds sont généralement décidés à ce niveau. Le niveau tactique consiste à établir des plans agrégés qui résultent d'arbitrage entre la charge estimée et les capacités disponibles à moyen terme. Le S&OP intervient précisément à ce niveau. Ce processus de décision tactique multi-acteurs consiste à définir une feuille de route commerciale et industrielle en conciliant demande prévisionnelle d'un côté et capacités de production, d'approvisionnement de de transport de l'autre, tout en considérant les objectifs financiers qui traduisent les orientations stratégiques [Antonio Márcio Tavares Thomé et al., 2012].

Les plans établis selon une granularité qui correspond à des familles de produits sont généralement mis à jour mensuellement sur un horizon de planification tactique de plusieurs mois (de 6 à 18 mois en pratique) [Blackstone et al., 2005; Grimson et al., 2007; Antônio Márcio Tavares Thomé et al., 2012].

Le dernier niveau de planification est le niveau opérationnel qui s'inscrit dans le court terme. La demande et les capacités y sont exprimées selon des niveaux de granularité très fins, correspondant le plus souvent à l'expression complète du produit fini [Pereira et al., 2020].

La décomposition hiérarchique des processus de planification [Hax et al., 1973] est un sujet largement discuté depuis la fin des années 70 et a contribué à façonner la plupart des systèmes de planification actuels. Dans ce papier, nous nous intéressons au S&OP qui joue un rôle central dans la coordination des chaînes logistiques en tant que chaînon entre les niveaux de planification stratégique et tactique [Tuomikangas et al., 2014]. [Goh et al., 2019; Antônio Márcio Tavares Thomé et al., 2014] soulignent l'importance de la coordination pour améliorer l'efficacité globale de la chaîne logistique.

Dans une revue de littérature récente, [Pereira et al., 2020] examinent de nombreuses études visant à optimiser le S&OP via des approches analytiques qui cherchent à réduire le coût de production global. Les modèles recensés s'appuient majoritairement sur des programmes linéaires, des programmes multi-objectifs ou des approches floues. Ces modèles prennent en considération des paramètres liés à la production, à l'approvisionnement, à la distribution et aux ventes. Notre contribution fait appel à des outils analytiques similaires mais se différencie de ce qui est proposé dans la littérature par son contexte d'application dans la production de masse fortement diversifiée où les données manipulées dans le S&OP prennent des formes très particulières (forte variété de produits et de contraintes industrielles et nomenclature modulaire).

D'autre part, le succès d'une démarche S&OP dépend beaucoup du degré d'alignement entre les différents acteurs qui ont souvent des objectifs contradictoires [Feng et al., 2008; Gelders et al., 1982; Antônio Márcio Tavares Thomé et al., 2012]. De nombreux auteurs ont souligné l'importance de l'alignement dans la planification tactique de la chaîne logistique [Goh et al., 2019; Oliva et al., 2011; Tuomikangas et al., 2014; Wagner et al., 2014]. D'autres contributions, qui s'intéressent à l'évaluation de la maturité du S&OP, révèlent à leur manière l'importance de l'alignement, considéré un critère de performance essentiel pour le S&OP [Ghrab et al., 2020; Grimson et al., 2007].

L'alignement et la coordination dans un contexte de forte diversité implique des difficultés supplémentaires évoqués dans la littérature [Blecker et al., 2006; ElMaraghy et al., 2013]. Une grande variété de produits finis pèse sur l'efficacité de la planification [Kashkoush et al., 2016; Um et al., 2017] en complexifiant le processus de prévision [Ramdas, 2003] et en générant des coûts opérationnels importants [Stäblein et al., 2011]. Plusieurs solutions ont été proposées en matière de conception, d'approvisionnement et d'ingénierie des flux pour limiter l'impact de la diversité. La standardisation et la conception modulaire de produits ont très largement été étudiées dans la littérature. [Chatras et al., 2016; Daaboul et al., 2011; Tang et al., 2017] mettent en avant l'intérêt de ces approches pour limiter la diversité des pièces, réduire l'effort de conception et limiter les coûts de production

La littérature concernant les règles d'agrégation de la demande en phase de planification tactique dans un contexte de forte diversité reste très limitée. Le sujet a été discuté de manière marginale en tant que sous-problème de la planification

hiérarchique [Gelders et al., 1982] sans que le contexte de la forte diversité ne soit évoqué de manière explicite. Seuls quelques principes généraux sont évoqués concernant les types d'agrégation envisageables pour regrouper des demandes [Gelders et al., 1982; Herrmann et al., 1994].

Compte tenu de la spécificité du contexte de la production fortement diversifiée, ce document s'intéresse à la question de l'agrégation de la demande en famille de produit. Certaines règles empiriques pour l'agrégation des produits ont été fournies dans la littérature. Elles comprennent la constitution de familles basées sur des coûts d'installation similaires [Hax et al., 1973], des saisonnalité et/ou des tendances comparables [Liberatore et al., 1985] ou encore sur l'utilisation de ressources industrielles communes [Herrmann et al., 1994; Kanyalkar et al., 2005]. Une recherche plus récente s'est intéressée aux modèles analytiques permettant de constituer des familles de produits en analysant les similitudes. Dans leur proposition, [Kashkoush et al., 2016] se sont intéressés aux paramètres opérationnels liés, entre autres, aux machines, aux processus, aux gammes de production et à la structure de la nomenclature (BOM) pour proposer des logiques d'agrégation dans un contexte de forte diversité.

Notre proposition se différencie des contributions existantes pour deux raisons principales. Premièrement, en raison de la nature combinatoire de la diversité considérée dans notre cas, il n'est pas possible de s'appuyer sur les représentations arborescentes classiques de la nomenclature pour constituer des familles de produit sur la base d'un simple critère de similarité. Un système particulier de représentation des produits change radicalement les perspectives et critères de regroupement en famille homogènes. Deuxièmement, alors que les contributions existantes n'abordent la question de l'agrégation que d'un seul point de vue (commercial ou industriel), nous proposons une approche qui combine les deux perspectives dans le but de garantir un meilleur alignement entre les acteurs du S&OP.

À notre connaissance, il n'existe pas de modèle analytique permettant d'identifier le meilleur niveau d'agrégation dans un contexte de production de masse fortement diversifiée. Le présent document vise à combler cette lacune en fournissant un modèle d'aide à la décision qui permet d'identifier des niveaux d'agrégation compatibles avec les objectifs des acteurs impliqués dans le processus de planification tactique.

3 CADRE DE L'ETUDE

Comme ce document se concentre sur le contexte de la production de masse fortement diversifiée dans le secteur automobile, il semble important d'introduire et de décrire certains concepts clés utiles à la compréhension de la problématique et de son environnement. C'est l'objet de la première partie de cette section. La deuxième partie identifie les enjeux liés au choix du niveau d'agrégation. Cette étape est essentielle à la compréhension de l'approche de modélisation et de résolution proposée en section 4.

3.1 Contexte

3.1.1 Variété de produits

Le contexte de la forte diversité dans le secteur automobile est la trame de fond de notre contribution. Poussés par le désir des clients de disposer de produits personnalisés, les catalogues automobiles ont explosé. Si on prend l'exemple de la Clio (une voiture à hayon populaire commercialisée par Renault), une comparaison des données récentes du tableau 3-1 et des données fournies par Pil & Holweg (2004) montre que le nombre total de variantes est passé de 10^5 à 10^{30} . Cette diversité résulte des nombreuses combinaisons possibles qu'un client peut réaliser et de la grande variété des spécifications techniques spécifiques à

chaque pays de commercialisation (normes anti-pollution, sécurité, etc.). Bien que limitée par certaines contraintes techniques, la diversité qui résulte des choix possibles au niveau des options et équipements proposés peut être qualifiée de combinatoire. Le tableau 1 donne les données récentes concernant la taille du catalogue pour quelques modèles de véhicule du constructeur Renault.

Tableau 1: Nombre de variants pour différents modèles

Modèle	Nombre de variants
Twingo	10^{16}
Trafic	10^{19}
Clio 4	10^{30}
Master	10^{47}

Pour donner au lecteur une idée de l'ampleur de cette variété, prenons l'exemple du modèle Trafic, qui est un véhicule utilitaire léger (VUL). Le tableau 3.1 indique que le catalogue comporte 10^{19} produits finis disponibles à la vente. Si l'on considère un volume moyen de 330 000 véhicules utilitaires légers vendus par an et un cycle de vie moyen de trois ans pour la commercialisation, la probabilité de produire un véhicule spécifique parmi cette diversité offerte est d'environ 10^{-13} . En outre, 10^9 To d'espace de stockage seraient nécessaires pour stocker sur un disque dur la définition complète du produit pour chaque variante. Plutôt que de stocker en mémoire des descriptions complètes correspondant chacune à un des 10^{19} véhicules commercialisables, la représentation des produits se fait selon une logique de description fonctionnelle dont il est question dans la sous-section suivante.

3.1.2 Description fonctionnelle du produit

Les constructeurs automobiles ont établi une représentation fonctionnelle qui décrit un véhicule comme une combinaison d'attributs associés chacun à une caractéristique générique. Ainsi, au lieu de stocker l'ensemble du catalogue de produits, seules les caractéristiques et leurs attributs correspondants sont stockés dans les systèmes d'information. En moyenne, un véhicule peut être défini à travers environ 200 caractéristiques. Notez que le nombre de caractéristiques augmente avec la complexité du produit.

Théoriquement, une caractéristique fait généralement référence à une fonction. Par exemple, certaines caractéristiques courantes sont le moteur, le toit, les roues, la couleur et le niveau d'équipement. Chaque caractéristique j possède un ensemble n_j d'attributs alternatifs. Par exemple, le blanc, le rouge et le bleu sont des attributs possibles pour la caractéristique couleur. Pour définir un produit spécifique, chaque caractéristique doit être renseignée à travers un seul et unique attribut. Par exemple, une voiture ne peut jamais avoir deux attributs pour la caractéristique moteur. Par conséquent, un produit fini est identifié par l'union d'attributs provenant chacun d'une caractéristique différente. La diversité des produits commercialisable correspond donc au produit du nombre d'attributs pour chaque caractéristique auquel on soustrait les combinaisons non autorisées [Chatras et al., 2016].

Un élément important pour l'analyse réside dans les contraintes qui peuvent exister entre les attributs. En effet, pour des raisons de compatibilité, l'attribut associé à une caractéristique peut impliquer automatiquement la présence d'autres attributs associés à d'autres caractéristiques pour un même véhicule. Par exemple, admettons que pour chacun des attributs de la caractéristique moteur on ne puisse utiliser qu'une seule et unique boîte de vitesse. Dans ce cas-là, le choix d'un attribut de la caractéristique moteur implique le choix d'un attribut de la caractéristique boîte de vitesse. Pour simplifier, on dira que la caractéristique moteur induit la caractéristique boîte de vitesse. Dans une logique de planification, notre exemple montre qu'il

suffit de connaître le moteur pour déduire la boîte de vitesse. Donc la présence de la caractéristique moteur dans un niveau d'agrégation donne des indications concernant les contraintes qui concernent la caractéristique boîte de vitesse. Ainsi, la prise en compte de ces liens permet de réduire drastiquement la quantité d'information à traiter en phase de planification pour l'expression de la demande et la prise en compte des contraintes industrielles.

Dans une telle structure, un niveau d'agrégation peut se définir comme un sous-ensemble de caractéristiques. Les combinaisons valides d'attributs associés à ces caractéristiques pointent sur une population de produit partageant les mêmes propriétés. Nous retrouvons ici la notion de famille de produit. Un niveau d'agrégation peut n'être constitué que d'une seule caractéristique comme le modèle de véhicule par exemple. Dans ce cas-là, les populations de produit ne distinguent que par les modèles auxquelles elles font référence. Un niveau plus détaillé peut être constitué de plusieurs caractéristiques et isoler ainsi des populations plus ciblées (par exemple des populations qui partagent les mêmes attributs des caractéristiques moteur et niveau d'équipement). Le cas extrême d'un niveau d'agrégation constitué de l'ensemble des caractéristiques permet d'identifier de manière univoque chaque variant produit du catalogue.

De la même manière, les attributs d'une caractéristique j , peuvent être utilisés pour décrire les contraintes capacitaires. En effet, une contrainte industrielle est définie par la combinaison des attributs qui caractérisent la population de véhicules concernés par une telle contrainte. Par exemple, supposons qu'une ligne d'assemblage ne fabrique que des véhicules du modèle X, qui peuvent avoir l'un des trois moteurs : M1, M2 ou M3. Notez ici que M1, M2 et M3 sont des attributs de la caractéristique Moteur, et que X est un attribut de la caractéristique Modèle. Supposons que, pour une raison technique, la capacité de production journalière de la ligne soit limitée à 100 véhicules équipés de M1 ou M2. Selon le système de description fonctionnelle du produit, cette contrainte, que nous nommons C, s'écrirait comme suit:

$$[C : X \text{ ET } (M1 \text{ OU } M2) \leq 100]$$

C'est-à-dire, on ne peut pas assembler plus de 100 véhicules présentant la combinaison d'attributs X et M1 ou X et M2. Cette expression logique utilisée pour décrire le terme gauche est appelée "expression d'une contrainte industrielle". Dans un cas réel, cette expression peut impliquer un grand nombre d'attributs associés à plusieurs caractéristiques différentes.

3.1.3 Description du processus de planification

La configuration actuelle du S&OP chez le partenaire industriel de cette étude comprend quatre étapes principales : examen de la demande commerciale, examen de la réponse industrielle, réunion d'alignement et réunion d'arbitrage. On peut se référer à [Ghrab et al., 2020] pour plus de détails concernant ces étapes pour l'entreprise étudiée.

La demande traitée dans le S&OP est exprimée à un niveau d'agrégation prédéfini. Les équipes industrielles procèdent ensuite à un examen de cette demande pour vérifier sa conformité aux contraintes capacitaires et fournir une réponse industrielle. Pour ce faire, une opération de désagrégation partielle est nécessaire car, le plus souvent, les caractéristiques du niveau d'agrégation utilisées pour exprimer la demande commerciale ne suffisent pas à contrôler les contraintes. La désagrégation consiste à calculer les volumes pour certaines combinaisons d'attributs associés aux caractéristiques représentées dans l'expression des contraintes industrielles et absentes du niveau d'agrégation initial. Cette étape est réalisée sur la base de taux historiques qui se substituent à l'information

commerciale manquante. Pour bien comprendre l'utilité de ce mécanisme, supposons que la demande prévue soit exprimée en utilisant uniquement la caractéristique moteur. S'il existe une contrainte industrielle critique qui concerne un des attributs de la caractéristique radio, les équipes industrielles doivent connaître les volumes prévus non seulement pour chaque moteur mais aussi pour l'attribut contraint de la caractéristique radio. Par conséquent, la demande initiale doit être exprimée sur des combinaisons d'attributs associés aux caractéristiques moteur et radio. Cette décomposition se fait grâce à des taux historiques, relatifs à chaque combinaison, appliqués au volume global. La demande décomposée peut ainsi être réajustée en tenant compte des contraintes avant d'être à nouveau consolidée selon le niveau d'agrégation initial pour formuler une réponse aux équipes commerciales à la fin de la phase d'examen de la réponse industrielle.

Les étapes de désagrégation sont très chronophages pour les équipes industrielles et introduisent un biais important car elles reposent sur des taux historiques qui ne reflètent pas les dynamiques d'un marché connu pour sa grande volatilité. Par conséquent, la définition d'un niveau d'agrégation approprié est cruciale pour limiter les effets de bord induits par les opérations de désagrégation. En effet, plus les caractéristiques liées à des contraintes industrielles sont prise en compte dans le niveau d'agrégation de la demande commerciale, moins il est nécessaire d'opérer des calculs basés sur des taux historiques. Poussé à l'extrême, cette logique peut aboutir à un niveau d'agrégation riche en caractéristiques et à une demande exprimée à un degré de finesse incompatible avec la nature tactique du processus S&OP. Un arbitrage doit donc être fait.

3.2 Propriétés d'un bon niveau d'agrégation de la planification

Sur la base des observations faites sur le terrain et de la description du contexte proposée en section 3.1, une série de propriétés qui serviront à qualifier ce qui est un "bon niveau d'agrégation" a été identifiée. Cette section présente ces propriétés souhaitables dont la traduction quantitative servira de base à la construction du modèle présenté dans la section 4.

Dans l'entreprise étudiée, le choix actuel du niveau d'agrégation relève d'un consensus empirique, dont les origines sont difficiles à dater, entre les équipes commerciales et industrielles. Il est intéressant de noter que le consensus s'est concentré sur l'utilisation de caractéristiques commerciales liées aux fonctionnalités visibles ou choisies par le client. Il s'agissait entre autres, du moteur, du niveau d'équipement et de la boîte de vitesses. Si la présence de ces caractéristiques dans le niveau d'agrégation actuel résulte d'un choix empirique, elle ne peut être totalement remise en question pour deux raisons principales. Premièrement, ces caractéristiques servent de référence au processus de prévision et au dialogue avec le réseau commercial. Deuxièmement, ces caractéristiques font référence à des fonctionnalités majeures d'un véhicule et constituent donc un vocabulaire commun entre les clients et les acteurs du réseau commercial. Par conséquent, maximiser la présence de ces caractéristiques dans le niveau d'agrégation est un gage de stabilité organisationnelle et de significativité commerciale de la demande. Nous utiliserons le concept de "caractéristiques commerciales" pour désigner les caractéristiques constituant le niveau d'agrégation initial. Sur la base des explications ci-dessus, on peut déduire une première propriété triviale d'un bon niveau d'agrégation : *un bon niveau d'agrégation doit être significatif d'un point de vue commercial en intégrant un maximum de caractéristiques commerciales*. Cette propriété se traduit par l'objectif de minimiser la sélection des caractéristiques non-commerce.

D'un autre côté, les contraintes industrielles critiques concernent des caractéristiques qualifiées de "caractéristiques industrielles". Comme expliqué précédemment, les équipes industrielles sont amenées à décomposer la demande agrégée en utilisant des taux historiques pour identifier, et éventuellement ajuster, les volumes associés à certaines contraintes. Cette opération, basée sur des taux historiques, conduit à l'apparition de biais importants lorsqu'il s'agit d'ajuster la demande commerciale pour la rendre compatible avec les contraintes. Deux situations sont à distinguer. Premièrement, les volumes calculés à partir des taux historiques peuvent être surestimés, et les contraintes industrielle correspondantes peuvent donc sembler violées. Cette situation peut entraîner une distorsion inutile de demande initiale pour respecter les contraintes et/ou des investissements inutiles pour augmenter la capacité. Deuxièmement, les volumes calculés peuvent être sous-estimés. Dans ce cas, alors que la demande prévue semble être conforme aux contraintes capacitaires, des pénuries peuvent apparaître lorsque la demande réelle est connue. Des livraisons d'urgence coûteuses peuvent être déclenchées pour éviter les ruptures ou limiter les retards de livraison. Par conséquent, la prise en compte des "caractéristiques industrielles" liées aux contraintes permet d'éviter les biais induits par les opérations de décomposition de la demande et donc d'assurer un meilleur alignement entre les deux principaux acteurs du S&OP. La deuxième propriété d'un bon niveau d'agrégation peut ainsi être déduite de cette courte analyse : *un bon niveau d'agrégation doit être significatif d'un point de vue industriel en intégrant un maximum de caractéristiques industrielles*. Cette propriété se traduit par l'objectif de maximiser la couverture des contraintes capacitaires.

Un autre aspect important qui caractérise un niveau d'agrégation est la diversité qu'il engendre et donc la charge de travail nécessaire à la préparation et à l'analyse des prévisions de la demande ainsi qu'aux arbitrages en cours de processus S&OP. Par ailleurs, plus la diversité est importante, moins les prévisions sont fiables [Wan et al., 2017]. Plus le niveau d'agrégation est riche en caractéristiques présentant un grand nombre d'attributs avec un faible niveau d'incompatibilité, plus la diversité associée à un niveau d'agrégation sera grande. La compacité du niveau d'agrégation est la garantie d'une meilleure qualité de prévisions, d'une charge de travail maîtrisée pour les prévisionnistes et d'un processus S&OP plus fluide. Par conséquent, une troisième propriété d'un bon niveau d'agrégation peut se formuler comme suit : *un bon niveau d'agrégation doit être suffisamment compact pour limiter la diversité gérée en phase de prévision de planification*. Cette propriété se traduit par une contrainte sur la taille de la diversité résultante.

On peut constater que ces trois propriétés sont antagonistes. Plus de significativité commerciale signifie plus de caractéristiques commerciales et moins de caractéristiques industrielles dans le niveau d'agrégation étant donné un certain degré de compacité. Par ailleurs, un niveau d'agrégation comportant un grand nombre de caractéristiques, qu'elles soient commerciales ou industrielles, va à l'encontre de la propriété de compacité. Pour trouver des situations de compromis entre ces trois propriétés, ce document propose une approche quantitative basée sur un modèle multi-objectifs dont la finalité est d'identifier les meilleurs niveaux d'agrégation en termes de significativité commerciale, de significativité industrielle et de compacité.

4 MODELE

Un bon niveau d'agrégation vise à satisfaire et à aligner les différents acteurs du S&OP. Le modèle analytique développé

dans cette section a pour but d'identifier les sous-ensembles de caractéristiques qui contribuent à faire d'un niveau d'agrégation un compromis efficace du point de vue des propriétés mentionnées en section 3.2. Sans méthodologie adaptée, la recherche d'un niveau d'agrégation peut s'avérer très complexe. Une approche exhaustive par énumération conduirait à comparer un nombre de solutions égale à la somme des combinatoires $\sum_{p=1}^N \frac{N!}{p!(N-p)!}$ où p représente le nombre de caractéristiques présentes dans un niveau d'agrégation. Par ailleurs, l'antagonisme des critères de décision, d'une part, et le caractère combinatoire lié à la diversité, d'autre part, nous conduisent à faire le choix d'une modélisation sous la forme d'un problème multi-objectif dont il sera question en section 4.2. Pour identifier un front de Pareto constitué de niveaux d'agrégations efficaces, une approche de résolution lexicographique sera présentée en section 4.3. Les résultats obtenus suite à l'instanciation de ce modèle sur des données réelles seront discutés et comparés à la situation actuelle en section 5.

4.1 Notations

Le reste du document utilise les notations du tableau 2.

Tableau 2: notations du modèle

Notation	Description
Indices et ensembles	
C	Ensemble des contraintes industrielles
F	Ensemble des caractéristiques
S	Ensemble des caractéristiques commerciales
M	Nombre total de contraintes industrielles
N	Nombre total de caractéristiques
i	Indice des contraintes industrielles, $i = 1.. M$
j	Indices des caractéristiques, $j = 1.. N$
Paramètres	
b_{ij}	$\begin{cases} 1, \text{ si la contrainte } i \text{ contient la} \\ \text{caractéristique } j \\ 0, \text{ sinon} \end{cases}$
$Co_{o_{j,k}}$	$\begin{cases} 1, \text{ si la caractéristique } j \text{ implique} \\ \text{la caractéristique } k \\ 0, \text{ sinon} \end{cases}$
t_i	Nombre de caractéristiques présentes dans l'expression d'une contrainte industrielles $i (t_i \geq 1 \forall i)$
cs_j	$\begin{cases} 1, \text{ si } j \text{ est une caractéristique commerciale } S \\ 0, \text{ sinon} \end{cases}$
cr_i	Poids de criticité pour la contrainte de capacité i
n_j	Taille (nombre d'attributs) de la caractéristique j ($n_j \geq 1 \forall j$)
B_1, B_2	Grand nombre ($B_1 \geq t_i \forall i = 1.. M$ et $B_2 \geq N - 2$)
$DivMax$	Diversité résultante du niveau de granularité actuel
Variables de décision	
Y_j	$\begin{cases} 1, \text{ si la caractéristique } j \text{ est choisie} \\ 0, \text{ sinon} \end{cases}$
Z_i	$\begin{cases} 1, \text{ si la contrainte } i \text{ est couverte} \\ 0, \text{ sinon} \end{cases}$
A_j	$\begin{cases} 1, \text{ si la caractéristique } j \text{ est choisie et n'est pas induite} \\ 0, \text{ sinon} \end{cases}$

Le modèle est alimenté par cinq données d'entrées principales. La première donnée d'entrée représente la liste des caractéristiques commerciales S . Cette liste est utilisée pour calculer le coefficient binaire de signification commerciale cs_j . La deuxième donnée d'entrée est la liste des contraintes capacitaires où chaque contrainte est repérée par l'indice i . Cette liste est utilisée pour générer une matrice binaire dont les composantes b_{ij} , prennent la valeur 1 lorsque la contrainte i s'exprime à travers un ensemble d'attributs dont au moins un est associé à la caractéristique j . b_{ij} prend la valeur 0 sinon. La troisième donnée d'entrée est un score de criticité, cr_i , associé à chaque contrainte i calculé en mesurant la saturation relative d'une contrainte industrielle en tenant compte du nombre d'ordres planifiés concernés par cette contrainte et de la capacité globale de la ligne de production. En d'autres termes, une

contrainte industrielle critique concerne une variété de produit très demandée pour laquelle un petit pourcentage de la capacité totale de la chaîne de montage est alloué. La quatrième donnée d'entrée représente la liste des liens d'induction entre les caractéristiques. Cette liste est exploitée pour calculer la matrice binaire de d'induction Co_{o} entre les caractéristiques. Les composantes $Co_{o_{j,k}}$ de cette matrice prennent la valeur 1 lorsque la caractéristique j induit la caractéristique k . $Co_{o_{j,k}}$ prend la valeur 0 sinon. Une dernière donnée d'entrée fournit le nombre d'attributs, n_j , pour chaque caractéristique j . Les valeurs de n_j sont utilisées pour mesurer la compacité d'un niveau d'agrégation. Il permet également, avec cs_j et $Co_{o_{j,k}}$ de calculer $DivMax$ qui représente la diversité actuelle :

$$DivMax = \prod_{j|cs_j=1 \text{ et } \sum_k Co_{o_{k,j}}=0 \forall k \in S} n_j$$

Le modèle s'appuie sur trois variables de décisions binaires. La variable Y_j nous renseigne sur l'appartenance ou pas d'une caractéristique à un niveau d'agrégation. Après optimisation, un niveau d'agrégation correspond à l'ensemble des caractéristiques j pour lesquels Y_j est égal à 1. La variable Z_i nous indique si la contrainte i est rendue visible ou pas par un niveau d'agrégation. Un niveau d'agrégation rend une contrainte visible lorsque tous les attributs, utilisés pour décrire la contrainte, sont représentés par leurs caractéristiques associées dans le niveau d'agrégation. En réalité, il n'est pas toujours nécessaire d'avoir une représentativité exhaustive en présence de liens d'induction entre les caractéristiques. La variable A_j permet d'introduire cette nuance en précisant le caractère induit ou non induit d'une caractéristique.

4.2 Modèle multi-objectifs

4.2.1 Fonction objective

Le modèle de sélection du niveau d'agrégation est un programme linéaire multi-objectifs en nombres entiers mixtes, où les deux premières propriétés introduites en section 3.2 se traduisent en objectifs d'optimisation, tandis que la propriété de compacité est considérée comme une contrainte du modèle. La formulation qui en résulte est la suivante :

$$OF_1 = \text{Min} \sum_{j=1}^N A_j * (1 - cs_j)$$

$$OF_2 = \text{Max} \sum_{i=1}^M cr_i * Z_i$$

La première fonction objectif (**OF1**) est liée aux préférences des équipes commerciales définies par la première propriété. Elle calcule le nombre de caractéristiques non commerciales présentes dans le niveau d'agrégation. Pour garantir la significativité commerciale, ce nombre doit être minimisé.

La deuxième fonction objective (**OF2**) représente la deuxième propriété liée aux préférences des équipes industrielles. Elle mesure le ratio de contraintes capacitaires critiques rendues visibles par le niveau d'agrégation. Pour garantir la significativité industrielle du niveau d'agrégation, cette part doit être maximisée.

4.2.2 Contraintes du modèle

Le modèle d'optimisation cherche à identifier des sous-ensembles de caractéristiques qui minimisent **OF1** et maximisent **OF2** sans dépasser un certain seuil de compacité qui correspond à la diversité maximale gérable en processus S&OP. La première contrainte du modèle (**MC1**) force Z_i à être égal à 0 si au moins une caractéristique j impliquée dans la contrainte industrielle i n'est pas présente dans le niveau d'agrégation (c'est-à-dire lorsque $Y_j = 0$). Afin d'assurer ce calcul, B doit être au moins égale à $\text{Max } t_i, \forall i \in M$.

$$\sum_{j=1}^N b_{ij} \times Y_j - t_i \geq B_1(Z_i - 1), \forall i = 1..M \text{ (MC1)}$$

La deuxième contrainte du modèle (MC2) fixe Z_i à 1 si toutes les caractéristiques, j , impliquées dans la contrainte industrielle i sont présentes dans le niveau d'agrégation (lorsque $Y_j = 1$).

$$\sum_{j=1}^N b_{ij} * Y_j - t_i \leq (Z_i - 1), \forall i = 1..M \text{ (MC2)}$$

D'autre part, la prise en compte des liens d'induction entre les caractéristiques est garantie par deux contraintes. Les deux contraintes (MC3' & MC3'') introduisent la variable de décision A_j qui permet d'identifier les caractéristiques j suffisantes (c'est à dire celles qui ne sont pas induites par une autre caractéristique présente dans le niveau d'agrégation). $B_2 \geq N - 2$ est un grand nombre utilisé pour garantir ce calcul. La quatrième contrainte (MC4) sélectionne automatiquement les caractéristiques induites lorsque l'inducteur est présent dans le niveau d'agrégation.

$$\forall j \in 1..N, \sum_{k=1}^N Y_k * \text{Coo}_{k,j} - Y_j \geq -B_2 * A_j, \text{ (MC3')}$$

$$\forall j \in 1..N, \sum_{k=1}^N Y_k * \text{Coo}_{k,j} - Y_j \leq B_2 * (1 - A_j) - 1, \text{ (MC3'')}$$

$$Y_p \geq Y_j * \text{Coo}_{j,p} \quad \forall p, j = 1..N \text{ (MC4)}$$

La dernière contrainte du modèle (MC5) traduit la propriété de compacité que doit avoir un bon niveau d'agrégation. La contrainte de compacité garantit que la diversité qui résulte d'un niveau d'agrégation ne dépasse pas la charge de travail actuelle. Étant donné que la diversité résulte de la combinaison des attributs des caractéristiques non-induites retenues dans le niveau d'agrégation, la contrainte peut être simplement exprimée comme suit :

$$\prod_{j|A_j=1} n_j \leq \text{DivMax}$$

Avant d'introduire cette formulation dans le modèle, il est nécessaire de la linéariser. Une simple transformation logarithmique permet d'obtenir la formulation suivante (les étapes de linéarisation sont fournies en annexe) :

$$\sum_{j=1}^N \log(n_j) * A_j \leq \log(\text{DivMax}) \text{ (MC5)}$$

4.3 Méthode de résolution

Différentes méthodes peuvent être utilisées pour résoudre les problèmes d'optimisation multi-objectifs [Marler et al., 2004]. Nous distinguons les approches qui cherchent à optimiser un critère unique de synthèse comme la somme pondérée ou le goal-programming qui visent à identifier un ensemble de solutions efficaces au regard des objectifs formant un front de Pareto. Compte tenu de la nature du problème où il est question d'identifier des solutions alternatives à la situation actuelle, des besoins d'explicabilité a posteriori et d'incomparabilité des objectifs commerciaux et industriels, c'est ce deuxième type d'approche qui a été retenu dans notre cas.

En raison de sa capacité à identifier des solutions de Pareto, nous optons pour une méthode de résolution lexicographique largement utilisée pour résoudre ce genre de problèmes.

Le principe de cette méthode consiste à rechercher de manière itérative la solution optimale au regard d'une des fonctions objectives tout en contraignant la seconde fonction objective en fonction des résultats de l'itération précédente. Chaque itération consiste donc à résoudre un programme linéaire en nombres entiers mixte mono-objectif. La résolution itérative est répétée jusqu'à ce que le problème devienne irréalisable.

Nous choisissons de résoudre **OF1** et d'appliquer itérativement une contrainte sur **OF2**. Ce choix est motivé par le fait que la minimisation de **OF1** permet de réduire le nombre des caractéristiques non-commerciales et donc de maintenir une diversité associée au niveau d'agrégation proche de celle qui est actuellement en vigueur. Une maximisation de **OF2** conduirait à des solutions riches en caractéristiques industrielles avec une tendance à saturer rapidement la contrainte de diversité.

La procédure de résolution se résume comme suit :

Objectif: Minimiser le nombre des caractéristiques non-commerce

$$\text{Min} \sum_{j=1}^N A_j * (1 - cs_j)$$

Contraintes:

$$\sum_{j=1}^N b_{ij} * Y_j - t_i \geq B_1 * (Z_i - 1), \forall i = 1..M$$

$$\sum_{j=1}^N b_{ij} * Y_j - t_i \leq (Z_i - 1), \forall i = 1..M$$

$$\sum_{k=1}^N Y_k * \text{Coo}_{k,j} - Y_j \geq -B_2 * A_j, \forall j \in 1..N$$

$$\sum_{k=1}^N Y_k * \text{Coo}_{k,j} - Y_j \leq B_2 * (1 - A_j) - 1, \forall j \in 1..N$$

$$Y_p \geq Y_j * \text{Coo}_{j,p}, \forall p, j = 1..N$$

$$\sum_{j=1}^N \log(n_j) * A_j \leq \log(\text{DivMax})$$

$$\left(\sum_{i=1}^M cr_i * Z_i \right)^k > \left(\sum_{i=1}^M cr_i * Z_i \right)^{k-1}, \text{ pour } k \in \text{iterations}$$

5 ÉTUDE DE CAS

Le modèle ainsi que la méthode de résolution présentés en section 4 ont été instanciés sur des données réelles mises à disposition par le partenaire industriel de cette étude. Plusieurs scénarios relatifs à différentes gammes de produit ont été analysés. La section 5.1 décrit sommairement les scénarios étudiés. La section 5.2 détaille et examine les résultats observés pour chaque scénario.

5.1 Paramètres du problème

Une instanciation numérique du modèle a été réalisée sur quatre véhicules qui se différencient par le segment du marché auquel ils appartiennent. Le modèle proposé a été programmé en utilisant CPLEX 12.9 sur un PC équipé d'un processeur Intel i5 6th Gen 2.4GHz CPU et une RAM de 8Gb. Le solveur est lancé dans chacun des cas pour obtenir une liste de solutions efficaces qui constituent des compromis alternatifs au niveau d'agrégation actuel. Le tableau 3 donne les paramètres du problème liés à chacun des quatre véhicules retenus pour l'analyse.

Tableau 3: paramètres de l'étude du cas

Véhicule	Toutes les caractéristiques	Caractéristiques commerciales	Contraintes industrielles	DivMax	Temps d'execution (s)
Compacte	246	18	40	159344640	10,49
Bestseller	207	19	39	4462920	9,16
Electrique	171	19	11	23328	9,52
Premium	214	19	35	2640000	10,26

5.2 Résultats et discussion

En suivant la procédure lexicographique, un ensemble de solutions a été obtenu pour chacun des quatre scénarios. Pour chaque véhicule, ces solutions sont représentées par des points bleus sur la figure 1. Chaque point bleu représente donc un niveau d'agrégation faiblement ou fortement efficace au sens de Pareto. La courbe orange représente le front de Pareto tracée à partir des solutions fortement efficaces. La courbe jaune représente quant à elle le ratio entre la diversité engendrée par un niveau d'agrégation efficace et la diversité associée au niveau d'agrégation actuel.

Ce premier résultat constitue un outil d'aide à la décision pour le choix d'un niveau d'agrégation qui garantit un compromis efficace entre les objectifs commerciaux et industriels tout en respectant une limite supérieure pour la diversité résultante. Il appartient aux décideurs, en fonction des différentes solutions proposées, de choisir le niveau d'agrégation le plus approprié.

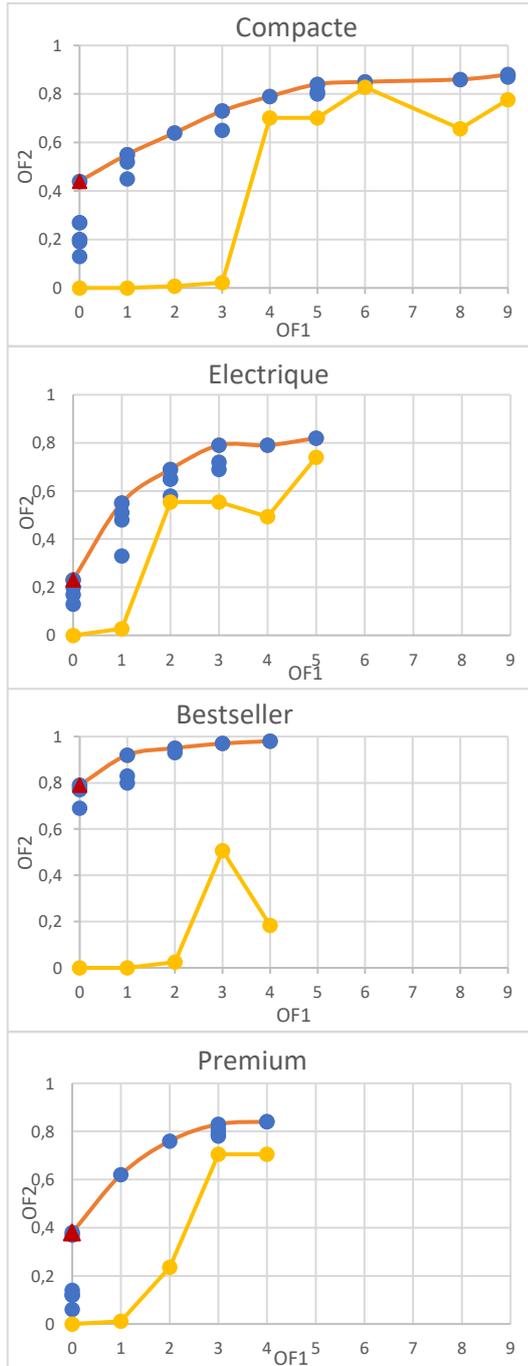


Figure 1 : Résultats de l'optimisation (bleu : ensemble de solutions, orange : solution pareto, rouge : solution actuelle, jaune : charge résultante)

5.3 Situation actuelle

La mesure de performance de la situation actuelle est représentée sur la figure 1 par un triangle rouge en bas à gauche du front de Pareto. En effet, le niveau d'agrégation actuel ne contient que des caractéristiques commerciales et correspond donc à une borne inférieure de **OF1** (comme aucune caractéristique commerciale n'est absente du niveau d'agrégation actuel, le score de **OF1** est de 0). D'autre part, comme toutes les caractéristiques commerciales pertinentes (non induites) sont présentes dans le niveau d'agrégation actuel, le score **OF2** du triangle rouge correspond à la valeur maximale possible lorsque **OF1** est égal à 0. Les points bleus qui se situent sur le même axe vertical et en dessous de la solution actuelle (triangle rouge) représentent des solutions composées de sous-ensembles de caractéristiques commerciales pertinentes.

En outre, à l'exception du modèle *Bestseller*, qui présente un score initial de **OF2** = 79%, les autres modèles de voitures obtiennent des scores **OF2** compris entre 23 % et 38 % lorsque **OF1** = 0. Le modèle *Compacte* obtient le score **OF2** le plus faible pour le niveau d'agrégation actuel. Cela signifie que l'ensemble des caractéristiques présentes dans le niveau d'agrégation actuel ne permettent de rendre visible que 23% de contraintes industrielles. Un tel score signifie que la demande exprimée selon ce niveau d'agrégation doit faire l'objet d'une décomposition importante pour être en mesure de vérifier la conformité de la demande avec les contraintes capacitaires. Ce deuxième résultat illustre la faible significativité industrielle d'un niveau d'agrégation composé exclusivement de caractéristiques commerciales. Tous les véhicules ne présentent pas le même comportement. Le véhicule *Bestseller* affiche un bon score sur **OF2** pour le niveau d'agrégation actuel. Cela révèle une possible tendance à faire des choix d'agrégation uniformes en ne tenant compte que des contraintes industrielles relatives aux véhicules les plus vendus. Cette intuition doit être confirmée par des analyses complémentaires.

5.4 Solutions alternatives

Une analyse de la forme du front de Pareto montre que de petites modifications par rapport aux choix actuels peuvent apporter des améliorations significatives du point de vue de la significativité industrielle du niveau d'agrégation. Plus précisément, les points correspondants à **OF1**=1 montrent qu'une légère détérioration de **OF1** peut avoir un impact significatif sur la visibilité des contraintes critiques mesurée par **OF2**. L'introduction dans le niveau d'agrégation d'une caractéristique non commerciale entraîne une amélioration de **OF2** de +11% pour le modèle *Compacte*, de +13% pour le modèle *Bestseller*, de +32% pour le modèle *Electrique* et de +24% pour le modèle *Premium*. Un tel compromis est de nature à minimiser les risques liés aux opérations de désagrégation.

La forme des courbes obtenues montre une disparité entre les différents véhicules. La longueur du front de Pareto (donnée par la valeur maximale sur l'axe des abscisses qui représente **OF1**) peut s'expliquer par la diversité des contraintes relatives à chaque véhicule. La pente initiale du front de Pareto nous renseigne quant à elle sur la pertinence du niveau d'agrégation actuel (plus cette pente est prononcée, plus il est intéressant de faire évoluer le niveau d'agrégation actuel).

A titre d'exemple, la pente du front de Pareto du véhicule *compacte* est relativement faible tandis que sa longueur est plus importante comparée aux autres véhicules. Cela traduit, pour le véhicule *compacte*, à la nécessité d'introduire dans le niveau d'agrégation un grand nombre de caractéristiques non-commerciales pour atteindre un bon niveau de prise en charge

de contraintes. D'un autre côté, la forte pente initiale des fronts de Pareto relatifs aux véhicules *Electrique et Premium* montre qu'une amélioration importante de la significativité industrielle est obtenue en introduisant seulement quelques caractéristiques non-commerciales dans le niveau d'agrégation.

5.5 Analyse de diversité engendrée

De manière générale, nous pouvons constater que la diversité engendrée par les premières solutions alternatives sur le front de Pareto (celles qui se situent le plus à gauche sur l'axe des abscisses) est très faible par rapport à la *diversité* associée au niveau d'agrégation actuel (*DivMax*). Ces solutions améliorent rapidement la significativité industrielle sans déstabiliser le processus actuel par une augmentation de la diversité à gérer pour les prévisionnistes, les planificateurs et les décideurs. Pour conclure, il peut être établi, sans perte de généralité, qu'en apportant de légers changements à la situation actuelle, il est possible de fournir aux équipes industrielles des informations plus pertinentes pour l'analyse des contraintes et ainsi de garantir un meilleur alignement des acteurs impliqués dans le S&OP sans bouleverser profondément les habitudes.

6 CONCLUSION ET PERSPECTIVES

Le S&OP est un processus de planification tactique dont l'efficacité repose largement sur l'alignement et la coordination de ses principaux acteurs. Le niveau d'agrégation actuel retenu pour soutenir les arbitrages en phase de planification favorise le point de vue commercial au détriment des contraintes industrielles. Cette situation implique des opérations de désagrégation de la demande basées sur des informations datées issues des historiques pour vérifier la compatibilité de la demande, et éventuellement l'ajuster, aux contraintes industrielles. Ce processus s'appuie sur une hypothèse de stabilité de la demande rarement observée sur le terrain et conduit à des arbitrages charge/capacité réalisés sur la base d'informations erronées.

Ce document fournit un modèle analytique d'aide à la décision qui tente d'apporter des solutions au problème du choix du niveau d'agrégation en phase de planification tactique.

Trois propriétés de ce qui fait un bon niveau d'agrégation ont été décrites puis traduites en objectifs et contraintes dans un modèle d'optimisation multi-objectif. Le modèle proposé a été testé sur des données réelles fournies par un constructeur automobile. Quatre cas relatifs à des véhicules qui se distinguent par leur positionnement en termes de segment de marché ont été étudiés. Les résultats obtenus nous ont permis de mesurer la performance et d'évaluer la pertinence du niveau d'agrégation actuel appliqué à chacun des quatre véhicules. Les solutions alternatives à la situation actuelle ont été analysées du point de vue de leur significativité commerciale, de leur significativité industrielle et de leur compacité. Cette analyse a permis de mettre en avant des comportements hétérogènes d'un véhicule à l'autre et d'identifier des compromis plus équilibrés pour remplacer le niveau d'agrégation actuel et permettre une meilleure coordination des acteurs impliqués dans le S&OP.

Bien que le modèle proposé fournisse des indications précieuses quant au choix du niveau d'agrégation, il a fallu simplifier le problème réel en adoptant quelques hypothèses qu'il s'agira de lever progressivement dans de futurs travaux. En effet, cette étude s'est limitée aux objectifs des entités commerciales et industrielles, mais elles ne sont pas les seules à jouer un rôle dans le S&OP. Un modèle plus riche pourrait intégrer des objectifs et des contraintes portés par d'autres acteurs impliqués dans ce processus. Par exemple, des objectifs financiers peuvent être pris en compte en définissant une liste des caractéristiques

prépondérantes, comme le moteur par exemple, dans les arbitrages financiers. À cette fin, un objectif d'optimisation supplémentaire visant à maximiser la présence de caractéristiques financières dans le niveau d'agrégation pourrait être introduit.

Par ailleurs, un axe de recherche à explorer concerne la prise en compte de caractéristiques soumises à des réglementations nationales, régionales ou internationales comme celle de la Corporate Average Fuel Economy (CAFE) qui impose des pénalités en cas de non-respect de seuils d'émission de CO₂. Les caractéristiques dont dépendent les émissions de CO₂ peuvent être intégrées dans le niveau d'agrégation pour contrôler le mix-produit et faire en sorte qu'il soit compatible avec les limites imposées par la réglementation. Enfin, d'un point de vue technique, le modèle proposé considère la diversité engendrée par le niveau d'agrégation comme une contrainte et non comme une valeur à minimiser. La transformation de cette contrainte en objectif peut enrichir la palette des solutions alternatives à la situation actuelle. De plus, le calcul de la diversité engendré et de *DivMax* représente une borne supérieure de la diversité réelle. En effet, plusieurs contraintes de compatibilité inter-critères permettant de limiter la diversité résultante ne sont pas prises en compte vu la complexité du calcul de diversité tenant en compte ces contraintes de manière dynamique. Cette étude peut bénéficier énormément d'un futur travail permettant de fiabiliser ce calcul.

7 ANNEXE

La contrainte originale est formulée comme suit :

$$\prod_{j|A_j=1} n_j \leq DivMax \quad (1)$$

Comme $A_j \in \{0,1\}$, (1) peut être transformé comme suit:

$$\prod_{j=1}^N 1 - A_j + n_j \times A_j \leq DivMax \quad (2)$$

En appliquant le logarithme de base e à (2) on obtient:

$$\sum_{j=1}^N \log(1 - A_j + n_j \times A_j) \leq \log(DivMax) \quad (3)$$

De plus, comme $A_j \in \{0,1\}$, on aura:

$$\begin{cases} \text{quand } A_j = 0 \Rightarrow \log(1 - A_j + n_j \times A_j) = 0 = A_j \times \log(n_j) \\ \text{quand } A_j = 1 \Rightarrow \log(1 - A_j + n_j \times A_j) = \log(n_j) = A_j \times \log(n_j) \end{cases}$$

Ainsi, (3) devient équivalente à la formulation suivante:

$$\sum_{j=1}^N \log(n_j) \times A_j \leq \log(DivMax) \quad (MC5)$$

8 REFERENCES

- Blackstone, J., & Cox, J. F. (2005). APICS. In 2005 Amer Production & Inventory (Ed.), APICS: The Associations for Operation Management. APICS.
- Blecker, T., & Abdelkafi, N. (2006). Complexity and variety in mass customization systems: analysis and recommendations. *Management Decision*, 44(7), 908–929. doi: 10.1108/00251740610680596
- Chatras, C., Giard, V., & Sali, M. (2016). Mass customisation impact on bill of materials structure and master production schedule development. *International Journal of Production Research*, 54(18), 5634–5650. doi: 10.1080/00207543.2016.1194539
- Daaboul, J., Da Cunha, C., Bernard, A., & Laroche, F. (2011). Design for mass customization: Product variety vs.

- process variety. *CIRP Annals - Manufacturing Technology*, 60(1), 169–174. doi: 10.1016/j.cirp.2011.03.093
- ElMaraghy, H., Schuh, G., Elmaraghy, W., Piller, F., Schönsleben, P., Tseng, M., & Bernard, A. (2013). Product variety management. *CIRP Annals - Manufacturing Technology*, 62(2), 629–652. doi: 10.1016/j.cirp.2013.05.007
- Feng, Y., D'Amours, S., & Beauregard, R. (2008). The value of sales and operations planning in oriented strand board industry with make-to-order manufacturing system: Cross functional integration under deterministic demand and spot market recourse. *International Journal of Production Economics*, 115(1), 189–209. doi: 10.1016/j.ijpe.2008.06.002
- Fontan, G., Merce, C., Hennet, J. C., & Lasserre, J. (2005). Hierarchical scheduling for decision support. *Journal of Intelligent Manufacturing*, 16(2), 235–242. doi: 10.1007/s10845-004-5891-9
- Gelders, L. F., & Van Wassenhove, L. N. (1982). Hierarchical integration in production planning: Theory and practice. *Journal of Operations Management*, 3(1), 27–35. doi: 10.1016/0272-6963(82)90019-5
- Ghrab, Y., & Sali, M. (2020). A New S&OP Maturity Assessment Model for Mass Customization Production Systems: A Case Study of an Automotive OEM. *Proceedings of The Eight International Conference on Information Systems, Logistics and Supply Chain*, 70–78.
- Goh, S. H., & Eldridge, S. (2019). Sales and Operations Planning: The effect of coordination mechanisms on supply chain performance. *International Journal of Production Economics*, 214(May 2018), 80–94. doi: 10.1016/j.ijpe.2019.03.027
- Grimson, J. A., & Pyke, D. F. (2007). Sales and operations planning: An exploratory study and framework. *The International Journal of Logistics Management*, 18(3), 322–346. doi: 10.1108/09574090710835093
- Hax, A. C., & May, H. C. M. (1973). A Hierarchical Production Planning System. *Interfaces*.
- Herrmann, J. W., Mehra, A., Minis, I., & Proth, J. M. (1994). Hierarchical production planning with part, spatial and time aggregation. *Proceedings of the 4th International Conference on Computer Integrated Manufacturing and Automation Technology, CIMAT 1994*, 430–435. doi: 10.1109/CIMAT.1994.389037
- Kanyalkar, A. P., & Adil, G. K. (2005). An integrated aggregate and detailed planning in a multi-site production environment using linear programming. *International Journal of Production Research*, 43(20), 4431–4454. doi: 10.1080/00207540500142332
- Kashkoush, M., & ElMaraghy, H. (2016). Product family formation by matching Bill-of-Materials trees. *CIRP Journal of Manufacturing Science and Technology*, 12, 1–13. doi: 10.1016/j.cirpj.2015.09.004
- Lalami, I., Frein, Y., & Gayon, J.-P. (2017). Production planning in automotive powertrain plants: a case study. *International Journal of Production Research*, 55(18), 5378–5393. doi: 10.1080/00207543.2017.1315192
- Laurent Lim, L., Alpan, G., & Penz, B. (2014). Reconciling sales and operations management with distant suppliers in the automotive industry: A simulation approach. *International Journal of Production Economics*, 151, 20–36. doi: 10.1016/j.ijpe.2014.01.011
- Liberatore, M. J., & Miller, T. (1985). Hierarchical Production Planning System. *Interfaces*, 15(4), 1–11. doi: 10.1287/inte.15.4.1
- Marino, G., Zotteri, G., & Montagna, F. (2018). Consumer sensitivity to delivery lead time: a furniture retail case. *International Journal of Physical Distribution and Logistics Management*, 48(6), 610–629. doi: 10.1108/IJPDLM-01-2017-0030
- Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. In *Structural and Multidisciplinary Optimization* (Vol. 26, Issue 6, pp. 369–395). doi: 10.1007/s00158-003-0368-6
- Olhager, J. (2003). Strategic positioning of the order penetration point. *International Journal of Production Economics*, 85(3), 319–329. doi: 10.1016/S0925-5273(03)00119-1
- Oliva, R., & Watson, N. (2011). Cross-functional alignment in supply chain planning: A case study of sales and operations planning. *Journal of Operations Management*, 29(5), 434–448. doi: 10.1016/j.jom.2010.11.012
- Özdamar, L., Bozyel, M. A., & Birbil, S. I. (1998). A hierarchical decision support system for production planning (with case study). *European Journal of Operational Research*, 104(3), 403–422. doi: 10.1016/S0377-2217(97)00016-7
- Pereira, D. F., Oliveira, J. F., & Carravilla, M. A. (2020). Tactical sales and operations planning: A holistic framework and a literature review of decision-making models. In *International Journal of Production Economics* (Vol. 228). doi: 10.1016/j.ijpe.2020.107695
- Pil, F. K., & Holweg, M. (2004). Linking product variety to order-fulfillment strategies. *Interfaces*, 34(5), 394–403. doi: 10.1287/inte.1040.0092
- Ramdas, K. (2003). Managing product variety: An integrative review and research directions. *Production and Operations Management*, 12(1), 79–101. doi: 10.2139/ssrn.290774
- Shao, X. F. (2020). What is the right production strategy for horizontally differentiated product: Standardization or mass customization? *International Journal of Production Economics*, 223(October 2019), 107527. doi: 10.1016/j.ijpe.2019.107527
- Stäblein, T., Holweg, M., & Miemczyk, J. (2011). Theoretical versus actual product variety: How much customisation do customers really demand? *International Journal of Operations and Production Management*, 31(3), 350–370. doi: 10.1108/01443571111111955
- Staeblein, T., & Aoki, K. (2015). Planning and scheduling in the automotive industry: A comparison of industrial practice at German and Japanese makers. *International Journal of Production Economics*, 162, 258–272. doi: 10.1016/j.ijpe.2014.07.005
- Tang, M., Qi, Y., & Zhang, M. (2017). Impact of Product Modularity on Mass Customization Capability: An Exploratory Study of Contextual Factors. *International Journal of Information Technology & Decision Making*, 16(04), 939–959. doi: 10.1142/S0219622017410012
- Thomé, A. M. T., Scavarda, L. F., Fernandez, N. S., & Scavarda, A. J. (2012). Sales and operations planning and the firm performance. *International Journal of Productivity and Performance Management*, 61(4), 359–381. doi: 10.1108/17410401211212643
- Thomé, A. M. T., Scavarda, L. F., Fernandez, N. S., & Scavarda, A. J. (2012). Sales and operations planning: A research synthesis. *International Journal of Production*

- Economics*, 138(1), 1–13. doi:
10.1016/j.ijpe.2011.11.027
- Thomé, A. M. T., Sousa, R. S., & Do Carmo, L. F. R. R. S. (2014). Complexity as contingency in sales and operations planning. *Industrial Management and Data Systems*, 114(5), 678–695. doi: 10.1108/IMDS-10-2013-0448
- Tuomikangas, N., & Kaipia, R. (2014). A coordination framework for sales and operations planning (S&OP): Synthesis from the literature. *International Journal of Production Economics*, 154, 243–262. doi: 10.1016/j.ijpe.2014.04.026
- Um, J., Lyons, A., Lam, H. K. S., Cheng, T. C. E., & Dominguez-Pery, C. (2017). Product variety management and supply chain performance: A capability perspective on their relationships and competitiveness implications. *International Journal of Production Economics*, 187, 15–26. doi: 10.1016/j.ijpe.2017.02.005
- Wagner, S. M., Ullrich, K. K. R., & Transchel, S. (2014). The game plan for aligning the organization. *Business Horizons*, 57(2), 189–201. doi: 10.1016/j.bushor.2013.11.002
- Wan, X., & Sanders, N. R. (2017). The negative impact of product variety: Forecast bias, inventory levels, and the role of vertical integration. *International Journal of Production Economics*, 186(November), 123–131. doi: 10.1016/j.ijpe.2017.02.002

PERFORMANCE EVALUATION OF A CENTRALIZED MAINTENANCE WORKSHOP WITH REMANUFACTURING

SINA AKBARI HAGHIGHI¹, MARIA DI MASCOLO², ZINEB SIMEU-ABAZI³

Univ. Grenoble Alpes, CNRS, Grenoble INP*, G-SCOP, 38000 Grenoble, France

*Institute of Engineering Univ. Grenoble Alpes

¹sina.akbari-haghighi@grenoble-inp.org

²maria.di-mascolo@grenoble-inp.fr

³zineb.simeu-abazi@grenoble-inp.fr

Abstract – Having a Centralized Maintenance Workshop (CMW) allows companies to centralize their resources to assure production equipment availability. The introduction of Circular products and remanufacturing processes helps industry players avoid excessive raw material consumption. In a CMW, the remanufacturing process, besides repairing and purchasing new components, recovers a portion of parts that are not repairable. This paper proposes a closed-loop model considering the lifetime of components, which well represents the real-life components' flow through the CMW. Adding the remanufacturing process to the system has some impacts on the performance of the CMW. Main indicators like the waiting time of equipment in CMW and maintenance cost must be studied to be able to evaluate the performances. To this purpose, a new CMW organization is proposed based on a real case in the automotive industry. Considering different possible simulation scenarios is crucial to get reliable results from the proposed closed-loop queuing network model. Several simulations are conducted to evaluate the CMW performance and find the optimal combinations of decision variables.

Keywords – Centralized maintenance workshop, Closed-loop queuing network, Circular economy, Remanufacturing, Performance evaluation, System modelling, Optimization, Arena simulation.

1 INTRODUCTION

In recent years, circular products absorbed more attention from industry players. Companies try to implement sustainability strategies, and the aim is to reduce the waste and environmental effects and maximize the productivity and efficiency of raw material consumption (Popa et al., 2016).

Maintenance is a process to maintain the functionality of a part of a facility, which is required in production sites. Traditionally, the maintenance process was limited to repairing a failed component or its replacement by a new one.

Centralized maintenance workshops permit companies to gather all required resources to perform maintenance. The goal is to achieve maximum operational availability of the production lines (Simeu-Abazi et al., 2014). Therefore, maintenance efficiency directly impacts a manufacturing line's capacity and service level and the product's final cost (Driessen et al., 2020). All these factors can affect customer satisfaction, which is the final goal of every business unit.

Knowing the importance of the maintenance in customer satisfaction and the company's global performance, evaluation and optimization of CMW became an essential need of the companies.

In the literature, the main focus of the CMW is on the corrective maintenance and periodic overhauls by replacing the malfunctioned parts with the existing spare parts. All the repaired components are considered to have the same quality as the new ones (Sleptchenko et al., 2019). However, one of these models' main assumptions is that the failed units visiting the CMW are always repairable. In the literature, some models

consider that a portion of the components entering the CMW are not repairable (Li et al., 2013).

There are different processes in the CMW to make a failed unit usable (Djeunang Mezafack et al., 2020). Repairing the failed unit, replacing it with a new one are traditional processes. In addition to these two methods, the remanufacturing process can help to extend the product's life (Kleber et al., 2011).

Depending on the maintenance process, the equipment's failure rate will be different and they will return to the CMW at different times even if they leave it simultaneously.

Therefore, to have more realistic outputs, the quality of maintained equipment should be considered. The lifetime of equipment in the production site is a suitable parameter to indicate the quality of the maintenance process. To be able to differentiate between maintenance processes from the lifetime perspective and track it, it is needed to have a closed-loop of equipment flow. This approach well represents real-life equipment flow through the CMW (Assadi et al., 2018).

This study, which is inspired by a real case in the automotive industry, extends the works of (Simeu-Abazi et al., 2014) and (Djeunang Mezafack et al., 2020) in two ways:

- Considering a closed-loop flow between the production line and CMW, therefore the components visit the CMW more than once,
- Taking into account the lifetime of the components leaving the CMW.

The study's main objective is to evaluate the CMW's performance in the presence of remanufacturing and new component purchasing. To achieve this goal, we have

structured this study into six sections. Section 1 contains the introduction and a short explanation of the objectives. Problem description and the organization of the CMW are presented in section 2. Section 3 gives more details about the constructed model properties, queuing network of the system, and the decision variables. A detailed explanation of the experiments and the critical indicators of the CMW's performance are represented in section 4. Several simulations are performed to obtain the results. They are well discussed in section 5. Finally, section 6 is dedicated to the conclusion and further research opportunities.

2 PROBLEM DESCRIPTION

A CMW is designed first to find the origin of the equipment's failure in the production workshop and second to repair it (Simeu-Abazi et al., 2014). Depending on the nature of the failure, severity of the failure, and experts' recognition, different processes are defined in CMW. According to the required procedures, failed units are directed to the stations with varying service time and costs.

Serving different production lines by a CMW makes it necessary to trace the failed units. It is more crucial when there is a closed-loop system where the equipment visits the CMW more than once during its life cycle.

In this study, the failed unit is a spindle of the machining center used to mass production of main engine parts in the automotive industry. It is composed of 2 parts; one is the rotating union shown in Figure 1, and the other is the rest of the spindle. From now on, we will call them "component" and "equipment," respectively. The rotating union provides high pressure through spindle coolant liquid during the machining process. It is necessary for high precision machining of automotive parts to reach high surface quality.

The CMW supports different production lines with 100 identical machining centres in total. It is assumed that only components fail, and the equipment is reusable without any further action.

The case study is taken from a private Iranian company that is active in automotive aluminium engine parts' mass production. The company is one of the leaders in its sector with significant market share.

The organization of CMW is shown in Figure 2. After failed unit diagnosis and disassembly, the equipment waits for component arrival in a buffer before the assembly station.

There are different possible paths through the CMW regarding components' last number of visits to CMW and their condition. In case the component is repairable, the tracing system checks the component's previous visits and sends them to one of the three repairing stations. Each component can be repaired a maximum of 3 times. Repair 1, 2, and 3 in Figure 2 are the stations for a first, second, and third time repairing, respectively.

If the component is not repairable, either because of its condition, or because it reaches the maximum three times repair, it may transfer to the remanufacturing station or be replaced by a new component in the station called NC Purchase.

Remanufacturing is a process in which all wearable components have been replaced and mechanical tolerances have been restored to their original standards.

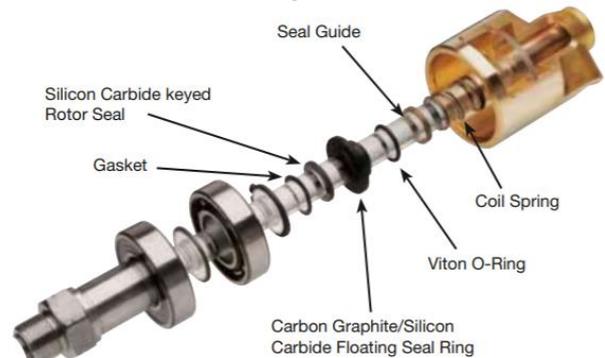


Figure 1. Rotating union¹

Waiting equipment is always assembled with the first available component, which can be repaired, remanufactured, or new. If all types of components are available, the priority of use will be 1- new component, 2- remanufactured, 3- first-time repair, 4- second-time repair, 5- third-time repair.

The goal is to evaluate the system's performance under different conditions and scenarios and identify the situation that leads to the best system performance. The system

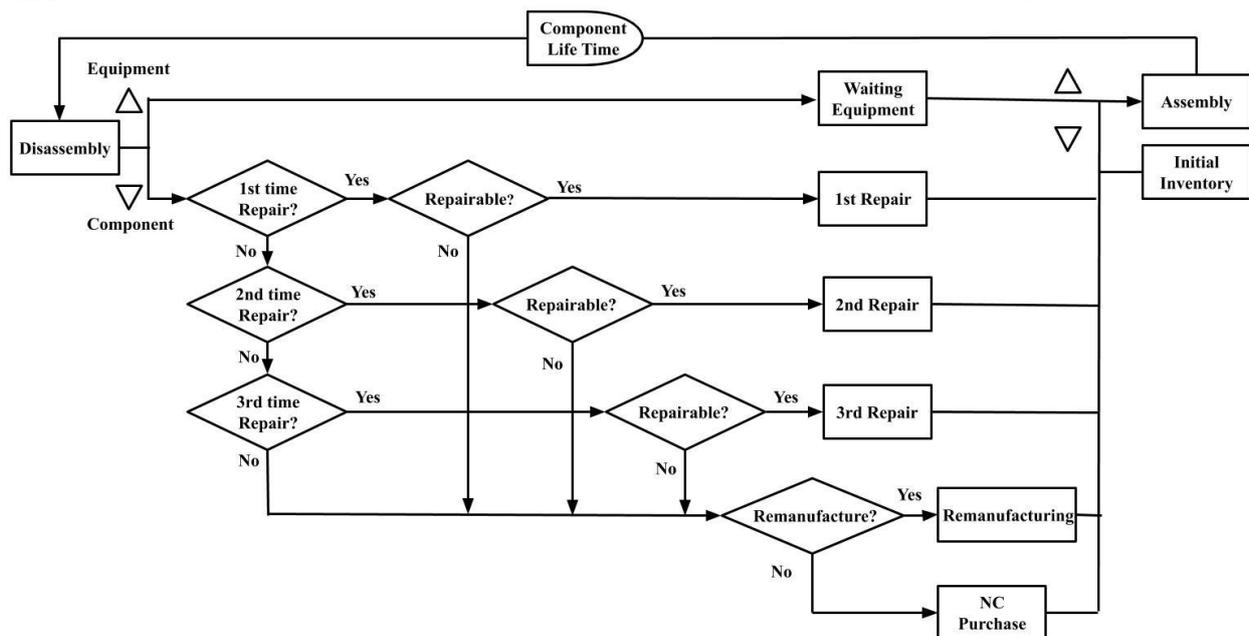


Figure 2. CMW organization

¹<https://www.deublin.eu/rotating-unions>

performance's key indicators are the average waiting time of the system's equipment and average maintenance cost per machine in 10 years. Thus, we have taken the methodology of the queuing network used in (Simeu-Abazi et al., 2014) to evaluate the CMW performance.

3 SIMULATION MODELING

3.1 Model properties

Based on queuing network modelling, a simulation model in Arena is conducted to study the system performance under different conditions. General properties of the system are as follows:

- The total number of machines that are supported by the CMW is 100,
- Each component can be repaired a maximum of 3 times, which means more than one visit is possible.
- At the beginning of the simulation, it is assumed that all machines are equipped with new components.
- Assembly and disassembly are single server stations (components processed one by one).
- There is no blocking in the system thanks to unlimited buffer capacities.
- There are two common servers for all three repairing stations (Two equipment can be repaired at the same time).
- NC purchase and remanufacturing stations have infinite servers. It means there is no restriction in supplying remanufactured and new components, and the suppliers have enough capacity.
- Classification and test stations are very rapid with little service time.
- First arrival of 100 failed units from different production lines follow a Poisson distribution with a rate of λ .
- All operational stations follow an exponential distribution with rates shown in Table 1.

Table 1. Service rates of the stations

Station	Service rate	
disassembly	μ_{ds}	
1st repair	μ_{rp1}	
2nd repair	μ_{rp2}	
3rd repair	μ_{rp3}	
remanufacturing	μ_{rm}	
NC purchase	μ_{nw}	
assembly	μ_{as}	
Component life cycle	1st repair	μ_{L1}
	2 nd repair	μ_{L2}
	3rd repair	μ_{L3}
	remanufacturing	μ_{L4}
	NC purchase	μ_{L5}

3.2 Queuing network representation

Figure 3 represents the queuing network of understudy CMW considering all the mentioned properties in the previous

section. In addition to parameters in Table 1, the following parameters are used to characterize the system:

- r : proportion of remanufacturing components $\in [0, 1]$,
- r_{p1} , r_{p2} , and r_{p3} : proportion of repairable components for the first, second, and third times, respectively $\in [0, 1]$,
- B : initial inventory safety stock
- λ : first arrival rate of 100 failed units from different production lines,
- μ_{cl} : classification station service rate,
- μ_{r1} : service rate of the first-time repairability test stations,
- μ_{r2} : service rate of the second-time repairability test stations,
- μ_{r3} : service rate of the third-time repairability test stations,
- N_{rp1} : number of components that are repaired for the first time,
- N_{rp2} : number of components that are repaired for the second time,
- N_{rp3} : number of components that are repaired for the third time,
- N_{rm} : number of remanufactured components,
- N_{nw} : number of purchased components,
- s_{rp1} : buffer of the first time repaired components,
- s_{rp2} : buffer of the second time repaired components,
- s_{rp3} : buffer of the third time repaired components,
- s_{rm} : buffer of remanufactured components,
- s_{nw} : buffer of purchased components.

3.3 Decision variables and key indicators

This study aimed at studying the CMW performance based on two main indicators of time and cost. From the time point of view, the system's agility in providing the required service to the failed units is evaluated by measuring the equipment's average waiting time (T_w). To assess the cost, the best indicator that can measure the system's efficiency is the average cost spent to maintain production line machines. The last indicator can be calculated by dividing the total CMW cost by the number of machines, 100 in this study.

Defined indicators must be studied under different scenarios and with various values of decision variables. Repairability rate (r_{p1} , r_{p2} , r_{p3}), remanufacturing rate (r), maximum allowed number of repairing for each component, and initial inventory level (B) are considered as decision variables. Various combinations of these variables will enable us to evaluate CMW's performance under different circumstances.

4 EXPERIMENTS

Experiments have been conducted based on the data from the automotive industry. We used the Arena simulation tool to study the CMW performance by analysing the variation of key indicators. Obviously to study the system accurately, the simulation length was crucial. The length of simulation is 2400 days, which is equal to almost 10 years to obtain reliable and accurate results. Using several combinations of decision variables leads to a more accurate evaluation of system behaviours and performances.

To validate the proposed model, we ran a simplified version of current model to compare its output with the results reported by (Djeunang Mezafack et al., 2020) for an open-loop CMW. The same results were obtained which validates the accuracy and efficiency of the current model.

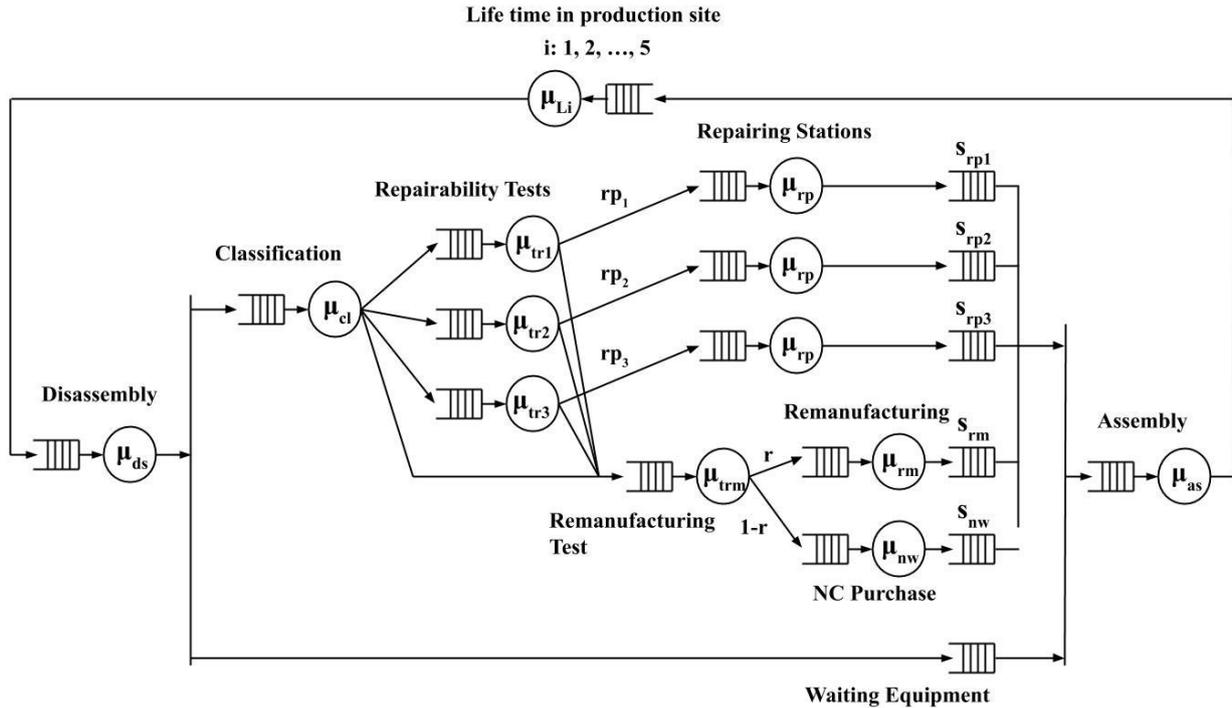
4.1 Experiments' characteristics

We have considered three different groups of repairability rate, which is shown in Table 2. In Group 1, 90 percent of components visiting the repairing station for the first time, are

repairable. 70 percent of components which have been repaired one time, are repairable. 50 percent of parts with two times previous repairs, can be repaired for third time. The remaining components can be processed in other stations (NC purchase or remanufacturing). The repairing rates of Group2 are 80, 60,

- $T_{rp1} = T_{rp2} = T_{rp3} = 5$ days (1 week)
- $T_{rm} = 30$ days (6 weeks)
- $T_{nw} = 40$ days (8 weeks)

Based on a real case in the automotive industry, the lifetime of components processed in different stations are reported as



and 40 percent for first, second, and third time repairs, respectively. The repairing rates for Group 3 are 70, 50 and 30 percent.

For each group in Table 2 we performed several experiments by changing the value of remanufacturing rate (r : 0, 0.1, 0.2, 0.4, 0.6, 0.8, 1).

Table 2. Repairability rate

	Repairability rate (r_{p1}, r_{p2}, r_{p3})
Group 1	(0.9, 0.7, 0.5)
Group 2	(0.8, 0.6, 0.4)
Group 3	(0.7, 0.5, 0.3)

Various initial inventory level (B) is considered to analyse the influence of this variable on the CMW performance efficiency: B : 0, 2, 5, 10, 15, 20, 25

We defined three scenarios for the repairing process:

1. Scenario 1: each component can be repaired a maximum of three times,
2. Scenario 2: each component can be repaired a maximum of two times,
3. Scenario 3: each component can be repaired one time.

In case all types of components are available, the priority of use will be 1- new component, 2- remanufactured, 3- first-time repair, 4- second-time repair, 5- third-time repair.

We performed 441 simulations to study all possible combinations of the variables.

The components' first visit follows a Poisson distribution with rate (λ) of 1 equipment per week (5 days).

Other stations service times are as follows:

- $T_{ds} = T_{as} = 1$ day

follows:

- $T_{L1} = 160$ days
- $T_{L2} = 120$ days
- $T_{L3} = 80$ days
- $T_{L4} = 360$ days
- $T_{L5} = 480$ days

4.2 Cost function

One of the key indicators of the CMW is the average cost spent on maintenance of each machine over a period of time. To be able to calculate the key indicator, the total cost of the CMW must be calculated. Different methods can be used to calculate the cost. In this study, Total cost is the sum of the waiting cost of the equipment in the CMW and the operational cost of the CMW itself.

$$C_{total} = C_{waiting} + C_{operation}$$

$$C_{waiting} = T_w * C_{ew} * (N_{rp1} + N_{rp2} + N_{rp3} + N_{rm} + N_{nw})$$

C_{ew} represents the cost of waiting equipment per day, which can be calculated with different methods. One most common practice is to consider the opportunity cost. We assumed that the presence of a failed unit in CMW is the same as the absence of a machining centre in the production line. Thus, C_{ew} will be equal to the revenue that the producer gains in a day when they use their machine in the production line, and it is 12 EUR/day.

Operational cost can be calculated by:

$$C_{operation} = C_{rp1} * N_{rp1} + C_{rp2} * N_{rp2} + C_{rp3} * N_{rp3} + C_{rm} * N_{rm} + C_{nw} * N_{nw}$$

- where:
- C_{rp1} : First time repairing unit cost (200\$)
 - C_{rp2} : Second time repairing unit cost (250\$)
 - C_{rp3} : Third time repairing unit cost (300\$)
 - C_{rm} : Remanufactured component unit price (600\$)
 - C_{nw} : New component unit price (900\$)

The costs were estimated based on our real case in the automotive industry.

Since the components are small and do not take a lot of space in the warehouse, the inventory cost has been neglected.

5 RESULTS

Having our model constructed, we simulated 441 different combinations of the decision variables to precisely study the performance of the CMW for ten years. By analysing the key indicators, we can obtain the optimum combination of the decision variables introduced in section 3.

5.1 Average waiting time of equipment (T_w)

Average waiting time as a key indicator represents the agility of the CMW in processing the visiting failed units. We measured the variation of T_w under various scenarios by choosing several combinations of variables explained in detail in section 4.

Globally we found that T_w decreases as we increase the initial inventory level up to B equal to 20. As it is shown in Figure 4, the deduction rate of T_w decreases while the value of B increases. Increasing the initial inventory level from 20 to 25 has no impact on the amount of T_w . This trend is true for all other combinations of the decision variables.

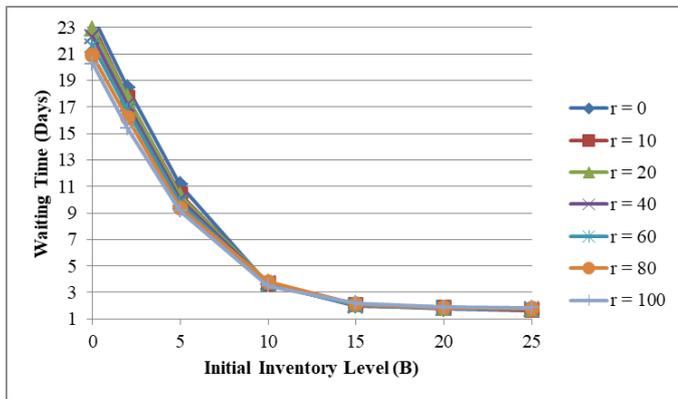


Figure 4. T_w versus B for Group 2 and maximum two times repairing

For the initial inventory level of 20, the simulation results of all combinations of the maximum allowed number of repairing and repairability rates show that the 1-time repair is the most time-efficient option. Therefore, from time point of view, it is better to repair the components only once and in future visits the failed components should be processed in the remanufacturing or NC purchase stations.

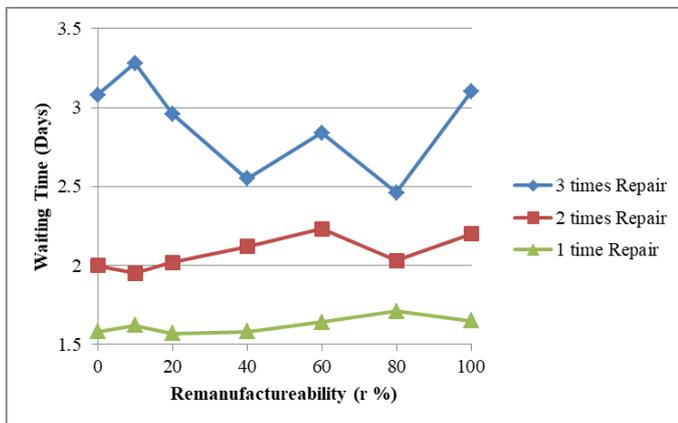


Figure 5. T_w versus r for Group 1 and $B = 20$

Depending on the repairability rate, the optimum value for remanufacturing is different. In Figure 5, it is shown that for group 1, the optimum remanufacturing rate is 0.2.

For other combinations of repairability rates, the optimum values of r are 0.1 for group 2 and 0.4 for group 3. According to Figures 6 and 7, after the optimum point, the increase of remanufacturing rate increases the T_w .

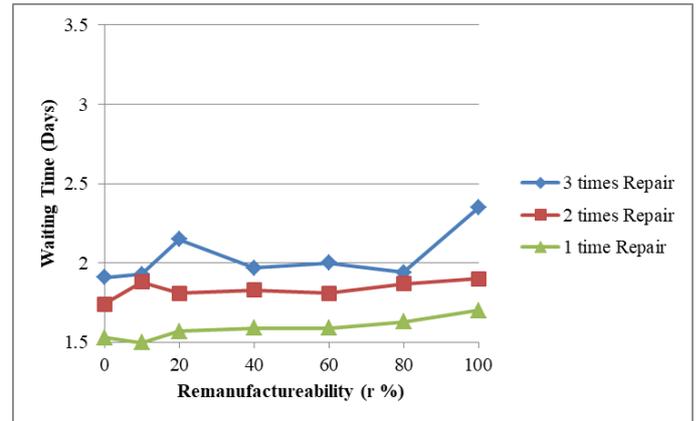


Figure 6. T_w versus r for Group 2 and $B = 20$

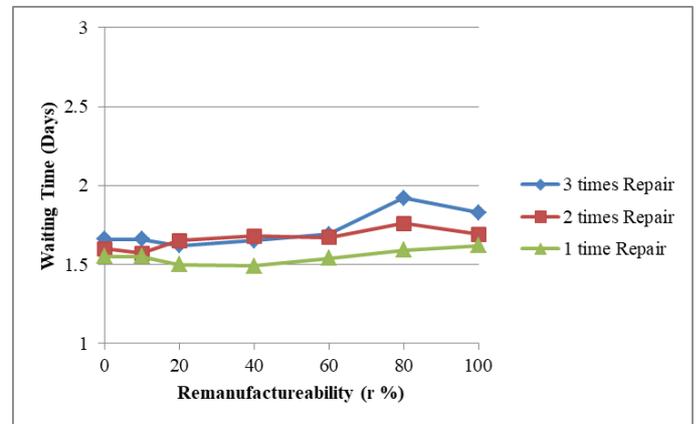


Figure 7. T_w versus r for Group 3 and $B = 20$

5.2 10-year maintenance cost per machine (C_{10})

10-year maintenance cost per machine as a key indicator represents the cost efficiency of the CMW in processing the visiting failed units. We measured the variation of C_{10} under various scenarios by choosing several combinations of variables that are explained in detail in section 4.

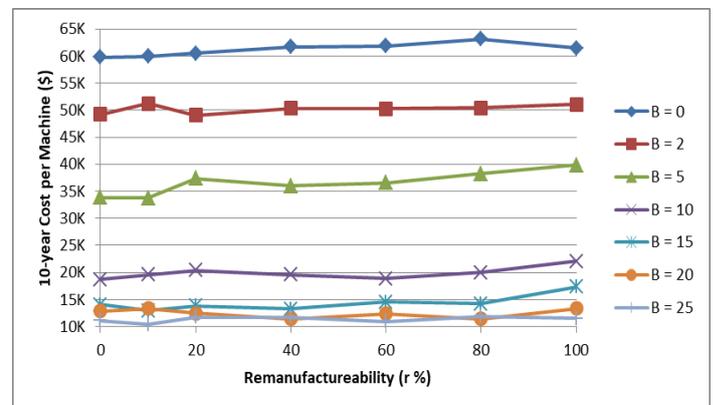


Figure 8. C_{10} versus r for Group 1 and maximum three times repairing

The number of components in the initial inventory has a significant impact on the maintenance cost per machine. Generally, C_{10} decreases by an increase in the initial inventory level. Figure 8 indicates how the value of C_{10} changes versus the remanufacturing rates for different initial inventory levels. For groups 2 and 3 and other maximum allowed number of repairing, the same variation trend is observed. The deduction rate of C_{10} decreases while the value of B increases. The deduction rate of C_{10} is the slope of the curve shown in Figure 9.

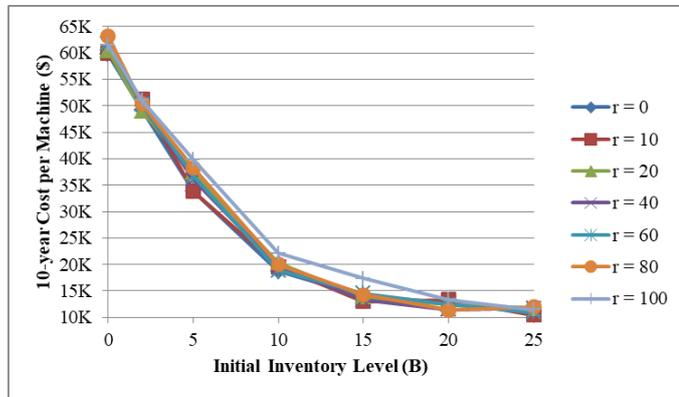


Figure 9. C_{10} versus B for Group 1 and maximum three times repairing

Top figures show that CMW is the most cost-efficient when the initial inventory level is 25. Analyzing the simulation results for $B = 25$ indicates that the system is more cost-efficient if the component is repaired only once. Figure 10 shows the variation of C_{10} versus remanufacturing rate (r) when B is equal to 20, and only one-time repair is allowed. Different curves in the figure indicate different repairability rates (Group1, 2, and 3).

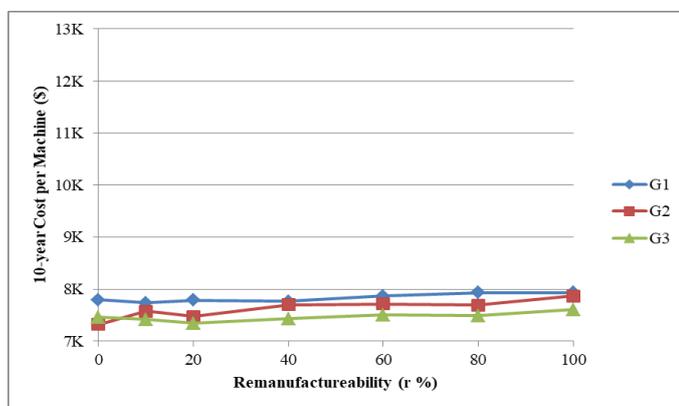


Figure 10. C_{10} versus r for one-time repair and $B = 20$

For Group 1, the optimum remanufacturing rate is 0.1. Remanufacturing is not cost-efficient for Group 2, which means that the optimum value of r is equal to 0. A remanufacturing rate of 0.2 is the optimum point for Group 3.

6 CONCLUSION

In this study, we measured the performance of CMW by evaluating the average waiting time of equipment in the system and 10-year maintenance cost per machine.

The impact of various combinations of initial inventory level, maximum allowed number of repairing, remanufacturing, and repairing rates were studied.

It is concluded that using the remanufacturing process can positively affect the CMW performances from the time and cost perspectives.

We also found that there is not a unique point that we can optimize both indicators of the system. But the trends of the variation of the indicators were similar because T_w had a significant effect on the cost function. The simulation results show that, although the optimum point for T_w and C_{10} are not the same, if a company chooses one of these options, it can be sure that the other indicator is in an acceptable range not far from the optimal.

The impact of an increase in initial inventory level is more in the lower values of B. For the more significant amounts of B, the initial inventory level's growth has less impact on CMW performance.

For an initial inventory level of less than 10, a maximum of the three-time repair was the optimal option. By increasing the value of B, the maximum one time repair became more efficient.

The repairing rate of the stations has a significant effect on the optimal remanufacturing rate. Also, the maximum number of allowed repairs had a considerable impact on the optimal remanufacturing rate.

For further research, different cost calculation scenarios can be considered. The transportation time between the production lines and the CMW can be taken into account. The priority of the components according to the maintenance process can be studied.

ACKNOWLEDGEMENT

This work has been supported by the French government; through the Université Grenoble Alpes and the IDEX Grenoble (Cross Disciplinary Program Circular) with the reference number C7H-ID17-PR28-CIRCULARWP1.

REFERENCES

- Driessen, M. A., van Houtum G. J., Zijm W. H. M., & Rustenburg W. D. (2020). Capacity assignment in repair shops with high material uncertainty. *International Journal of Production Economics*, 221.
- Sleptchenko, H. H., A. Turan, S. Pokharel, and T. Y. ElMekkawy, 2019. Cross-training policies for repair shops with spare part inventories. *International Journal of Production Economics*, 209, 334–345.
- Popa, V. N., & , Popa L. I. (2016). Green Acquisitions and lifecycle management of industrial products in the circular economy. *IOP Conference Series: Materials Science and Engineering* (Vol. 161).
- Simeu-Abazi, Z., Di Mascolo, M., & Gascard, E. (2014). Queuing network-based methodology for designing and assessing performance of centralized maintenance workshops. *Journal of Manufacturing Technology Management*, 25(4), 510–527.
- Kleber, R., Zanoni, S., & Zavanella, L. (2011). On how buyback and remanufacturing strategies affect the profitability of spare parts supply chains. *International Journal of Production Economics*, 133(1), 135–142.
- Asadi, M., Mobin, M., Cheraghi, S., & Li, Z. (2018). An algorithm for performance evaluation of closed-loop spare supply systems with generally distributed failure and repair times. *IEEE Transactions on Reliability*, 99, 1–11.
- Djeunang Mezafack, R., Di Mascolo, M., Simeu-Abazi, Z., & Akbari Haghghi, S. (2020). Performance

evaluation of centralized maintenance workshop in the circular economy context. *13th International Conference on Modeling, Optimization and Simulation, MOSIM 20*.

Li, B., Li, X., Guo, W., & Wu, S. (2013). A generalized stochastic Petri-net model for performance analysis and allocation optimization of a particular repair system. *Asia-Pacific Journal of Operational Research*, 30(1), 1250042.

(This paper is not included in this version, please contact authors for more information)

Session 1.4 : Design et Pilotage des systèmes de production de biens et de services

Chairman : Martin Trépanier - Yannick Frein

CIGI 2021

Redesign du mode de pilotage d'une chaîne logistique minière

Bassma AZZAMOURI^{1,2}, Vincent HOVELAQUE^{1,2}

¹EMINES-SCHOOL OF INDUSTRIAL MANAGEMENT, UNIVERSITE MOHAMMED VI POLYTECHNIQUE,
43150 Ben-Guerir, Maroc
{bassma.azzamouri, vincent.hovelaque }@emines.um6p.ma

²UNIV RENNES 1, CNRS, CREM - UMR 6211,
F-35000 Rennes, France
bassma.azzamouri@etudiant.univ-rennes1.fr, vincent.hovelaque@univ-rennes1.fr

Résumé – Nous adressons une problématique de redesign du mode de pilotage d'une chaîne logistique minière afin d'améliorer la satisfaction du client en termes de quantité, de qualité définie par les caractéristiques d'un ensemble d'éléments chimiques et de délai de livraison et d'optimiser la mobilisation des ressources. La décision intégrée consiste à piloter le flux de matière tout au long de la chaîne en tenant compte du comportement de tous les éléments chimiques caractérisant d'abord la matière première, passant par le produit intermédiaire, pour arriver au produit final, face à différents paramètres tels que : le choix entre les stocks de la chaîne logistique, le traitement du produit et la constitution des mélanges. L'évaluation de l'impact de cette décision est réalisée à travers un modèle mathématique linéaire de planification de la production, afin de satisfaire le client par la quantité et qualité demandée, en choisissant des configurations de production avec moins de ressources. Le modèle donne comme résultats, les flots de produits et les itinéraires définis par le chemin de production, y compris le stock de provenance, les entités à traverser, la gamme de traitement à adopter pour chaque ingrédient et les mélanges à effectuer pour chaque produit fini.

Abstract – We are addressing an issue of a redesign of the management mode of a phosphate production mining supply chain to improve customer satisfaction in terms of quantity, quality defined by the characteristics of chemical components, and delivery time and to optimize the mobilization of resources. The integrated decision consists in controlling the flow of material throughout the chain, taking into account the behavior of all the chemical components characterizing the raw material, the intermediate and the final product, and the impact of different parameters such as the choice between the stocks of the supply chain, the choice between different treatment routings and the mixtures to constitute. The evaluation of the impact of this decision is carried out through a linear mathematical model of production planning, to satisfy the customer with the best quality requested, by choosing production configurations with fewer resources. The model gives as results, the product flows and the routes defined by the production path, including the entity which is the source of each product and the entities to be crossed, the treatment routings to be adopted for each ingredient and the mixtures or blendings to be constituted for each finished product.

Mots clés - Redesign, chaîne logistique minière, configurations de production, gammes de traitement, constitution des mélanges.

Keywords – Redesign, mining supply chain, production configurations, treatment routings, blending.

1 INTRODUCTION

Chaque industrie manufacturière vise à satisfaire au mieux ses clients et à mettre à leurs dispositions une large diversité de produits. De leurs côtés, les clients deviennent de plus en plus exigeants. C'est pour cette raison que chaque entreprise procède à l'intégration de nouvelles lignes de production, nouveaux moyens de transport, recettes de fabrication améliorées voire de nouveaux produits conçus avec beaucoup de créativité. L'objectif principal est de répondre au mieux aux exigences de ses clients et de garder/augmenter sa part de marché. L'entreprise tient compte aussi des différentes ressources disponibles pouvant être mobilisées, dans le but de minimiser les dépenses et de garantir la rentabilité financière souhaitée [Lim et al, 2020]. Elle se trouve donc, parfois, face à l'obligation d'améliorer le mode de gestion au sein de son processus de production et de s'adapter aux améliorations qu'elle désire intégrer. Il s'avère que le design d'une nouvelle chaîne logistique répondant à des critères sophistiqués, est

coûteux et nécessite de grands investissements financiers et organisationnels. Les bénéfices de ces derniers pourront être obtenus lentement. La solution alternative reste donc pour chaque entreprise celle d'adopter le concept du redesign du système productif actuel en gardant les mêmes infrastructures. Ce concept consiste à intégrer une décision qui change ou modifie le mode de pilotage actuel et d'évaluer son impact pour des objectifs bien particuliers [Persson, 1995]. Ces derniers tiennent généralement compte de : la productivité de l'entreprise, l'amélioration de la performance, la satisfaction des clients, les coûts de production et de transport, la gestion des stocks, la gestion du temps de production et de mise à disposition du produit aux clients (Lead time) [Persson, 1995] ainsi que les aspects environnementaux, liés au développement durable (par exemple de la gestion efficace des déchets alimentaires dans le cadre d'une chaîne de distribution efficace d'aides alimentaires [Martins et al, 2018]).

Le présent travail s'inscrit dans cette veine du redesign et s'applique à l'un des sites d'extraction et de traitement du phosphate nommé Khouribga du groupe OCP. Ce dernier est leader sur le marché international du phosphate et de ses dérivés, expédiant ses produits à travers le monde via un réseau de plus de 150 clients. Les commandes satisfaites au niveau de la chaîne logistique étudiée sont dédiées à deux types de clients : *i*) client local qui est l'unité chimique de production des engrais et d'acide phosphorique, pour lequel la commande est transportée sous forme de pulpe via pipeline *ii*) clients à l'export pour lesquelles les commandes sont expédiées sous forme de roche phosphatée et acheminées via trains vers le port. Nous allons nous focaliser sur la partie de la chaîne concernée par la fabrication des produits exportés.

La chaîne logistique étudiée est considérée comme un réseau de production composé de plusieurs entités (nœuds) : d'extraction, d'enrichissement (ou de traitement) et de séchage. Ces entités sont liées entre elles via des convoyeurs, permettant l'acheminement des produits de différentes natures (brutes et lavées). La satisfaction des clients est traduite par le respect du triplet quantité, qualité et délai. Les exigences en termes de qualité sont traduites par une charte client définie par des plages de valeurs avec des bornes supérieures et inférieures d'un ensemble d'éléments chimiques (BPL, SiO₂, MgO, Cd, CO₂, Al₂O₃, Fe₂O₃,...). La spécificité du présent contexte industriel est l'hétérogénéité des matières premières qui peuvent être utilisées pour la fabrication des produits finis. De plus, elles se trouvent sur la chaîne sous différentes formes : les mines fournissent le produit brut nommé qualité source (QS), les laveries procèdent à l'enrichissement de ces QS pour fournir des qualités sources intermédiaires ou lavées nommées QSL. C'est au niveau de l'unité de séchage que les mélanges nécessaires pour préparer les qualités marchandes (QM) demandées par les clients, sont constitués. Il existe donc une coordination entre les entités de la chaîne dans le cadre de la satisfaction des commandes clients.

Puisque les volumes de matière première généralement disponibles sont supposés suffisants pour répondre aux clients, le souci principal reste donc celui de la qualité du produit fini caractérisée par les teneurs en différents éléments chimiques. Les produits extraits et présents en amont peuvent ne pas avoir les caractéristiques chimiques respectant la charte client et ils ont besoin d'une série d'opérations d'enrichissements et de mélanges pour atteindre l'objectif souhaité. Il est rare qu'un simple enrichissement par un procédé de lavage suffise à satisfaire le client ; c'est pour cette raison que l'obtention d'une QM passe par un mélange de minerais.

Les flots de produits ne doivent pas être figés par QM et doivent faire appel à des études régulières d'optimisation pour les choix à effectuer en termes des quantités d'intrants et des itinéraires définis par la gamme opératoire de transformation (ou de non transformation). Le pilotage de la chaîne logistique étudiée se fait en se basant principalement sur le BPL et en tenant compte des conséquences liées au comportement des autres éléments chimiques. Cependant, pour garantir une satisfaction totale de tous les composants chimiques de la charte client, nous proposons de remonter en amont de la chaîne et d'intégrer tous les éléments chimiques dans sa gestion et son pilotage depuis la sortie des mines, en passant par les différentes opérations, jusqu'à l'arrivée au dernier stade. C'est pour cette raison, que nous nous intéressons à un travail de redesign du mode de pilotage de la chaîne logistique étudiée, afin d'améliorer la satisfaction des clients en mobilisant moins de ressources. Dans cette étude, le redesign ne consiste pas à une remise en cause des structures physiques existantes mais à une redéfinition du

mode de pilotage de l'ensemble. Ceci se réalisera à travers l'intégration d'une décision de type tactique basée sur le pilotage des flux et d'évaluer son impact sur l'amélioration de la satisfaction du client sur un horizon court terme. Pour chaque QM, les flots de produits sont à définir en fonction de la disponibilité des QS et des QSL, des ressources à mobiliser et des spécificités de chaque commande. Ces flots seront dynamiques car variant selon chaque commande et selon la nature et la quantité des QS livrées par les mines. Une flexibilité est garantie par le nouveau mode de gestion, via la proposition de solutions alternatives en cas d'indisponibilité d'une ou de plusieurs ressources ou de manque de QS. Le schéma ci-dessous est une illustration de la chaîne logistique étudiée :

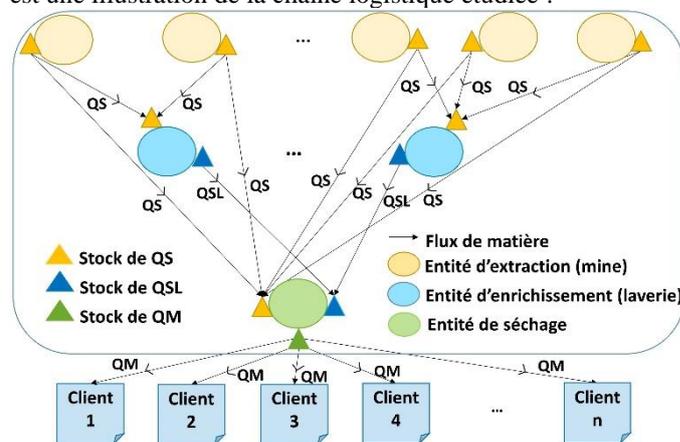


Figure 1. Réseau de production et d'acheminements possibles du produit sur la chaîne logistique étudiée

Le présent papier est organisé comme suit : la section 2 est consacrée à une revue de littérature au sujet du redesign des chaînes logistiques ; la section 3 décrit les hypothèses et les règles de gestion ainsi que le modèle mathématique. La section 4 est une illustration numérique des résultats obtenus par l'exécution du modèle. L'article s'achève par une brève conclusion du travail réalisé.

2 REVUE DE LITTÉRATURE ET CHAMP EMPIRIQUE D'ANALYSE

Le redesign d'une chaîne logistique est depuis plus d'une vingtaine d'années [Persson, 1995] sujet de nombreux travaux de recherche menés théoriquement ou par étude de cas réels. Il consiste à supposer qu'une chaîne logistique existe déjà, de prendre en considération ses infrastructures actuelles et de se concentrer sur des décisions de reconfiguration des activités de production ou des modes de gestion et en tirer des bénéfices [Hammami et Frein, 2014]. Les décisions sont prises à un niveau stratégique/tactique et évaluées au niveau opérationnel. Le redesign a été traité dans plusieurs industries avec des objectifs qui diffèrent selon le domaine d'étude et la nature des produits traités. [Persson, 1995] a étudié l'impact du redesign sur l'amélioration de la productivité et de la performance de n'importe quel processus productif à travers l'intégration de nouveaux concepts et stratégies. Une année plus tard, [Berry et Naim, 1996], ont étudié l'impact de l'intégration d'une stratégie de redesign sur une grande industrie européenne de production d'ordinateurs personnels. Sur les 25 années de recherche précédentes, le concept du redesign a été appliqué à travers le développement de programmes mathématiques et d'outils, notamment dans l'agroalimentaire [Vorst et al, 2008], [Martins et al, 2018], cimenterie [Jahani et al, 2018], pharmaceutique [Cortinhal et al, 2015], agriculture [Feito-Cespon et al, 2017], [Berry et Naim, 1995], [Krishnan et al, 2019] ainsi que la production de nouvelles technologies [Martins et al, 2017].

L'analyse des travaux menés sur ces différents domaines, permet de les classer selon les critères d'analyse suivants :

- *Niveau de la chaîne logistique* : Le Redesign peut être appliqué aux trois niveaux de la chaîne logistique, i) approvisionnement, à travers le redesign des zones de réception des produits approvisionnés en réaffectant les fournisseurs aux quais et le stockage des pièces reçues aux différents espaces de stockage [Saez-Mas et al, 2020] ii) réseau de production, traité par [Berry et Naim, 1996] ou ii) au réseau de distribution des produits intermédiaires ou finaux, à travers non seulement la définition des chemins de distribution optimaux mais aussi les choix d'investissement par rapport à l'ouverture ou la fermeture d'entrepôts de stockage, traité par [Jahani et al, 2018], [Vorst et al, 2008], [Martins et al., 2018], [Lim et al, 2020].

- *Type de décision* : le Redesign est une décision de type tactique voire stratégique. Elle sert à intégrer des améliorations sur i) un horizon moyen et long terme [Melo et al., 2011], [Martins et al., 2017], [Jahani et al., 2018]. L'impact du redesign est souvent évalué sur ii) le court terme dans le but de vérifier le degré de robustesse et la pertinence de la décision prise sur le quotidien opérationnel, le cas de : [Martins et al, 2007]

- *Mode de résolution* : les problématiques de type redesign d'une chaîne logistique productive font appel à différentes méthodes de résolution, à travers : i) des programmes d'optimisation traités par [Jahani et al, 2018], [Martins et al, 2018],[Feito-Cespon et al, 2017], [Razmi et al, 2013], [Razmi et al, 2019],[Lim et al,2020] ii) des modèles de simulation : développés par [Vorst et al, 2008], [Pissonnier et al, 2019] ou iii) Une combinaison entre l'optimisation et la simulation traitées par : [Martins et al, 2017], des frameworks Personnalisés traités par :[Krishnan et al, 2019]

- *Objectif de modélisation* : La majorité des papiers analysés sont basés sur une combinaison d'objectifs, notamment : l'objectif économique traduit par une minimisation des coûts opératoires, l'objectif environnemental ainsi que l'objectif de satisfaction client. Ceci est traité par : [Vorst et al, 2008],[Feito-Cespon et al, 2017],[Martins et al,2018]. D'autres chercheurs se sont focalisés sur un seul objectif, économique par exemple, tel que le cas de [Jahani et al, 2018], [Melo et al, 2011], [Cortinhal et al,2015], [Lim et al,2020], un objectif environnemental qui constitue une indication de l'inefficacité des opérations et de l'utilisation des ressources, ceci a été traité par [Krishnan et al, 2019]. Tandis que pour un ensemble d'autres auteurs, le souci principal était traduit par la satisfaction client à moindre coût, comme c'est le cas pour [Razmi et al, 2013], [Martins et al, 2017].

A notre connaissance, le redesign des chaînes logistiques productives a eu moins d'attention dans le secteur minier. Dans un travail effectué sur l'axe Centre du groupe OCP, [Azzamouri et al, 2020] ont proposé un modèle d'optimisation des mélanges à utiliser en intégrant la gestion des stocks de sécurité, de la qualité des QM et des gammes de traitement. Cependant, nous traitons ici un problème spécifique de redesign d'un réseau logistique de production minière sous l'objectif de satisfaire la demande des clients tout en minimisant la mobilisation des ressources. Il permet également de considérer la combinaison originale des éléments suivants, qui n'a pas été traité dans la littérature :

i) Le redesign est effectué sur un réseau de production (et non de distribution auquel il est généralement associé dans la littérature).

ii) Il concerne non seulement le produit final, mais plusieurs natures de produits depuis la matière première (QS) en passant par des produits intermédiaires (QSL) pour arriver au produit final (QM) demandé par le client.

iii) Les caractéristiques des intrants sont évolutives au fil du temps.

iv) Les flots de production sont dynamiques au fil du temps en tenant compte des infrastructures disponibles.

v) Le redesign est effectué sur un horizon tactique évalué sur le quotidien opérationnel.

vi) La constitution des mélanges fait partie du redesign à effectuer à travers le choix d'un mélange optimal à adopter pour la préparation des qualités marchandes.

vii) Le produit final peut être préparé à partir de la matière première seulement, ou de la matière intermédiaire seulement ou d'une combinaison de matières premières et de produits intermédiaires, tout dépend de la disponibilité de ces deux types de produits et de la demande à satisfaire.

viii) L'utilisation d'une seule formule dans le cadre de la fonction objectif pour gérer implicitement et en même temps : les délais de préparation des QM, la mobilisation des ressources, la priorisation des configurations de production et la priorisation de la consommation des QS.

ix) La gestion des commandes des clients est impactée par un ensemble de paramètres : nature des gisements, choix des configurations de production ainsi que les mélange à effectuer, ces paramètres sont à tenir en compte d'une façon complémentaire l'une à l'autre.

3 MODELISATION DE LA CHAINE

3.1 Hypothèses de modélisation

Le présent travail est une étude de cas qui prend en compte l'ensemble des infrastructures composant la chaîne logistique minière étudiée pour travailler sur le redesign souhaité. La complexité du terrain étudié rend nécessaire l'intégration d'un ensemble d'hypothèses de modélisation afin de considérer les détails liés au mode de gestion adopté. En complément des caractéristiques décrites précédemment, quelques hypothèses sont proposées ci-dessous :

3.1.1 Gammes de production

L'ensemble des opérations composant le processus étudié permet de proposer la notion de « gammes de traitement ». Ces derniers sont au nombre de quatre gammes à adopter dans le présent travail de modélisation, gamme 1 : aucun enrichissement + séchage, gamme 2 : Lavage simple + séchage, gamme 3 : Lavage et flottation+ séchage, gamme 4 : Lavage, broyage et flottation+ séchage.

3.1.2 Configurations de production

Différents chemins de production (ou configurations) sont possibles. Leurs choix dépendent principalement des QS disponibles sur les stocks et de la QM à satisfaire. Généralement, cinq configurations de production existent et se représentent par le couple (i,h) qui détermine l'intrant choisi et par conséquent le stock mobilisé ainsi que la gamme de traitement h. Généralement, si la QS ne nécessite pas d'enrichissement, elle s'achemine directement vers l'unité de séchage. Concernant les QS pauvres, le passage par la laverie s'avère nécessaire. Ceci permet d'obtenir un produit semi-fini ou intermédiaire. Au niveau de l'unité de séchage, s'effectuent les mélanges nécessaires pour satisfaire le client en termes de quantité et de qualité de la roche phosphatée. Ceci se poursuit par le séchage du mélange avant le chargement des trains par le produit final. Ceci est décrit sur le schéma en ci-dessous :

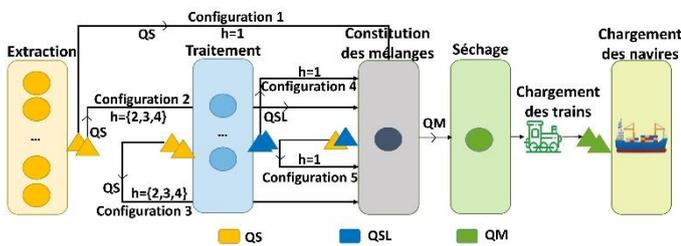


Figure 2. Illustration des configurations de production

3.1.3 Gestion temporelle par configuration

Chaque configuration (i,h) se caractérise par son temps opératoire, vu qu'elle résulte d'un passage par une ou plusieurs entités. Le produit au cours de sa préparation passe généralement par trois principales opérations : le traitement, l'acheminement et le stockage. Chacune de ses opérations prend un temps qui varie entre quelques minutes, quelques heures et quelques jours (respectivement). Il s'avère donc nécessaire de trouver l'unité temporelle exacte à adopter lors de la modélisation de la chaîne et qui s'accorde avec les temps nécessaires à chaque opération. De plus, les délais au plus tôt et au plus tard de livraison des différentes qualités marchandes peuvent se superposer, d'où un chevauchement des périodes sur lesquelles ces QM doivent être préparées. Il est donc judicieux de mettre en place des hypothèses simplificatrices pour déterminer les périodes sur lesquelles les intrants seront traités. Prenons l'exemple de la deuxième configuration qui consiste d'abord à acheminer le produit brut depuis la mine vers l'unité de traitement (laverie), puis de le traiter au niveau de la laverie, ensuite de stocker le produit lavé pour éliminer les volumes d'eau supplémentaires et enfin de l'acheminer vers l'unité de séchage. Nous supposons donc que chacune de ces quatre opérations soit effectuée sur des périodes différentes mais successives et en respectant la contrainte liée aux débits maximaux d'alimentation des ressources mobilisées. Donc la première opération qui consiste à acheminer les QS depuis les mines vers l'unité de traitement doit se réaliser sur la période t. Le traitement des QS et le stockage des QSL obtenues doivent s'effectuer sur la période suivante t+1 et l'acheminement des QSL vers l'unité de séchage est à réaliser sur la période t+2.

3.2 Intrants et extrants du modèle

Le présent modèle mathématique a pour but d'évaluer l'impact du pilotage de la chaîne en tenant compte du comportement de chaque composant chimique caractérisant le produit, sur l'amélioration de la satisfaction des clients. Les données d'entrée utilisées tiennent compte d'une part des commandes des clients définies par quantité, qualité et délai de livraison des produits finis et d'autre part des infrastructures mobilisées. Concernant ces derniers, il s'agit généralement des capacités des ressources (laveries et convoyeurs) et de leurs disponibilités par période, de l'état initial des stocks, des alimentations prévues en QS par les mines, définies par la qualité et la quantité des QS, des temps opératoires nécessaires par configuration. Pour la gestion de la qualité, les rendements liés au traitement par les entités d'enrichissement sont pris en compte.

Quant aux extrants du modèle, ce dernier permet de déterminer les flots de produits nécessaires par QM. Chaque flot lui-même caractérisé par une quantité d'ingrédients (QS ou QSL) et un chemin de production défini par passage ou non par les entités d'enrichissement.

3.3 Formulation mathématique

Le présent travail est un modèle mathématique développé en utilisant un ensemble d'indices, paramètres, variables et contraintes, traduisant la chaîne logistique traitée. Le modèle d'optimisation a pour objectif la satisfaction d'un carnet de commande (quantité, délai, qualité) à travers une fonction de pénalité. Les tableaux ci-dessous illustrent les notations utilisées dans le modèle avec leurs descriptions :

Tableau 1. Notations utilisées

Indices	Descriptions
t	Indice de temps (t=0..T), t couvrant un mois.
i	Indice des intrants (i=1..I)
k	Indice des commandes (k=1..K)
c	Indice des composants chimiques (c=1..C)
h	Indice des gammes (h=1..H)
r	Indice des ressources (r=1..R)
Paramètres	Descriptions
$\alpha_{c,i}$	Poids du composant c du minerai i
S_i^{init}	Quantité initiale de l'intrant i à t=0
$A_{i,t}$	Quantité approvisionnée par intrant i sur la période t
$\omega_{c,i,h}$	Coefficient d'enrichissement du composant c par minerai i, lors du passage par la gamme de traitement h.
$\eta_{i,h}$	Coefficient de réduction du poids de l'intrant i lors du passage par la gamme de traitement h
$\beta_{c,k}^{min}$	Borne inférieure (%) du composant c de la commande k
$\beta_{c,k}^{max}$	Borne supérieure (%) du composant c de la commande k
$\tau_{c,k}$	Valeur cible (%) du composant c de l'extrait appartenant à la commande k
σ_c	Pénalité de l'écart par rapport aux valeurs cibles des composants c
$\sigma_{i,h}$	Pénalité lié au choix de la configuration (i,h)
D_k	Quantité demandée de l'extrait λ_k
$U_{i,h}$	Paramètre binaire=1 si l'intrant i admet d'être traité par la gamme de traitement h, sinon 0
$U_{i,h,r}$	Paramètre binaire=1 si l'intrant i passant par la gamme h admet r comme ressource pour son traitement, sinon 0
$g_{i,h}$	Temps moyen par configuration, défini par intrant et par gamme de traitement
$g_{i,h,r}$	Période de mobilisation des ressources pour chaque configuration (i,h)
ρ_r	Débit journalier d'alimentation par ressource
$\phi_{r,t}$	Taux de disponibilité journalier par ressource
E_k	Date de préparation au plus tôt de la commande k
L_k	Date de préparation au plus tard de la commande k
N	Quantité minimale d'intrant à consommer
M	Quantité maximale d'intrant à consommer
Variables	Descriptions
$x_{i,h,k,t}$	Variable de décision – quantité à prélever du minerai i, traitée par la gamme h pour produire la commande k durant la période t
$O_{i,h,k,t}$	Variable de décision - binaire valant 1 si le minerai i est traité par la gamme de traitement

	h pendant la période t pour produire la commande k	$\forall i \in \{QS_{entrée_laverie}\}, t$
$y_{i,h,k,t}$	Variable intermédiaire- le quotient des quantités $x_{i,h,k,t}$ par la quantité M maximale à consommer d'un intrant	Avec : $\{QS_{entrée_laverie}\} = \{QS_{entrée_laverie-1}\} \cap \dots \cap \{QS_{entrée_laverie-n}\}$ (16) $S_{i,t} = S_{i,t-1} - \sum_{h,k U_{i,h}=1 \cap h \in \{2,3,4\}} x_{i,h,k,t}$
$z_{i,h,k,t}$	Variable intermédiaire- le quotient des quantités $x_{i,h,k,t}$ sur la quantité N minimale à consommer d'un intrant	$\forall i \in \{QSL_{Sortie_laverie}\}, t$ Avec : $\{QSL_{Sortie_laverie}\} = \{QSL_{Sortie_laverie-1}\} \cap \dots \cap \{QSL_{Sortie_laverie-n}\}$ (17)
$S_{i,t}$	Variable intermédiaire- Stock de l'intrant i à la fin de la période t	$S_{i,t} = S_{i,t-1} - \sum_{h,k U_{i,h}=1, h=1} x_{i,h,k,t}$
$g_{c,k}$	Variable intermédiaire - poids intermédiaire du composant c de l'extrait de l'ordre k	$\forall i \in \{QS_QSL_séchage\}, t$ $S_{i,t} = S_{i,t-1} - \sum_{h,k U_{i,h}=1 \cap h=1} x_{i,h,k,t}$ (18)
$\lambda_{c,k}$	Variable intermédiaire – poids du composant c dans la commande k	$S_{i,t} \geq 0, \forall i, t$ (19)
$\Delta_{c,k}$	Variable intermédiaire – Ecart absolu entre la solution optimale proposée par le modèle et la valeur cible de chaque composant de la commande k	$\Delta_{c,k} \geq v_{c,k} - D_k \times \tau_{c,k}, \forall c, k$ $\Delta_{c,k} \leq v_{c,k} - D_k \times \tau_{c,k}$ (20)

Définition du problème :

$$\min(\sum_{i,h,k,t|U_{i,h}=1} \sigma_{i,h} \times (Q_{i,h} \times o_{i,h,k,t}) + \sum_{c,k} \sigma_c \times \Delta_{c,k}) \quad (1)$$

Tenant compte des contraintes suivantes :

$$D_k = \sum_{i,h,t|U_{i,h}=1} x_{i,h,k,t} \times \eta_{i,h}, \forall k \quad (2)$$

$$v_{c,k} = \sum_{i,h,t} x_{i,h,k,t} \times \eta_{i,h} \times \alpha_{c,i} \times \omega_{h,c,i}, \forall c, k \quad (3)$$

$$\beta_{c,k}^{\min} \times D_k \leq v_{c,k} \leq \beta_{c,k}^{\max} \times D_k \quad (4)$$

$$0 \leq y_{i,h,k,t} \leq 1, \forall i, h, k, t$$

$$y_{i,h,k,t} = \frac{x_{i,h,k,t}}{M} \quad (5)$$

$$y_{i,h,k,t} \leq o_{i,h,k,t}$$

$$z_{i,h,k,t} = \frac{x_{i,h,k,t}}{N}, \forall i, h, k, t \quad (6)$$

$$z_{i,h,k,t} \geq o_{i,h,k,t}$$

$$\sum_{h|U_{i,h}=1} o_{i,h,k,t} = 1, \forall i, k, t \quad (7)$$

$$(1 - U_{i,h}) \times \sum_{k,t} o_{i,h,k,t} = 0, \forall i, h \quad (8)$$

$$\sum_{i,h,t|U_{i,h}=1} o_{i,h,k,t} \geq 1, \forall k \quad (9)$$

$$E_k \times o_{i,h,k,t} \leq (t + g_{i,h}) \times o_{i,h,k,t} \leq L_k \times o_{i,h,k,t} \quad (10)$$

$$\sum_{i,h,k|U_{i,h,r}=1 \cap t - g_{i,h,r} \geq 1} x_{i,h,k,t} - g_{i,h,r} \leq \rho_r \times \phi_{r,t}, \forall r \in \{1,3,4\}, t \quad (11)$$

$$S_{i,0} = S_i^{init}, \forall i \quad (14)$$

$$\forall i \in \{QS_{min e}\}, t$$

$$avec : \{QS_{min e}\} = \{QS_{min e-1}\} \cap \{QS_{min e-2}\} \cap \dots \cap \{QS_{min e-n}\} \quad (15)$$

$$S_{i,t} = S_{i,t-1} + A_{i,t} - \sum_{h,k|U_{i,h}=1} x_{i,h,k,t}$$

Commentaires :

- Relation (2) définit les quantités d'intrants qui doivent composer chaque commande, en tenant compte de la gamme de traitement nécessaire par intrant. Ce traitement selon sa nature permet de réduire le poids du $x_{i,h,k,t}$, d'où la nécessité d'y associer le rendement poids correspondant. Pour la gamme $h=1$ (pas d'enrichissement par traitement mécanique), le rendement poids $\eta_{i,h} = 1$.

- Relation (3) représente la structure chimique de l'extrait k pour chaque composant c. Elle dépend des quantités $x_{i,h,k,t}$, des caractéristiques chimiques de chaque intrant choisis $\alpha_{c,i}$, du coefficient de réduction du poids $\eta_{i,h}$ ainsi que du coefficient chimique lié à l'enrichissement des intrants $\omega_{h,c,i}$.

Relation (4) représente la définition attribuée aux exigences des clients en termes de la qualité demandée. Cette dernière est traduite par des intervalles avec des bornes inférieures $\beta_{c,k}^{\min}$ et supérieures $\beta_{c,k}^{\max}$ pour chaque composant c de la commande.

- Relation (5) et (6) définissent les quantités minimales et maximales à consommer de chaque intrant i. Ceci est au but de ne pas consommer des quantités insignifiantes d'intrants par rapport à la quantité totale demandée par le client.

- Relation (7) décrit le nombre de gammes de traitement à utiliser (gamme unique par intrant) pour chaque commande.

- Relation (8) traduit l'opérabilité de la gamme de traitement choisie.

- Relation (9) définit le nombre d'intrants et par conséquent de gammes à utiliser pour chaque commande et qui doit être d'au moins un intrant.

- Relation (10) est une contrainte traduisant le respect du délai de livraison du produit final. Ce délai est défini par un intervalle de temps avec deux bornes représentant la date au plus tôt et au plus tard d'accostage du navire qui va transporter le produit vers le client final, au niveau du port.

$$(\sum_{i,h,k|h \in \{2..4\} \cap U_{i,h,r}=1 \cap t - g_{i,h,r} \geq 1} x_{i,h,k,t} - g_{i,h,r} \times \eta_{i,h}) + (\sum_{i,h,k|h=1 \cap U_{i,h,r}=1 \cap t - g_{i,h,r} \geq 1} x_{i,h,k,t} - g_{i,h,r}) \leq \rho_r \times \phi_{r,t}, \forall r \in \{2,5\}, t \quad (12)$$

$$\sum_{i,h,k|h \in \{2,3,4\} \cap U_{i,h,r}=1 \cap t - g_{i,h,r} \geq 1} x_{i,h,k,t} - g_{i,h,r} \leq \rho_r \times \phi_{r,t}, \forall r \in \{6,7,8\}, t \quad (13)$$

- Relations (11) et (12) représentent les contraintes d'acheminements des produits entre les entités selon la nature de ces produits et en tenant compte de la capacité maximale journalière des convoyeurs mobilisés. La relation (11) concerne l'acheminement des QS vers l'entrée de la laverie tandis que la relation (12) concerne l'acheminement des QS et des QSL vers l'unité de séchage.
- Relation (13) détermine la quantité totale d'intrants à traiter par unité d'enrichissement en tenant compte de sa capacité maximale journalière
- Relation (14) traduit l'état du stock sur la période $t=0$ qui est S_i^{init} .
- Relation (15) décrit l'état des stocks de QS sur la période t au niveau des mines, en tenant compte de la situation du stock sur la période précédente, du programme prévu d'approvisionnement $A_{i,t}$ ainsi que des quantités consommées sur la même période t en passant par l'une des gammes h .
- Relation (16) représente l'état des stocks de QS située à l'entrée des unités d'enrichissement (laveries). Il est décrit en fonction de la situation de ces stocks sur la période $t-1$ et les quantités consommées sur la période t .
- Relations (17) et (18) sont des contraintes liées à l'état des stocks de QS lavées à la sortie de chaque unité de traitement ainsi que les QS et les QSL stockées au niveau de l'unité de séchage (respectivement). Elles sont décrites en tenant compte du stock sur la période $t-1$ ainsi que la somme des quantités consommées sur la période t . Sur les deux entités, la nature du produit stocké (lavé) fait que la gamme de traitement à adopter est uniquement $h=1$ qui n'impacte pas la composition chimique de ces produits déjà enrichis.
- Relation (19) impose que chaque stock d'intrants soit supérieur ou égal à 0.
- Relation (20) est un calcul de déviance $\Delta_{c,k}$ entre les compositions chimiques des produits finis proposées par le modèle et les valeurs cibles des composants chimiques imposées par le client.
- Relation (1) est une fonction objectif du modèle. Elle permet à la fois la gestion de la qualité du produit fini, à travers la minimisation d'une pénalité fictive liée à la déviance $\Delta_{c,k}$ ainsi qu'une mobilisation optimale des ressources à travers la priorisation de l'utilisation des configurations, en choisissant un chemin de production, dont la durée opératoire est la plus courte et par conséquent qui mobilise le moins de ressources, et en attribuant des pénalités fictives au choix de chaque configuration.

des données et résultats, le modèle intègre sept intrants seulement, mais la diversité et le nombre d'intrants réellement existant sont beaucoup plus importants que ce qui est proposé pour l'exécution du modèle. Ils sont pris en compte en gardant l'information sur la mine source et l'entité au niveau de laquelle ils étaient stockés au moment de consommation. Concernant le carnet de commandes, on propose de satisfaire quatre commandes clients. Les tableaux ci-dessous illustrent les commandes à satisfaire en termes de quantité, qualité et délais :

Tableau 2. Quantités commandées par les clients

	Commande 1	Commande 2	Commande 3	Commande 4
Dj	80	80	100	50

Tableau 3. Caractéristiques chimiques des qualités marchandes

	Commande 1	Commande 2	Commande 3	Commande 4
BPL (%)	$67.5 < \beta_{1,1} < 68.75$	$69.92 < \beta_{1,2} < 70$	$65 < \beta_{1,3} < 66$	$67.5 < \beta_{1,4} < 68.75$
SiO ₂ (%)	$2.2 < \beta_{2,1} < 3.5$	$1.7 < \beta_{2,2} < 1.7$	$2.2 < \beta_{2,3} < 3.9$	$2.2 < \beta_{2,4} < 3.5$
MgO (%)	$0.5 < \beta_{3,1} < 0.65$	$0.38 < \beta_{3,2} < 0.4$	$0.45 < \beta_{3,3} < 0.65$	$0.5 < \beta_{3,4} < 0.65$
Cd (ppm)	$19 < \beta_{4,1} < 25$	$20 < \beta_{4,2} < 21$	$18 < \beta_{4,3} < 25$	$19 < \beta_{4,4} < 25$

Tableau 4. Délais de préparation des commandes

	Commande 1	Commande 2	Commande 3	Commande 4
Début de planche	5	11	13	15
Fin de planche	9	15	17	19

Les quantités d'intrants supposées disponibles sur les différents stocks de la chaîne sont de l'ordre de 100 pour chaque intrant, $\forall i, S_i^{init} = 100$ et leurs caractéristiques chimiques sont représentés sur les tableaux en ci-dessous :

Tableau 5. Caractéristiques chimiques des intrants

$\alpha_{i,c}$	BPL (%)	SiO ₂ (%)	MgO (%)	Cd (ppm)
i=1	55	7.5	2.3	19
i=2	61.5	4.5	0.9	19
i=3	65	4.5	0.9	17
i=4	68	5	0.8	15
i=5	69.8	4.5	0.54	17
i=6	71.5	3.5	0.5	15
i=7	73.5	3	0.4	17
i=8	55	7.5	2.2	19
i=9	61	5.5	1.5	20
i=10	65	3.4	0.56	18

Le programme prévu d'extraction et d'approvisionnement provenant des mines est représenté par le tableau 6 sous forme de quantités d'intrants sur les 30 jours d'un mois donné. Dans la réalité, les intrants alimentés sont beaucoup plus diversifiés et les quantités diffèrent d'un mois à un mois et entre les jours du même mois :

Tableau 6. Programme d'alimentation prévu des mines

$A_{i,t}$	t=1	t=2	...	t=28	t=29	t=30
i=1	0	0	...	100	0	0
i=2	100	0	...	10	0	0
i=3	0	100	...	0	10	0
i=4	0	100	...	0	0	10
i=5	0	0	...	0	0	0
i=6	0	0	...	0	0	0
i=7	0	0	...	0	0	0
i=8	0	0	...	0	0	0
i=9	0	0	...	0	0	0
i=10	0	0	...	0	0	0

Les pénalités de la fonction objectif sont ici fictives et définies pour pondérer l'importance des différents objectifs. La première pénalité $\sigma_{i,h}$ concerne le choix des configurations de production, à laquelle nous attribuons des valeurs variantes entre 400 et 4000. Ces valeurs permettent de favoriser l'utilisation de quelques chemins de production par rapport à d'autres. La deuxième pénalité σ_c , paramétrée à 2000, est liée aux écarts par composant entre la qualité du produit fini demandé par le client et la qualité proposée par le modèle.

4 ILLUSTRATION NUMERIQUE DU REDESIGN ADOPTE

4.1 Données d'entrées

L'exécution du modèle est faite en tenant compte de cinq mines, deux laveries, une unité de séchage et quatre éléments chimiques décrivant la qualité des produits utilisés (BPL, SiO₂, MgO, Cd). Les QM se caractérisent par d'autres composants chimiques et il est possible de les prendre en considération selon les exigences des clients. Au but de limiter la taille des tableaux

4.2 Résultats numériques

Le modèle est exécuté sur le logiciel d'optimisation FICO-Xpress optimization. Il utilise 2440 variables et 2984 contraintes, avec un temps moyen d'exécution de 2.7s.

Le redesign de la chaîne logistique étudiée à travers le modèle mathématique développé permet de proposer des chemins de production par QM. Ces chemins mobilisent les ressources selon les périodes qui leurs sont attribuées et en respectant les capacités imposées des différentes installations.

Le tableau 7 et la figure 3 illustrent les résultats d'exécution du modèle. Un code couleur permet de distinguer sur la figure les différents composants de chaque QM ainsi que leurs chemins de préparation. Le même code couleur est utilisé sur le tableau pour représenter les détails des résultats obtenus par QM. Les mélanges s'effectuent au dernier stade de la chaîne juste avant l'expédition du produit fini.

Le tableau 7 contient plusieurs colonnes détaillant les flots de produits pour chaque commande ainsi que les itinéraires suivis par chaque ingrédient. Les caractéristiques chimiques des extrants sont également représentées sur le tableau 9 respectant bien les exigences des clients.

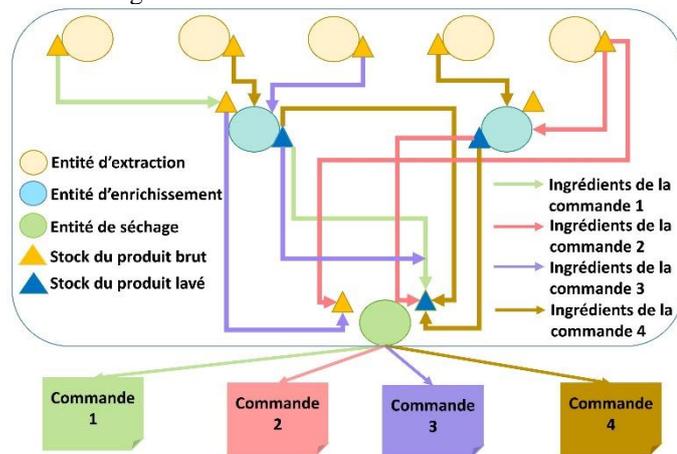


Figure 3. Illustration des chemins de production des commandes proposées par le modèle

4.3 Analyses des résultats

L'exécution du modèle est faite sur la base d'un carnet de commandes composé de quatre commandes, sur des périodes répartis tout au long du mois. Les résultats proposés par le modèle permettent une illustration de la flexibilité et du dynamisme discuté précédemment.

Tout d'abord, en analysant le tableau 7 on remarque des compositions différentes pour les commandes 1 et 4, relative à la même QM (mêmes caractéristiques chimiques) mais à

satisfaire sur des intervalles de temps différents avec des quantités différentes. Ceci reflète le dynamisme des solutions proposées par le modèle en agissant simultanément sur les choix entre les mines (tas de QS) et entre les gammes de traitements, pour la constitution des mélanges qui se fait au niveau du dernier stade de la chaîne avant l'expédition du produit fini.

Quant à la flexibilité, cette dernière est assurée à travers les flots de produits variant par QM selon l'état initial des stocks et la disponibilité des ingrédients ainsi que des installations. Le redesign du mode de pilotage adopté qui consiste à prendre en compte le comportement de chaque composant chimique face aux opérations décrites précédemment et l'impact de chaque trajectoire suivie, du tout amont au tout aval de la chaîne, permet de satisfaire les clients en termes des quatre composants chimiques (BPL, SiO₂, MgO, Cd). Il permet également de pénaliser le choix des configurations de production mobilisant le plus de ressources. Ceci se réalise à travers l'intégration du temps opératoire liée à chaque configuration et la pénalisation du choix du chemin qui a le temps opératoire le plus long. Dans le cas étudié, plus ce temps est long plus le nombre de ressources mobilisées est grand. Ceci pénalise par conséquent la mobilisation de plus de ressources et privilégie les chemins les plus courts.

Quant à la notion du délai, le modèle tient compte de la différence entre la période sur laquelle le produit est choisi pour l'utilisation et la période sur laquelle le produit est mis à disposition de la dernière opération du processus productif, avant l'expédition du produit fini. Il tient compte du temps opératoire nécessaire à chaque opération pour créer cette différence.

La trajectoire suivie par chaque ingrédient est décrite par le stock source et la gamme de production à adopter. L'enrichissement des produits ne peut dans certains cas être suffisant pour satisfaire les clients en termes de plusieurs éléments chimiques. Il est nécessaire donc de choisir entre les mines celle qui pourra répondre au mieux aux attentes des clients. Ceci sera poursuivi par le traitement et de la constitution des mélanges comme opérations complémentaires.

Il est important de signaler que le présent travail permet non seulement d'intégrer un redesign du mode de pilotage pour une meilleure satisfaction des clients à travers le présent outil de planification personnalisé, mais aussi d'obtenir un outil de traçabilité. Il garde l'information sur le chemin suivi par chaque QS et/ou QSL au cours de sa préparation. En cas d'anomalie, il sera facile aux responsables de l'entité de planification de récupérer l'information exacte sur les différents chemins de production par QM, d'étudier la cause racine et d'y remédier.

Tableau 7. Résultats d'exécution du modèle

Commandes (k)	Intrant (i)	Gamme (h)	Période de consommation de l'intrant (i) d'un stock	Période d'utilisation du convoyeur 1	Période d'utilisation de la laverie	Période d'utilisation du convoyeur r 2	Période de mise à disposition à l'unité de séchage (t+ $\mathcal{G}_{i,h}$)	Quantité ($X_{i,h,k,t}$)	Teneur en BPL (%)	Teneur en SiO ₂ (%)	Teneur en MgO (%)	Teneur en Cd (ppm)
1	2	2	2	2	3	4	5	68.11	65	3.168	0.63	20.9
2	2	2	10	10	11	12	13	72.97	67	3.12	0.57	20
	7	1	14	14	-	14	15	16.43				
3	2	2	12	12	13	14	15	44.10	68	3.10	0.54	19.52
	7	1	16	16	-	16	17	17.62				
4	1	2	16	16	17	18	19	33.84	65	2.83	0.58	21.14
	10	2	17	17	18	19	20	84.58				

5 CONCLUSION

Nous proposons dans ce papier le redesign d'un réseau de production du phosphate d'une industrie minière marocaine pour améliorer la satisfaction de ses clients en mobilisant moins de ressources. Le concept du redesign permet d'intégrer une décision au niveau tactique et d'évaluer son impact sur la chaîne ou une partie de la chaîne dans le but d'atteindre un objectif visé. Dans notre cas, le redesign a été appliqué afin d'améliorer la satisfaction des clients en mobilisant moins de ressources. Ceci a été réalisé en tenant compte de l'impact de chaque opération effectuée tout au long du processus productif sur les caractéristiques chimiques des produits, depuis la sortie de la mine qui fournit la matière première, en passant ou non par l'unité de traitement afin d'enrichir le produit, jusqu'à l'arrivée à la dernière station de préparation des qualités marchandes avant le chargement des trains à acheminer vers le port. Le redesign du mode de pilotage adopté a permis de passer de flots de produits figés par QM à des flots dynamiques en adoptant un mode de gestion flexible respectant parfaitement les exigences des clients et en mettant en lumière l'impact énorme d'étudier le comportement des caractéristiques chimiques sur la satisfaction des clients. Le modèle propose une solution par rapport aux choix entre les stocks participants, les itinéraires de production définie par le chemin et les gammes de traitement à adopter pour chaque QS. La constitution des mélanges se fait par l'identification des quantités de QS et de QSL à intégrer dans la production de chaque QM et le choix de l'itinéraire pour ces quantités (gammes de transformation). Plusieurs perspectives sont envisageables pour le présent travail. Il s'agit d'abord d'effectuer un nombre de tests suffisant sur le terrain afin d'évaluer la pertinence de la solution proposée par rapport au mode de gestion adopté actuellement. Il s'agira aussi, dans une deuxième phase, d'intégrer l'impact du processus extractif de la mine sur les caractéristiques chimiques des intrants. Ceci permettra de garantir une vision globale basée sur une complémentarité de fonctions entre les entités de la chaîne depuis le tout amont de la chaîne jusqu'à la dernière opération en aval.

6 REFERENCES

- Azzamouri, A., Bamoumen, M., Hilali, H., Hovelaque, V., & Giard, V. (2020). Flexibility of dynamic blending with alternative routings combined with security stocks: A new approach in a mining chaîne logistique. *International Journal of Production Research*, 0(0), 1-18.
- Berry, D., & Naim, M. M. (1996). Quantifying the relative improvements of redesign strategies in a P.C. chaîne logistique. *International Journal of Production Economics*, 46-47, 181-196.
- Cortinhal, M. J., Lopes, M. J., & Melo, M. T. (2015). Dynamic design and re-design of multi-echelon, multi-product logistics networks with outsourcing opportunities: A computational study. *Computers & Industrial Engineering*, 90, 118-131.
- Feitó-Cespón, M., Sarache, W., Piedra-Jimenez, F., & Cespón-Castro, R. (2017). Redesign of a sustainable reverse chaîne logistique under uncertainty: A case study. *Journal of Cleaner Production*, 151, 206-217.
- Hammami, R., & Frein, Y. (2014). Redesign of global chaînes logistiques with integration of transfer pricing: Mathematical modeling and managerial insights. *International Journal of Production Economics*, 158, 267-277.
- Jahani, H., Abbasi, B., Alavifard, F., & Talluri, S. (2018). Chaîne logistique network redesign with demand and price uncertainty. *International Journal of Production Economics*, 205, 287-312.
- Krishnan, R., Agarwal, R., Bajada, C., & Arshinder, K. (2020). Redesigning a food chaîne logistique for environmental sustainability – An analysis of resource use and recovery. *Journal of Cleaner Production*, 242, 118374.
- Lim, J., Norman, B. A., & Rajgopal, J. (s. d.) (2020). Redesign of vaccine distribution networks. *International Transactions in Operational Research*, n/a(n/a).
- Martins, C. L., Melo, M. T., & Pato, M. V. (2019). Redesigning a food bank chaîne logistique network in a triple bottom line context. *International Journal of Production Economics*, 214, 234-247.
- Martins, S., Amorim, P., Figueira, G., & Almada-Lobo, B. (2017). An optimization-simulation approach to the network redesign problem of pharmaceutical wholesalers. *Computers & Industrial Engineering*, 106, 315-328.
- Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2012). A tabu search heuristic for redesigning a multi-echelon chaîne logistique network over a planning horizon. *International Journal of Production Economics*, 136(1), 218-230.
- Persson, G. (1995). Logistics Process Redesign: Some Useful Insights. *The International Journal of Logistics Management*, 6(1), 13-26.
- Pissonnier, S., Dufils, A., & Le Gal, P.-Y. (2019). A methodology for redesigning agroecological radical production systems at the farm level. *Agricultural Systems*, 173, 161-171.
- Razm, S., Nickel, S., & Sahebi, H. (2019). A multi-objective mathematical model to redesign of global sustainable bioenergy supply network. *Computers & Chemical Engineering*, 128, 1-20.
- Razmi, J., Zahedi-Anaraki, A., & Zakerinia, M. (2013). A bi-objective stochastic optimization model for reliable warehouse network redesign. *Mathematical and Computer Modelling*, 58(11), 1804-1813.
- Saez-Mas, A., Garcia-Sabater, J. P., Garcia-Sabater, J. J., & Ruiz, A. (2020). Redesigning the in-plant supply logistics: A case study. *Computers & Industrial Engineering*, 143, 106422.
- Vorst, J. G. A. J., & Beulens, A. J. M. (2002). Identifying sources of uncertainty to generate chaîne logistique redesign strategies. *International Journal of Physical Distribution & Logistics Management*, 32(6), 409-430.
- Vorst, J. G. A. J. van der, Tromp, S.-O., & Zee, D.-J. van der. (2009). Simulation modelling for food chaîne logistique redesign; integrated decision making on product quality, sustainability and logistics. *International Journal of Production Research*, 47(23), 6611-6631.

CIGI 2021

Understanding Saturation of a dedicated Non-Elective Surgery Theater

GUSTAVO SANTAMARIA-ACEVEDO¹, BENJAMIN LEGROS², OUALID JOUNI¹, ZIED JEMAI³, MATHIEU RAUX⁴, GUILLAUME LAMÉ¹

¹ Université Paris-Saclay, CentraleSupélec, Laboratoire Genie Industriel, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France
gustavo.santamaria@centralesupelec.fr • oualid.jouini@centralesupelec.fr • guillaume.lame@centralesupelec.fr

² EM Normandie, Laboratoire Métis, 64 Rue de Ranelagh, 75016. Paris, France
benjamin.legros@centraliens.net

³ R-OASIS, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
zied.jemai@centralesupelec.fr

⁴ Sorbonne Université, INSERM, UMRS1158 Neurophysiologie Respiratoire Expérimentale et Clinique; AP-HP, Groupe Hospitalier Pitié-Salpêtrière Charles Foix, Département d'Anesthésie Réanimation, Paris, France.
mathieu.raux@aphp.fr

Résumé – Nous analysons un bloc opératoire d'urgence avec des opérations non-programmées. En se basant sur des données réelles, nous construisons un diagramme de flux détaillé afin d'intégrer la plupart des caractéristiques réelles en termes de ressources, de synchronisation et de parcours des patients. Ensuite, à partir de ce diagramme de flux et des données d'arrivée des patients et des durées de chirurgie, nous construisons un modèle de simulation qui nous permet de reproduire la situation actuelle de l'hôpital et de prédire son comportement futur. L'objectif de cette étude est de déterminer comment réduire les temps d'attente des patients en chirurgie non-programmée dans des conditions proches de la saturation. Les résultats des expériences menées montrent que même si les salles de chirurgie représentent la ressource la plus critique, l'ajout de salles de chirurgie supplémentaires n'est pas le seul moyen de réduire les temps d'attente. D'autres alternatives, telles que des équipes supplémentaires de soins infirmiers et d'anesthésie, peuvent réduire considérablement les temps d'attente et de traitement. Nos résultats montrent que la solution la plus rentable et la plus simple pour réduire les temps d'attente consiste à investir dans les ressources de transport de patients entre les chambres et les salles d'opérations.

Abstract – We analyze an Emergency Surgery Theatre (EST) with non-elective patients. Based on real data, we build a detailed flowchart in order to integrate most of the real-life characteristics in terms of resources, synchronization, and patients' paths. Next, from this flowchart and the data representing patients' arrivals and surgery durations, we construct a simulation model which allows us to replicate the current situation of the hospital and to predict its future behavior. The objective of this study is to determine how to reduce the waiting times of non-elective surgery patients under close-to-saturation conditions. The results of the undertaken experiments show that although surgery rooms are the most important resource, building additional surgery rooms is not the only way to reduce the waiting times. Other alternatives, such as additional nursing and anesthesia teams can reduce the waiting and processing times notoriously. Our results show that the most cost-effective and easiest solution to reduce waiting times in the EST is to hire additional transportation aids.

Mots clés – Chirurgie urgente, simulation à événements discrets, systèmes complexes, synchronisation des ressources, bloc opératoire

Keywords – Non-Elective surgery, Discrete Event Simulation, Complex Systems, Resource Synchronization, Operating Theater

1. INTRODUCTION

Surgery Theaters are considered to be the heart of the hospital, representing between 20-40% of the costs and 60-70% of the revenues of a hospital (Macario et al, 1995) (Denton et al, 2007) (HCUP, 2017). A special category of surgery theaters are the Emergency Surgery Theaters (EST), which are dedicated specifically to serve non-elective surgeries. Non-elective surgeries are those in which the patient arrives to the Emergency Department (ED) with symptoms or traumas that are so severe that a surgery needs to be performed in matter of hours. Even though non-elective surgeries are vital for the life of the patients who require them, most of the available literature and models are related to elective surgeries and the optimization of elective surgery wards. However, non-elective surgeries present distinct characteristics that require the development of dedicated models and management strategies

(Duma and Aringhieri, 2019). Among other issues, the stochasticity of arrivals, the dynamic and sometimes extremely quick evolution of the severity degrees of the patients, and the complexity of managing resource synchronization with short notice, makes the improvement of the efficiency of an EST a very challenging task and one that differs significantly from the management of elective surgeries. Because of these specific aspects, some hospitals have set up dedicated ESTs (Heng and Wright, 2013), which have empirically shown to increase utilization while reducing overtime and cancellations (van Riet et Demeulemeester, 2015) (van Veen Berkx et al, 2016).

Unlike elective surgeries, which are scheduled weeks or even months ahead, non-elective or emergency surgeries are scheduled less than 48 hours before the surgery takes place and are subject to schedule changes if the condition of the patient

worsens. Additional to the unpredictability of these procedures, hospitals are facing a sustained increase in the demand for emergency surgeries, which increases the utilization of all the EST resources as well as put additional pressure on the hospital management.

In this paper, we present the results of work conducted in collaboration with the EST of the largest French hospital, La Pitié-Salpêtrière. La Pitié-Salpêtrière serves a population of 12 million people in the Île-de-France department, and it is the referral center for several hospitals in the metropolitan area of Paris. It has 1,756 beds as of 2018, which makes it one of the largest university hospitals in Europe. The same year, the emergency services of the hospital received 66,027 patients of which 7,700 required hospitalization and 4,500 required non-elective surgery. The surgery ward of the Pitié-Salpêtrière hospital has 13 operating rooms. Among these, 2 are dedicated solely to Non-Elective surgeries. These 2 rooms make up what for purposes of this study will be called Emergency Surgery Theater (EST). Both the ED and the EST are open 24 hours a day and 365 days per year.

The EST of the Pitié-Salpêtrière hospital has been facing increasing challenges in the last years due to an increase in the number of patients who require emergency surgery, a phenomenon which has subsequently revealed problems in the internal processes. This increase is the main reason behind this work. The problems faced by the EST are intensified by the conflicting interests between diverse stakeholders. There are surgeons preferring to perform surgery at certain times of the day, as well as surgeons who do not perform surgery at night. This generates lower utilization rates of the surgery rooms and staff during the night shifts, but creates a bottleneck between 8:00 and 10:00a.m., when surgeons start registering their patients for the day. Anesthesiologists may schedule semi-urgent patients in their personal order of preference, creating unnecessary scheduling issues. There are also cases of nurses only working with a specific anesthesia team out of personal preference, generating additional delays. Another example of conflicting interest that impacts the performance of the EST are surgeons who prefer to perform several surgeries consecutively, but the nursing and anesthesia teams preferring to work in a different way, creating scheduling conflicts.

Additionally, being dedicated to non-elective surgery, the surgery theater faces the issues generated by the stochastic behavior of the demand for surgery rooms, which prevents the managers of planning ahead. This stochasticity creates not only scheduling problems but also problems such as a high utilization of transport aids at specific moments, which subsequently generates problems in the patient flow.

The hospital's management is aware of these difficulties and have intuitions about eventual solutions to improve the service quality provided to patients. From our discussions with decisions-makers in the hospital, four hypotheses were formulated to evaluate the current state of the EST and eventual decisions that should be taken. These hypotheses were formulated as follows:

- H1: The situation at the EST is close to saturation.
- H2: The bottlenecks at the EST are mainly created by a surgeon scheduling problem.
- H3: Adding an emergency surgery room may significantly improve the system performance.

- H4: Patients' transportation plays a moderate role in the system congestion.

Our aim in this paper is to discuss these hypotheses and quantify them in order to support decision-makers. Although specific to Pitié-Salpêtrière's hospital, our analysis and conclusions may be of interest for other hospitals facing similar saturation problems.

Based on a large amount of data, complemented with interviews and focus groups with hospital staff and managers, as well as fieldwork in the hospital, we characterize each process that takes place in an EST, with their corresponding resources, probabilities, and durations. From this, we used the simulation software Arena to develop a discrete-event simulation model of the EST which replicates the main characteristics of the real system. Our model considers requirements on resource synchronization, multiple patient specialties, and the dynamic evolution of patient severity. This simulation model allows us to evaluate the effect of having a higher demand, adding resources, or implementing new scheduling policies.

2. LITERATURE

Non-elective surgery theaters have been less studied than elective ones (Samudra et al., 2016). We refer to the review of (Cardoen et al., 2010) for an overview of the research related to surgery room management. First, we present the references related to the management of *non-elective* surgery rooms. Next, we describe those for *elective* surgery rooms employing a similar approach as ours. Finally, we give the literature related to policy making for the combined use of elective and non-elective surgery rooms.

Bowers and Mould (2004) created a simulation model to analyze the behavior of a non-elective surgery theater, to explore the effects in terms of waiting time of including scheduled elective procedures in this theater and maximize the utilization of the surgery rooms. Persson and Persson (2010) simulated and compared between three scenarios: the base scenario, a second scenario where the turnover between surgeries was reduced by 20 minutes, and a third scenario with an additional surgery room. Paul and MacDonald (2013) used a non-preemptive multi-priority queueing model to manage non-elective surgeries and optimize the resource allocation in the surgical department. Based on data from a hospital, Antognini et al. (2015) analyzed a non-elective surgery theater with Monte Carlo simulation. They determined the optimal number of operating rooms that are necessary during three different shifts to minimize the waiting time of emergency surgery patients. Additionally, they analyzed the effect of changing some parameters in the system, specifically the cleanup time and the patient volume. We have used a similar approach (i.e., a simulation model built from real data) to determine the effect of an increased demand, adding a third surgery room in the EST, having more resources, or reducing transportation delays.

For elective surgery rooms, most studies rely on simulations too Wullink et al. (2007). For example, van der Lans et al. (2006) used simulation to compare five methods to anticipate urgent surgery in a mixed-use surgery ward. Their tests show that allocating slack times for urgent surgeries in elective operating rooms can reduce the effects of demand variability. Lamiri et al. (2008) proposed a method that combines

simulation with mixed integer programming to solve the problem of planning a surgery room that serves both elective and emergency patients, which can be viewed as competing classes in some scenarios due to their different characteristics. The results of their experiments show that the costs for the hospital can be significantly reduced by using stochastic models where the uncertainty related to emergency surgery is explicitly considered. Stepaniak and Dexter (2016) showed that both theatre and surgeons can be limiting resources to start an emergency surgery, highlighting the need to account for the full range of resources when analyzing ESTs. This consideration was a key part of this study, since one of our objectives was to show that although the surgery room is the most important resource, it is necessary to consider all the resources and the synchronization between them to fully explain the phenomena that occur in the surgery room.

A key decision for hospital administrators is to combine or to separate elective and non-elective surgeries, by choosing whether to have dedicated ESTs. After reviewing the literature on managing elective and non-elective surgeries, Ferrand et al. (2014) concluded that there are still mixed opinions on the value of dedicated ESTs in the literature. Wullink et al. (2007) concluded that flexible ESTs can lead to better efficiency through reduced waiting times, reduced overtime and increased theatre utilization. However, in the review of van Riet and Demeulemeester (2015), the trade-offs involved in choosing to have separated or combined theatres are shown. Some results suggest that contrary to the conclusion of Wullink et al. (2007), dedicated ESTs such as the one in the studied hospital reduce non-elective waiting times to surgery, increase theatre utilization, reduce staff overtime, reduce disruptions in elective surgeries and increase throughput (Heng and Wright, 2013; Veen-Berkx et al., 2016). Finally, Duma and Aringhieri (2019) proposed models to evaluate which policy (dedicated, hybrid, or flexible) should be implemented according to the hospital characteristics.

3. PROCESS DESCRIPTION

We have identified each part of the process undertaken in the EST. In addition to a description of the resources involved, we analyze the steps of the standard path followed by a surgery patient in the hospital. The goal is to build a flowchart representation allowing us to replicate the behavior of the EST in a simulation model as shown in Figure 1. The path followed by a surgery patient in the hospital is composed by a series of assessments, examinations, preparations and procedures which are explained below.

3.1. Reception

Outpatients may arrive to the ED reception on foot or by ambulance, where they will be registered in the system by the responsible person in the ED. In this registration some details about the patient pathology and severity are already filtered. Afterwards, the patient will be subject to a triage examination performed by a triage nurse. Here an initial assessment of criticality is performed, as well as the possible patient pathway: medical or surgical. Afterwards, the possible surgical patient is examined by a general ED physician. If a surgery is needed, the patient is categorized into one of the five surgical specialties (Maxillo-Facial, Orthopedic, Urology, General, and Vascular) and the surgery resident of the respective specialty is called. At the hospital, from the data from 2013 to 2017, the

proportion of each specialty is 18.46%, 39.61%, 12.52%, 23.68%, and 5.72%, respectively.

3.2. Initial Consultation

After patients have been admitted, they must be checked by the surgery resident of the specific specialty, to reevaluate the necessity of performing a surgery. In an affirmative case, their criticality level is identified. Despite the most extended categorization in the literature includes 4 levels - Trauma, Emergent, Urgent, and Semi-Urgent, this hospital uses the last 3 levels: Emergent (surgery required within six hours or less), Urgent (surgery required within 24 hours) and Semi-Urgent (patients who can wait up to 48 and sometimes 72 hours). Each one of them is assigned with a color code, respectively Red, Orange, and Green. This classification results in priority rules. Red patients have priority over orange and green patients, and orange patients have priority over green patients. This means that if a red patient arrives while orange or green patients are waiting for surgery, then this red patient will be served before the orange and green patients already waiting in the system. For our simulation model, we assume that the service rule within one class (red, orange, or green) is first-come-first-served. In practice, with red patients, some extra priority can be provided to some red patients when their life is in danger. If the waiting time in one category of patient is too long, then this patient can become a red patient. Specifically, if an orange patient waits 12 hours, or if a green patient waits 72 hours, then these patients become red patients. In Table 1, we provide the proportion of green, orange, and red patients from the data from 2013 to 2017. Note that the priority among patients is only determined by the criticality and not by the specialty.

Table 1: Patient severity levels per specialty

	Green	Orange	Red
Maxillofacial	48.10%	35.51%	16.39%
Orthopedic	43.80%	52.68%	3.51%
Urology	26.62%	36.69%	36.69%
General	14.61%	61.48%	23.91%
Vascular	13.32%	15.76%	70.93%

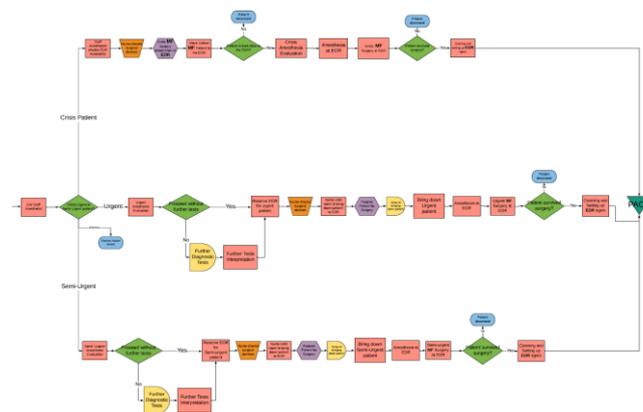


Figure 1: Flow of non-elective surgery patients in the OR

3.3. Scheduling, wait, and transportation.

In case the patient is assigned a Red code, the staff anesthetist will check the EST schedule, and will assign the patient to the next available slot that is within the estimated survival time of the patient. In the meantime, the patient will wait directly in the ED. In case of orange codes, the anesthetist will also check the ESR and the staff schedule and will assign the patient to the next suitable slot that does not create scheduling and staff conflicts, but the patient will usually wait for the surgery in a

hospitalization room (unless there is a slot in the next few hours, a case in which the patient will wait inside the ED) and will be brought to the surgery room when required. In case of a green code surgery, the patient will be sent either to a hospitalization room and in some cases the patient will be sent home and will wait for a call to return to the hospital. In all cases, when the patient has been scheduled by the anesthesiologist, the surgery nurse will call the floor to bring the patient down from the hospitalization room to undergo preanesthetic assessment. The time to call, the delay to find an available staff for transportation, and the transportation by itself represent a delay that is modeled in our simulations. Further details are provided for these delays in Section 4.2.

3.4. Surgical procedure

Before the surgical procedure itself begins, a surgery nurse will check that the required surgical devices are available. Next, the patient will go through an anesthesia evaluation with the anesthesia resident, and the anesthesia nurse. Afterwards, the patient will be taken to the surgery room and the anesthesia is performed on the patient by the anesthetist, the anesthesia resident, and the anesthesia nurse. Once the patient is anesthetized, they will be prepared for the surgery (in terms of cleansing and asepsis) by a surgery nurse and a surgical aide. After this, the surgery itself can start. There are five surgeons in this emergency surgery theater, one for each type of surgery: Maxillofacial, Orthopedic, Urologic, Vascular, and General. Not all of them are in the EST the whole time; there are Orthopedic, General and Maxillofacial surgeons always on duty in the EST, while the Urology and Vascular surgeons are usually available on call, since the number of surgeries for these specialties is lower. Each one of these surgeons is supported by a surgeon in training, usually called surgery resident. A surgery requires the presence of the specific surgeon and resident for each specialty, the staff anesthetist (who stays only for the beginning of the surgery with green patients and for the whole procedure with red patients), the anesthesia resident, a surgery nurse, an anesthesia nurse, and a surgical aide. After the surgery, the patient wakes up at the surgery room. In case of congestion, the patient is transferred to the Post-Anesthesia Care Unit (PACU) for the waking up.

3.5. Waking up at PACU

Once the surgical procedure is completed, the staff will decide where the patient will wake up. In most of the cases the patient will wake up inside the surgery room, but in cases in which the room is urgently needed, the patient will wake up at the PACU. We do not evaluate the time spent by patients at the PACU as it does not represent a bottleneck for the EST.

4. DATA ANALYSIS

To build the simulation model in Arena, we collected different sets of data. The hospital provided us with datasets which contained information of 23,127 programmed surgeries in the general EST between January 1st of 2013 and December 2017. The datasets contain records about the surgery patients for whom a surgery room has been booked, including among others, the motive of the emergency surgery, the name of the surgical procedure, and the duration of the surgery. From this data, in Section 4.1, we analyze the distribution of registrations which will play the role of the arrival process in our simulation model. Next, in Section 4.2, we investigate the distribution of the durations of the process followed by a patient. Finally, in

Section 4.3, we present the simulation model built to replicate the behavior of the EST.

4.1. Analysis of the registrations

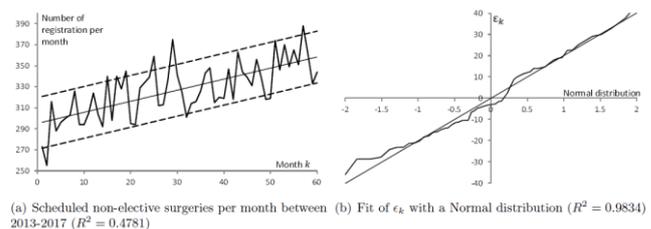
We analyzed the data of the registration pattern over time. We want to determine whether a trend or seasonality effects can be exhibited from the data. This in turn will enable us to provide an appropriate forecast for the simulations.

4.1.1. Long-term trend and seasonality.

In Figure a, we present the number of registrations per month, N_k , from January 2013 to December 2017, where k is the associated month starting in January 2013 with $k=1$. We want to determine an appropriate model \tilde{N}_k to represent this time series to replicate the arrival pattern and forecast future demand. The simplest model is a linear regression. We thus propose the following model to represent the time series:

$$\tilde{N}_k = 1.0533k + 295.07 + \varepsilon_k \quad (1)$$

where ε_k follows a Normal distribution with mean 0 and standard deviation 19.22. In Figure b, we provide the Normal probability plot showing the fit of the computed values of the difference between the data and the trend with a standard Normal distribution with mean 0 and standard-deviation 1. The coefficient of determination is $R^2=0.9834$, showing a very good fit of this model. A Kolmogorov-Smirnov Test provides a KS coefficient of 0.096. Given that with 60 values and a risk 5%, the threshold is $0.1141 > 0.096$, this confirms the good fit with the Normal distribution. In the simulations, we employ the model defined in Equation (2) to replicate the actual and future behavior of the monthly arrival process.



(a) Scheduled non-elective surgeries per month between 2013-2017 ($R^2 = 0.4781$) (b) Fit of ε_k with a Normal distribution ($R^2 = 0.9834$)

Figure 2 : Analysis of the monthly registration process.

4.1.2. Daily and hourly effects

In Figure, we present the distribution of the number of registered surgeries per day for different days of the week (Monday, Wednesday and Saturday) from the beginning of 2013 to the end of 2017. The other days of the week show a similar distribution. We exclude the days where no surgery is scheduled from this analysis. We observe a similar distribution from one day to the other with a high frequency of having only one surgery (around 18%) and the rest of the distribution behaving as a Normal distribution with mean 10. In Table 2, we provide the mean and the standard deviation of the number of surgeries per day. This confirms a similar behavior from one day to the other.

Table 2: Distribution of the number of surgeries per day

Day	Mean	Std. Dev.	χ^2 Coeff.
Sunday	7.792	4.194	25.836
Monday	8.271	4.127	10.674
Tuesday	8.269	4.590	22.538
Wednesday	7.920	4.403	14.801
Thursday	8.099	4.381	10.332
Friday	8.919	4.547	35.661

Saturday	8.367	4.249	10.206
----------	-------	-------	--------

We note that Friday is the busiest day while Sunday is the less busy one. However, these differences are not significant. In the last column of Table 2, we compute the cumulative χ^2 coefficient between the distribution of the number of surgeries in a given day as compared to the theoretical one obtained if all days behaved identically. The threshold with risk 5% and 23 degrees of freedom is 35.17. As all coefficients -except the one of Friday- are below this threshold, we can validate with risk 5% the hypothesis that all days of the week have statistically the same mean. Following these results, in our simulation model, we assume that the distribution of the number of surgeries per day is independent from the day of the week and follows the distribution of the number of surgeries for all the days of the week. Our analysis found that the majority of patients are booked for surgery between 8am and 6pm, with a first peak at 9am and a second peak between 6pm and 8pm. This observation is consistent with the observations made during the previous years. The surge in the registrations that starts at 8am is because this is the time at which the surgeons who did not spend the night at the hospital start working, and they start booking the patients whose condition worsened during the night and require surgery in the next hours, or patients for whom the surgeon wanted the kind of surgery they must undergo to be decided in a staff meeting, and therefore are booked after said meeting. The second peak, during the late afternoon, is mainly due to orange patients that the surgeon has been monitoring during the day, and in the meantime their condition has stabilized enough to get surgery. In our simulation model, we include the hourly seasonal effect by applying the multiplicative seasonal coefficients of 2017. Similar coefficients have been observed in the past and year 2018 has shown to follow the same behavior.

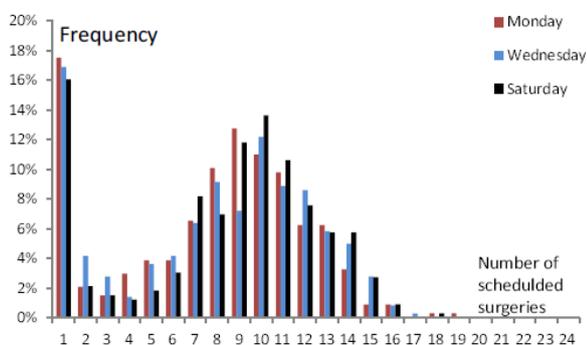


Figure 3 : Distribution of the number of surgeries per day from 2013 to 2017

4.2. Analysis of the procedure duration

To complete the analysis, we now investigate the procedure duration. The hospital provided us the data of the estimated surgery duration, defined as to the expected amount of time between the incision being performed on the patient, and the patient being closed up and real registered data for the overall surgery process including surgery, cleaning and anesthesia. Unfortunately, the data were imprecisely registered with standard durations like 30, 45, or 60 minutes. In Figure , we show the estimated values for the surgery duration of Vascular patients, where we can observe that the surgery duration has been registered as a discrete variable rather than a continuous one as it should be. With several identical or very close values in a data set, the classical statistical tests like Shapiro-Wilk or

Kolmogorov-Smirnov fail in showing the fit of the data with a standard distribution. One way to overcome this limitation is to apply a continuity correction to our data. However, from our discussion with the hospital, it is difficult to determine how the mistakes in registrations were made. It could be that there is a tendency to over or underestimate the real durations. Therefore, we do not wish to apply any changes in the data. Instead, for each duration, we provide the coefficient of determination between the logarithm of the data and a standard Normal distribution with mean 0 and standard-deviation 1. Although, this method is not as rigorous as a statistical test, it shows that our data may be well modeled by a Log-Normal distribution.

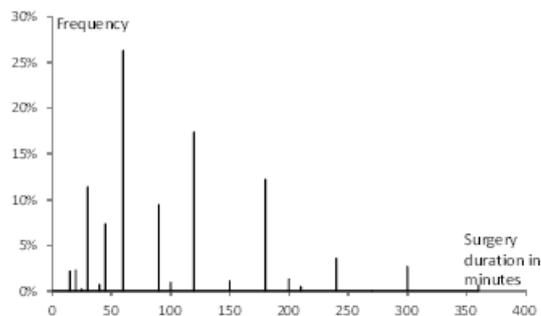


Figure 4: Distribution of the procedure length for vascular patients 2013-2017

In Table 3 , we provide information about the mean duration, the standard-deviation (SD), the coefficient of determination with the Normal distribution (R^2) as well as the proportion of each surgery type (%). These values will be used in the simulations as parameters of the procedure duration. Our results show that each specialty differ in durations for the surgery process as well as for the variability. Vascular operations are the longest operations in terms of surgery and preparation while Urology ones are the shortest. By computing the ratio between standard deviation and mean, we observe that each specialty behaves similarly in terms of variability as this ratio remains close to 0.6. Finally, the coefficients R^2 being close to one suggest that the Log-Normal distribution may be appropriate to model the surgery duration and the total process duration.

Table 3: Duration of the process per specialty (in minutes)

Specialty	%	Total Process Duration		
		Mean	Std. Dev.	R^2
M. Facial	18.46	135.68	87.40	0.9881
Ortho.	39.61	166.80	89.15	0.9604
Urology	12.52	102.16	64.10	0.9589
General	23.68	135.13	84.55	0.9824
Vascular	5.72	194.02	131.44	0.9510

4.3. The Simulation Model

Based on the process description of Section 3 and the statistical analysis of the arrival process, the surgical process, and the transportation process made in this section, we build a discrete event simulation model to replicate the behavior of the EST. The model is a non-terminating simulation, corresponding to the EST being open 24/7. We assume that each part of the procure duration follows a Log-Normal distribution with parameters corresponding to those found in this section or suggested by the hospital when the information was missing. Furthermore, we model the arrival of patients by a time-dependent Poisson distribution where the arrival rate is

modified every hour. For one simulation, we replicate a given year until the performance measures (waiting time, utilization rate...) converge to their stationary values. Assume that the expected number of arrivals for a given year is N , then the hourly arrival rate at hour i (for $i=1, 2, \dots, 24$) of month k (for $k=1, 2, \dots, 12$) is given by

$$\lambda_{i,k} = (N/12 + \varepsilon_{i,k}) / (24 \cdot n_k) * C_i \quad (2)$$

where n_k is the number of days in the considered month (28, 29, 30 or 31), $\varepsilon_{i,k}$ is the normal distribution defined in Section 4.1 with mean 0 and standard-deviation 19.22, and C_i is the hourly coefficient corresponding to the hour of the day (i.e., the average hourly coefficients of year 2017). As shown in Section 4.1, the month of the year and the day of the week do not impact the time-dependent arrival rate. Note that in case $\lambda_{i,k}$ becomes negative due to a very low value of $\varepsilon_{i,k}$, we set $\lambda_{i,k} = 0$ in the simulation. This scenario is however unlikely to happen in our simulations as the minimal value for $N/12$ is set at 291.67 for year 2013 and the standard deviation of $\varepsilon_{i,k}$ is 19.22.

One of the novelties of this simulation model is that it treats and analyzes every single resource independently, and the resources are not assumed to be readily available for every surgery. As described in Section 3, there are 16 agent types in the model, besides the surgery rooms. Hence, the simulation can consider, for example, that a surgery room is available, and both the surgical team and the patient are ready, but the surgery cannot begin because the anesthesiologist is in the other surgery room with another patient. Therefore, the model was designed to reflect the need for resource synchronization that is required in a real-life surgical environment.

5. SIMULATION STUDY

In this section, selected results of the different simulations will be presented and explained. In each of the four subsections we aim to discuss the hypotheses formulated in the Introduction.

5.1. H1: The situation at the EST is close to saturation.

To discuss this hypothesis, we evaluate what can happen in the surgery theater if the demand (the number of patients requiring non-elective surgery) keeps growing. In Table 4 we provide the expected wait for red, orange, and green orthopedic patients as a function of the expected number of patients per year. Similar observations are made with other specialties. Recall that in the current situation we have approximately 4500 patients per year. Therefore, in this table, we replicate the past performance and forecast the future evolution. The results of the simulation show that if the demand keeps growing at the current rates, in a close future the EST will not be able to provide surgery to semi-urgent (green) patients in a timely manner because the capacity of the system will be exceeded.

This increase in the waiting times would lead to patients having to wait longer times than recommended, which would have a negative impact on their health. Under the current conditions, it seems that the system is getting closer to its saturation point as the wait is significantly increasing, which means that there is a need for additional resources. Hence, Hypothesis H1 is clearly confirmed by our simulation study.

Table 4: Waiting times (in hours) for Ortho. patients

Number of patients per year	Red Code	Orange Code	Green Code
-----------------------------	----------	-------------	------------

3500	1:32	4:41	7:23
3750	1:41	4:59	8:42
4000	1:49	5:21	10:39
4250	1:55	5:44	13:13
4500	2:15	6:43	21:04
4750	2:20	7:04	25:32
5000	2:26	7:20	45:21
5250	2:53	8:29	177:09
5500	3:23	10:26	∞

5.2. H2: The bottlenecks at the EST are mainly created by a surgeon scheduling.

Among the experiments that were performed, was the inclusion of additional clinical staff members to test what effect they would have on the waiting times. In Table 5, we show the results for Orthopedic and Maxillofacial patients of different experiments like having 2 orthopedic surgeons instead of 1 (Case 1), 2 surgeons and 2 surgery residents for all specialties instead of 1 (Case 2), one additional anesthesia resident during the day and at night (Case 3), one additional anesthesia nurse and one additional aid (Case 4), one additional surgery nurse (Case 5), and one additional anesthesiologist (Case 6). The results for the remaining specialties are very similar.

Adding one surgeon (Case 1), regardless of the specialty, reduces the waiting times only marginally. Having 2 surgeons and 2 surgery residents for all specialties has an effect which is mostly beneficial for green patients. Having one additional anesthesia resident (Case 3) leads to a stronger improvement than adding 5 surgeons (Case 2). Surprisingly, having one additional anesthesia nurse and one additional aid (Case 4) provides more improvement than adding an anesthesia resident (Case 3). Having one additional surgery nurse (Case 5) has only a significant effect for red and orange patients. Finally, the highest improvement is obtained by adding an anesthesiologist to the EST.

Therefore, and unlike Hypothesis 2, from a clinical staff perspective the bottlenecks are created by the temporary lack of anesthesiologists or nurses, and not by a surgeon scheduling problem. This is explained by the fact that anesthesia teams (anesthesiologists and anesthesia residents), and nursing teams are required in all surgeries regardless of the specialty. Having more anesthesia and nursing teams increases the probability of the surgery to start without delays, since the synchronization levels are higher.

Table 5: Waiting times until procedure for Orthopedia and Maxillofacial patients (in hours)

		C. 1	C. 2	C. 3	C. 4	C. 5	C. 6
MFR	2:08	2:08	2:01	2:07	2:02	1:52	1:48
MFO	6:53	6:53	6:17	6:28	6:18	6:36	5:52
MFG	21:22	21:22	18:10	17:24	17:08	21:13	14:27
OPR	2:15	2:15	2:05	2:08	2:07	1:54	1:47
OPO	6:43	6:43	6:10	6:19	6:10	6:25	5:43
OPG	21:04	21:04	17:54	17:12	17:01	20:39	14:13

5.3. H3: Adding an emergency surgery room may significantly improve the system performance.

To address the high utilization levels of the surgery rooms, the hospital administration decided to build a third emergency surgery room. Although the surgery room will be built, the hospital has yet to determine how it will be used. We analyzed 5 scenarios: the third room is open 24/7; open from 8:00 until midnight every day (8-24E); open from 8:00 until midnight only on weekdays (8-24W); open from 8:00 until 18:00 every

day (8-18E) and open from 8:00 until 18:00 on weekdays (8-18W).

Table 6: Waiting times until beginning of procedure after the opening of a third surgery room (Ortho and MF)

		8-18W	8-18E	8-24W	8-24E	24/7
MF R	2:08	1:22	1:14	1:08	0:57	0:56
MF O	6:53	4:42	4:21	4:12	3:59	3:49
MF G	21:22	7:21	6:28	6:05	5:33	5:16
OP R	2:15	1:23	1:14	1:09	0:56	0:55
OP O	6:43	4:38	4:21	4:09	3:58	3:50
OP G	21:04	7:12	6:22	5:59	5:26	5:05

Table 6 summarizes the effect that adding a third surgery room has on the waiting times of maxillofacial and orthopedic patients. Adding a third surgery room that functions 24/7, can reduce the current waiting time of red patients of around an hour, of orange patients of 3 hours and of green patients of 17 hours. *Therefore, our analysis demonstrates the validity of Hypothesis 3; a third surgery room, even without hiring extra staff, has the ability to improve the service quality substantially.* It is then up to the hospital management to determine how this third surgery room will be employed. A 24/7 scenario is optimal in terms of performance, but it may be too costly or impractical due to some resources not willing to work at night. The consensus so far is to open the third surgery room between 8am until 6pm from Monday to Friday (code 800-1800W). This scenario has the ability to already significantly reduce the waiting times as compared to the current situations while being acceptable by the staff members.

5.4. H4: Patients' transportation plays a moderate role in the system congestion.

After the surgery has been scheduled, patients wait in their rooms or in the ED until they are ready to go to surgery. Once the surgery room and the patient are ready for the surgery, the patient must be brought down to the surgery theater. There are two elements that make this process more complex than expected. The first one is the size of the hospital: La Pitié-Salpêtrière is among the largest hospitals in Europe by area, and some buildings within the hospital can be up to 1 kilometer away from each other. This makes the internal transportation operations to take as much as one hour in some cases, since they require an ambulance to transport the patients between the buildings. The second element that complicates the patients' transportation regards the staff: the transport attendants (or stretcher-bearers) do not work exclusively for the surgery ward; they serve the entire hospital and provide transportation to any patient in the hospital who requires it. This synchronization problem between the transport attendants and the nurses and the EST in general, generates that orange and green patients in the hospital experience a long delay before being taken to the surgery room, since they might have to wait for a stretcher-bearer to be available before being transported to the surgery ward. On average, the delay for the transport of non-critical (orange and green) surgical patients to start is of 31.5 minutes. In this section, we investigate the effect of reducing the delay to find an available staff. In practice, reducing this delay can be achieved by hiring additional transport attendants who would work specifically for the emergency surgery service, or by increasing the coordination and synchronization levels between the actors using new information systems that would prioritize the picking up of non-elective surgical patients over any other patients. A third alternative is to have a special waiting room

such that the transport and its delay do not affect the schedule, since the patient is already waiting downstairs.

In Table 7, from the current situation with 4500 patients per year, we investigate the effect of reducing the delay before transport due to unavailable transport staff. Reducing this delay can notoriously reduce the total waiting times between the surgery booking and the patient entering the surgery room. As a matter of fact, by reducing the delay in picking up the patients in about 20 minutes (from 31.5 to 10 minutes), the waiting times for semi-urgent patients can be reduced by up to 10 hours. The effect is however less significant for red and orange patients with a reduction of approximately, 15 minutes and 1 hour and 20 minutes, respectively. *Therefore, contrary to Hypothesis 4, the transportation delay can have a major impact on the system congestion.* The underestimation of the transportation delay is due to the higher care given to red patients (i.e., those for which the improvement is the less significant). Such a considerable impact can be explained by the very high utilization levels that the surgery rooms currently present, which is already generating considerable waiting times for semi-urgent patients. If the patients arrive earlier to the surgery room, they might start their procedure on time and not lose the room to a patient with higher priority. A reduction of this order may also increase the available operating time, meaning that a short additional surgery could take place every day, increasing the total capacity of the EST.

Table 7: Waiting times in minutes as a function of the transportation delay (Ortho and MF)

	31'30	25'	20'	15'	10'	5'	1'
MF R	02:08	02:01	01:58	01:56	01:54	01:53	01:50
MF O	06:53	05:51	05:42	05:28	05:22	05:08	04:58
MF G	21:22	14:21	13:20	12:33	11:52	10:48	10:17
OP R	2:15	01:59	01:59	01:57	01:56	01:53	01:49
OP O	6:43	05:48	05:34	05:25	05:15	05:01	04:53
OP G	21:04	14:18	13:14	12:12	11:44	10:43	10:00

During the time spent in the surgery ward, it was evidenced that there were cases in which the patient transport to the surgery room started only when the previous patient had left the surgery room, or even when the surgery room cleansing is underway. Considering that the transport takes on average 30 minutes (without the delays in picking up the patient), a further improvement of the process could be achieved by transporting the patient at the same time that the previous patient is waking up. The combined effects of reducing the delay before the transport and making the transport process concurrent to the waking up of the previous patient could remarkably improve the performance measures of the surgery ward.

6. CONCLUDING REMARKS

In this paper, we demonstrated through different simulation scenarios that it is possible to achieve lower waiting times for all types of patients and subsequently prepare the hospital for increased demand, not only through increasing the amount of surgery rooms but also by improving the internal processes of the surgery ward, especially the transportation of patients inside the hospital. This was identified as an important bottleneck in the EST and is one that can be addressed in a much easier and less costly way than the other bottlenecks. By creating priority levels in patient transportation or hiring specific transportation aides for the EST and hence reducing the transportation delay, the hospital may increase its capacity up to 400 patients per year or reduce the waiting times for green-code patients up to 10 hours. This is a very relevant

finding, because most of the approaches in the literature focus on the clinical processes that take place in the surgery theater, but do not take into consideration how the logistical and managerial aspects can influence the patients waiting and processing times and the general performance of a surgery theater.

The results of this research have been partly implemented by the hospital. In particular, to solve the possible effects of the anticipated increase in the demand, the hospital opened a third surgery room, which will be increasingly used as demand keeps growing. It would be interesting to analyze if the actual impact of these changes, matches the predictions of the simulation model that we have presented here. An additional research perspective would be to analyze what would happen if the transportation delays were addressed with measures such as waiting rooms specifically for surgical patients. This research did not analyze the effects of adopting these measures along with the implementation of a hybrid or flexible policy in the hospital's surgery theater. Considering that a combined surgery theater would have 13 surgery rooms, this could be a very interesting path to explore. Finally, although the waiting times to surgery are important, future research would have to take into consideration the length of stay which would mean analyzing not only the EST and the arrival of surgical patients to the ED, but the patient flows in the whole hospital, from arrival, until they leave to home, including wards, as well as post-anesthesia and intensive care units.

7. REFERENCES

- Antognini, J. M. O., Antognini, J. F., and Khatri, V. (2015). How many operating rooms are needed to manage non-elective surgical cases? A Monte Carlo simulation study. *BMC Health Services Research*, 15(1):487.
- Bowers, J. and Mould, G. (2004). Managing uncertainty in orthopaedic trauma theatres. *European Journal of Operational Research*, 154(3):599–608.
- Cardoen, B., Demeulemeester, E., and Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3):921–932.
- Denton, B., Viapiano, J., and Vogl, A. (2007). Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science*, 10(1):13–24.
- Duma, D. and Aringhieri, R. (2019). The management of non-elective patients: shared vs. Dedicated policies. *Omega*, 83:199 – 212
- Ferrand, Y. B., Magazine, M. J., and Rao, U. S. (2014). Managing operating room efficiency and responsiveness for emergency and elective surgeries - a literature survey. *IIE Transactions on Healthcare Systems Engineering*, 4(1):49–64.
- Heng, M. and Wright, J. G. (2013). Dedicated operating room for emergency surgery improves access and efficiency. *Canadian Journal of Surgery*, 56(3):167–174
- Lamiri, M., Xie, X., Dolgui, A., and Grimaud, F. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research*, 185(3):1026–1037.
- Lane, D. C., Monefeldt, C., and Rosenhead, J. V. (2000). Looking in the wrong place for healthcare improvements: A system dynamics study of an accident and emergency department. *Journal of the Operational Research Society*, 51(5):518–531.
- Macario, Alex, M. M., Vitez, Terry, M., Dunn, Brian, B., and McDonald, Tom, M. (1995). Where Are the Costs in Perioperative Care?: Analysis of Hospital Costs and Charges for Inpatient Surgical Care. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 83(6):1138–1144.
- McDermott, K. W., Freeman, W. J., and Elixhauser, A. (2017). Overview of operating room procedures during inpatient stays in us hospitals, 2014: statistical brief# 233. *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. Rockville: Agency for Healthcare Research and Quality.
- Paul, J. A. and MacDonald, L. (2013). Determination of number of dedicated or's and supporting pricing mechanisms for emergent surgeries. *Journal of the Operational Research Society*, 64(6):912–924.
- Persson, M. J. and Persson, J. A. (2010). Analysing management policies for operating room planning using simulation. *Health Care Management Science*, 13(2):182–191.
- Riet, C. V. and Demeulemeester, E. (2015). Trade-offs in operating room planning for electives and emergencies: A review. *Operations Research for Health Care*, 7:52 – 69. ORAHS 2014 - The 40th international conference of the EURO working group on Operational Research Applied to Health Services.
- Samudra, M., Van Riet, C., Demeulemeester, E., Cardoen, B., Vansteenkiste, N., and Rademakers, F. E. (2016). Scheduling operating rooms: achievements, challenges and pitfalls. *Journal of Scheduling*, 19(5):493–525.
- Stepaniak, P. S. and Dexter, F. (2016). Constraints on the scheduling of urgent and emergency surgical cases: Surgeon, equipment, and anesthesiologist availability. *Perioperative Care and Operating Room Management*, 3:6 – 11.
- Van der Lans, M., Hans, E. W., Hurink, J. L., Wullink, G., Van Houdenhoven, M., and Kazemier, G. (2006). Anticipating urgent surgery in operating room departments. University of Twente, Tech. Rep. WP-158.
- Veen-Berkx, E. V., Elkhuizen, S. G., Kuijper, B., and Kazemier, G. (2016). Dedicated operating room for emergency surgery generates more utilization, less overtime, and less cancellations. *The American Journal of Surgery*, 211(1):122 – 128.
- Wullink, G., Van Houdenhoven, M., Hans, E., van Oostrum, J., van der Lans, M., and Kazemier, G. (2007). Closing emergency operating rooms improves efficiency. *Journal of Medical Systems*, 31(6):543–546.

CIGI 2021

Conception d'aménagement des pépinières forestières : Revue des pratiques et perspectives

Olivier BÉLANGER¹, Luc LEBEL^{1,3}, Nadia LEHOUX^{2,3}, Edith BROTHERTON^{2,3}

¹ Département des sciences du bois et de la forêt, UNIVERSITÉ LAVAL
Pavillon Abitibi-Price, 2405, rue de la Terrasse, Québec, G1V 0A6. Canada
Olivier.Belanger.6@ulaval.ca
Luc.Lebel@sbf.ulaval.ca

² Département de génie mécanique, UNIVERSITÉ LAVAL
Pavillon Abitibi-Price, 2405, rue de la Terrasse, Québec, G1V 0A6. Canada
Nadia.Lehoux@gmc.ulaval.ca
Edith.Brotherton.1@ulaval.ca

³ Consortium de recherche FORAC
Université Laval, Québec, Canada

Résumé – Les activités de reboisements sont au cœur des stratégies proposées par plusieurs nations pour lutter contre les changements climatiques. Les pépinières forestières sont fondamentales pour la production de plants servant au reboisement des forêts. Une revue de littérature spécifique à leur conception révèle toutefois un certain vide méthodologique associé au design et à l'opération d'un tel centre de production. Cet article propose donc une analyse des pratiques de conception associées au design d'usines, dans le but de définir comment de telles façons de faire pourraient être appliquées à la production d'un produit vivant (plant forestier). L'étude a notamment permis de mettre en relief une combinaison de méthodes quantitatives et qualitatives permettant d'apporter des pistes d'amélioration aux aménagements actuels. Le parallèle entre les deux domaines permet de proposer une nouvelle méthodologie visant la sélection et l'adaptation de méthodes pour la conception des pépinières. Cette nouvelle méthodologie guidera les gestionnaires de pépinière dans la sélection et l'adaptation d'une méthode appropriée en lien avec leurs objectifs de conception dans le but d'accroître leur productivité et leur performance.

Abstract – Tree nurseries are fundamental for reforestation, an activity that is adopted by several countries to mitigate climate change. Nevertheless, a literature review looking at the layout design of these production centers revealed a lack of design methods for this specific context. As a result, this article aims to analyze layout design approaches for the manufacturing sector to highlight best practices that could put in place in tree nurseries. The study pointed out that a combination of quantitative and qualitative approaches represents the most adapted way to improve the current system. Furthermore, the study reveals underlying link between those two domains allowing the proposition of a whole new methodology for the selection and the adaptation of appropriate design methods for tree nurseries. This methodology can be used to guide tree nursery managers to choose the appropriate method based on their design objectives to increase their productivity and performance.

Mots clés – Aménagement d'usines, pépinière forestière, processus de reboisement.

Keywords – Layout design, tree nurseries, reforestation process.

1. INTRODUCTION

Les pépinières sont des systèmes de production complexes qui visent à produire des plants forestiers de qualité répondant aux spécifications du reboisement des forêts naturelles. Lorsque la régénération est insuffisante dans les sites forestiers, le reboisement contribue à la remise en production du site, en plus d'être perçu comme un allié contribuant à la lutte contre les changements climatiques [Liao et al., 2010]. Ces plants doivent être produits en pépinière pour atteindre des cibles de qualité leur permettant de s'établir adéquatement dans leur futur milieu de vie. La production de plants forestiers de qualité requiert des travaux d'entretien et des régies de culture précises [Colas et al., 2003]. Ces travaux sont affectés par les processus de production, mais aussi par l'aménagement de la pépinière, car la conception d'un site a toujours un impact significatif sur la productivité [Don Taylor, 2007]. Il est probable que l'aménagement du site au sein d'une pépinière forestière a lui aussi un impact similaire sur la productivité. De ce fait, des gains en performance pourraient être possibles en adaptant aux besoins des pépinières certaines approches et méthodes de conception développées initialement pour le design d'usines. Le but de cette recherche consiste donc à mettre en évidence des méthodes de conception propres aux pépinières forestières, tout en présentant des possibilités de transpositions de méthodes de design d'usines au cas spécifique de production de plants forestiers. Pour ce faire, une revue de la littérature englobant la conception de pépinières forestières et l'aménagement d'usines a été réalisée. En considérant les particularités des pépinières et les méthodes existantes de design d'usines, une nouvelle méthodologie de conception des pépinières est mise de l'avant. Ce travail de synthèse a été réalisé en décrivant les caractéristiques observables selon l'échelle d'amélioration souhaitée. L'approche suggérée pourra être différente selon le niveau d'amélioration souhaité, variant d'une seule opération à l'ensemble de la pépinière.

Cet article est divisé de la façon suivante. Les bases méthodologiques permettant la réalisation de la revue de littérature seront d'abord présentées. Ensuite, les méthodes traitant de la conception d'aménagement pour les pépinières seront décrites à la section 2.1. Elles seront suivies des différentes catégories de méthodes existantes de conception d'usines à la section 2.2. Une analyse sera dès lors présentée concernant la sélection et la méthodologie d'adaptation de méthodes existantes pour traiter les cas relatifs aux pépinières à la section 3. Finalement, un bref retour sur les résultats et leurs perspectives d'applications concluront l'article à la section 4.

2. REVUE DE LA LITTÉRATURE SUR LES TECHNIQUES D'AMÉNAGEMENT : MÉTHODOLOGIE

Dans le but de mettre en relief les différentes techniques d'aménagement à déployer pour mieux supporter la production de plants forestiers, une revue de littérature a été réalisée. Les travaux se sont concentrés sur la production de plants forestiers en contenants, qui est plus utilisée au sein des pépinières québécoises, en plus de permettre un meilleur contrôle de l'environnement de croissance [Landis, 1995a]. Les différentes techniques de production de plants forestiers à racine nue ou en champs n'ont donc pas été prises en compte dans cette recherche. Cette revue a été menée selon la technique du cercle herméneutique [Boell et Cecez-Kecmanovic, 2010], en utilisant le chapitre écrit par Landis T. [1995b] comme point départ

puisqu'il propose quelques concepts qualitatifs fondamentaux reliés à la conception d'un aménagement pour une pépinière forestière. L'objectif de la revue visait à répondre aux deux questions de recherche suivantes :

- Existe-t-il des méthodes de conception d'aménagement particulières pour les pépinières forestières ?
- Quelles sont les principales méthodes d'aménagement d'usines développées au fil du temps ?

Pour répondre à ces questions, trois bases de données ont été principalement utilisées, soit : *Google Scholar*, *Springer link* et *ResearchGate*. Les mots clés utilisés étaient en anglais et comprenaient : « *tree nursery, factory design, plant production, layout optimization, greenhouse, operation, management* » ainsi que différentes combinaisons de ces derniers. Un total de 117 articles a été répertorié. Une lecture du résumé de chacun de ces articles a permis de classer ces derniers selon les thèmes couverts et d'évaluer leur niveau de pertinence en lien avec le projet. Ainsi, 10 articles principaux ont été retenus et jugés comme étant d'importance critique pour la réalisation du projet de recherche, en plus de 29 articles d'importance très élevée. Ces articles ont été révisés en profondeur et un résumé de leur avancée a permis de mieux identifier les points saillants pour répondre aux questions de recherche, tel que résumé dans les prochaines sous-sections.

2.1 Méthodes de conception d'aménagement pour les pépinières

Tout d'abord, il convient de proposer une terminologie permettant d'identifier les structures dédiées à la production des plants forestiers. Dans l'optique de se doter d'une appellation générale englobant tous les environnements de propagation et les environnements de croissance [Landis, 1995c], le terme d'environnement de culture sera utilisé pour les représenter de façon non distinctive. Il prend donc en compte les installations conçues pour la germination, la croissance, la propagation et l'aoûtement des plants, peu importe le degré de contrôle de l'environnement permis par la structure. Dans le contexte québécois, il est possible d'associer les tunnels, les serres et les aires de cultures extérieures à cette définition. Au Québec, la production d'un plant forestier en pépinière s'effectue en trois années distinctes. Chacune de ces années comprend différentes opérations. Certaines sont similaires telles que la régie des cultures tandis que d'autres sont spécifiques à certains moments de la production. Par exemple, l'ensemencement survient uniquement au début de la première année, alors que le classement survient uniquement à la troisième année, dans le but d'améliorer les cultures pour atteindre les critères de qualifications du client. La figure 1 présente un résumé de ces différentes étapes pour mieux saisir les bases de cette production.

La planification de la conception d'une pépinière forestière doit prendre en compte de nombreux critères. Certains de ces critères sont jugés indispensables lors de la sélection du site, alors que d'autres sont jugés comme étant désirables [Landis, 1995b]. Les critères indispensables sont l'accès au soleil, la qualité de l'eau, les sources d'énergie disponibles, la superficie utilisable du terrain et le contexte écopolitique. Chacun de ces critères impose différentes contraintes de proximité qui guideront le choix de la localisation pour une pépinière, mais aussi la disposition de chacune des aires d'activités du système. À ceci, il est possible d'ajouter des critères désirables lors de

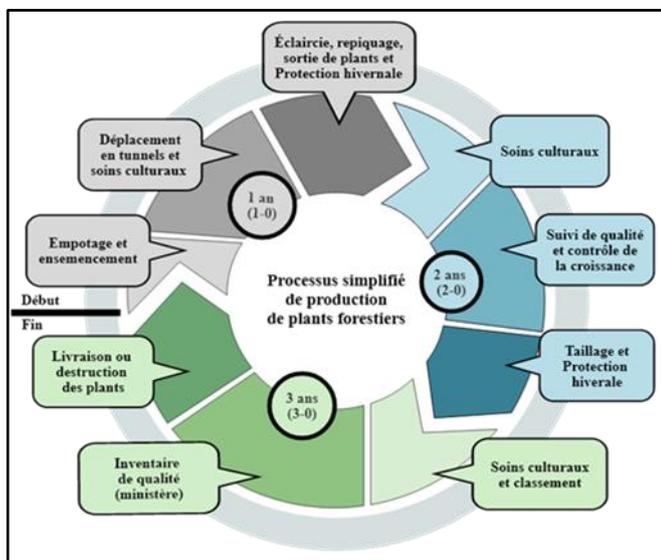


Figure 1. Représentation des trois années de production d'une pépinière incluant les processus simplifiés.

l'implantation d'une pépinière. Ces critères sont le microclimat, la topographie, la disponibilité de la main-d'œuvre, la distance des marchés et l'accessibilité [Landis, 1995b]. Pour intégrer davantage d'éléments reliés aux processus de production, certaines pépinières devront tenir compte de considérations supplémentaires qui influenceront la répartition spatiale.

Malgré le peu de publications attribuées à la conception des pépinières, un cas d'aménagement basé sur un regroupement par activités est présenté par Ingram et Vanek [2013]. Leur approche prend en compte le fait qu'une pépinière est caractérisée par de courtes opérations ayant lieu à un certain moment de la saison. En effet, les opérations sur les plants doivent avoir lieu en cohérence avec l'atteinte de différents stades phénologiques chez le plant tels que la germination, la croissance et l'endurcissement. Un regroupement des activités partageant des points communs tels que des intrants, des machines ou même du personnel a l'avantage de diminuer les activités n'ajoutant pas de valeur au produit [Ingram et Vanek, 2013]. Ainsi, une considération des activités lors de l'élaboration de principes de conception d'une pépinière permet de proposer de meilleures pratiques de gestion [Ingram et Vanek, 2013]. Il s'agit en fait d'un regroupement qui se rapproche d'un type d'organisation par processus [Montreuil et Lefrançois, 1996]. Par exemple, la préparation de mélange de tourbe et l'emportage peuvent être regroupés dans un même endroit [Ingram et Vanek, 2013]. Différentes notions d'adjacences, de proximité désirée et d'éloignement sont aussi considérées qualitativement. Par exemple, la salle des employés aurait avantage à se situer à proximité des bureaux tout en étant éloignée de l'entrepôt de pesticide. Il est à noter que l'approche proposée par Ingram et Vanek s'inspire considérablement du « *lean manufacturing* ». Ainsi, l'objectif est de diminuer les gaspillages de temps, d'intrants et d'espace, tout en augmentant les activités productives. À cet effet, une autre publication propose des moyens d'implanter ce mode de gestion dans les pépinières [Vanek et Ingram, 2014]. L'utilisation d'un 5S est également considérée pour mieux organiser l'environnement de travail et l'entreposage des équipements.

À la lumière des lectures, une pépinière forestière a souvent été considérée comme un centre produisant des matières premières destinées au reboisement. Cependant, la pépinière forestière devrait plutôt être perçue comme un centre processeur. Les

opérations faites sur les plants par les employés, mais aussi par les conditions environnementales sont utilisées à des fins de modification du produit pour en respecter les critères de classification du client, ce qui correspond à un centre processeur tel que l'entend Montreuil [2006]. Ainsi, une pépinière forestière réalise des opérations sur des semences selon les devis des clients pour en générer un plant conforme. Il est peu intuitif, mais juste d'affirmer que les conditions environnementales, au même titre qu'un employé, travaillent et possèdent la responsabilité d'activité productive sur le plant. Ceci explique en partie la vaste quantité de recherche pour optimiser les processus environnementaux dans les environnements de croissance afin de stimuler la productivité des cultures [Mazzanatti et al., 2015], [Gagnon. J. & DeBlois, 2017] et [Lamhamedi et al., 2020]. Ainsi, l'utilisation de l'environnement à des fins de production est bien documentée et a mené à la conception de diverses formes de structure utilisant différents matériaux pour régir les conditions environnementales appliquées aux produits [Landis, 1995c]. La plupart de ces processus environnementaux extérieurs sont difficiles à prévoir et les conditions environnementales peuvent générer des conditions défavorables aux plants. Certains centres de production iront même jusqu'à ignorer l'environnement extérieur pour se concentrer sur des cultures entièrement contrôlées, mais beaucoup plus coûteuses plutôt que de devoir assurer une gestion plus étroite du risque.

L'analyse quantitative des flux de produits dans différents postes de travail pour les opérations de classement avant l'expédition d'une pépinière a fait l'objet d'une étude pour permettre des améliorations [Gabriel-Santos et al., 2015]. Des récipients égaux contenant des plants de tailles différentes sont déplacés sur un convoyeur, alors que des employés effectuent un tri selon la taille en plus d'effectuer un contrôle de qualité. Les plants de tailles similaires sont replacés dans un même récipient selon trois catégories de hauteur. Les auteurs utilisent l'analyse de la conception axiomatique dans le but d'améliorer la productivité de cette opération. Ils introduisent différentes matrices pour transformer des paramètres de conception en besoins fonctionnels pour reconcevoir le système de convoyeur. Avec cette méthode, la capacité de production est augmentée de 133% et la productivité de 55% pour l'opération de classement et d'emballage des plants avant l'expédition [Gabriel-Santos et al., 2015]. L'utilisation de cette méthode permet de trouver l'origine des pertes de productivité et de proposer une conception évitant les pertes de temps. Une telle méthode est très avantageuse pour la considération d'opérations individuelles ou d'opérations regroupées partageant le même centre ou les mêmes ressources. Par exemple, un centre d'emportage et d'ensemencement pourrait certainement être amélioré avec cette méthode.

Les différents travaux recensés jusqu'à présent mettent en lumière des façons de concevoir l'aménagement d'une pépinière, sans toutefois aborder chacune des aires d'activités en détail ni présenter des schémas de conception détaillés pour améliorer une pépinière dans son ensemble. Les flux, lorsqu'ils sont évalués, sont considérés pour une seule opération. Pourtant, en pépinière, les opérations sont courtes, nombreuses et varient en fonction du moment de l'année, requérant souvent du matériel de manutention différent [Landis, 1995c]. Il devient alors pertinent d'évaluer les flux selon les différentes activités en cours à différents moments de l'année [Zhu et al., 2017]. Il semble donc intéressant d'explorer davantage les méthodes de conception d'aménagement issues du domaine manufacturier de

façon à en tirer de meilleures pratiques pouvant être adaptées au milieu des pépinières. Ces approches ont souvent été utilisées pour améliorer une entreprise selon diverses échelles, passant des opérations à une usine dans son ensemble. La prochaine sous-section aborde donc les principales méthodes de conception d'usines.

2.2 Principales méthodes de conception d'usines

Malgré quelques travaux réalisés en conception de pépinières, il semble qu'encore très peu de recherches aient abordé des méthodes de conception plus quantitatives d'aménagement pour ce secteur particulier [Eben-Chaïme et al., 2011]. D'autre part, des procédures ainsi que des approches permettant de développer des conceptions d'usines sont bien détaillées dans la littérature [Tompkins et al., 2010]. Dans l'objectif de résoudre concrètement le problème de disposition des installations (ou *facilities layout problems (FLP)*), diverses approches ont été développées au fil des ans et ont fait l'objet de plusieurs classifications [Singh et Sharma, 2005], [Moslemipour et al. 2011] et [Sharma et Singhal, 2016].

De telles méthodes impliquent typiquement d'identifier les caractéristiques de l'usine devant être conçue ou améliorée [Nordin et Lee, 2016]. La formulation mathématique du problème de conception peut ensuite être réalisée, puis résolue via des méthodes exactes, des méthodes heuristiques ou des approches intelligentes [Sharma et Singhal, 2016]. Ces catégories d'approches seront détaillées séparément, mais leur objectif est essentiellement le même, soit de minimiser la distance parcourue, les flux et les coûts de déplacement entre les centres, tel que présenté à l'équation 1 [Tompkins & al., 2010].

$$\text{Min } Z = \sum_{i=1}^m \sum_{j=1}^m f_{ij} c_{ij} d_{ij} \quad (1)$$

Où f_{ij} : sont les flux entre les centres i et j

c_{ij} : sont les coûts de déplacements entre i et j

d_{ij} : sont les distances de déplacements entre i et j

Les approches exactes reposent sur l'optimisation et la recherche de la meilleure solution d'aménagement possible. De nombreuses approches d'optimisation ont été développées et permettent d'obtenir une configuration idéale pour une usine. Le modèle BLOC [Montreuil, 1991] en est un exemple. La faiblesse principale de cette approche est qu'elle est limitée par le nombre de centres puisque le temps de résolution du problème devient rapidement très grand lorsque le nombre de centres à positionner augmente. Ainsi, diverses formulations de la représentation mathématique par programmation en nombres entiers mixtes (MIP) de Montreuil ont été proposées pour en augmenter la vitesse de résolution et le nombre de centres maximaux pouvant être traités [Meller et al., 1999]. Encore aujourd'hui, des travaux s'y intéressent pour rendre ces méthodes optimales applicables à des usines de plus grandes tailles [Anjos, 2021]. Cependant, il ne semble pas y avoir de modèles capables de garantir l'optimalité pour plus d'une vingtaine de centres [Brotherton, 2003], [Anjos, 2021].

Les approches par heuristique permettent la résolution de problèmes de conception complexes ayant un nombre élevé de centres, en assouplissant les contraintes formulées via les MIP ou par l'assignation quadratique (QAP) [Nordin et Lee, 2016]. Elles se distinguent en trois sous-catégories : les heuristiques de construction, les heuristiques d'amélioration et les métaheuristiques [Sharma et Singhal, 2016]. Les heuristiques de

construction telle que la méthode SLP [Muther, 1973], permettent de proposer des conceptions de départ pour une usine. La méthode SLP de Muther [1973] peut également être décrite comme une procédure puisqu'elle présente les étapes à suivre pour la conception d'une usine. Cette procédure se veut suffisamment flexible pour être aussi appliquée en tant qu'heuristique d'amélioration. Ces dernières représentent des méthodes adaptées pour proposer des conceptions alternatives d'une usine déjà existante. La méthode CRAFT est aussi un exemple d'heuristique d'amélioration [Amour & Buffa, 1963]. Elle a été appliquée récemment pour identifier de meilleurs arrangements de départements pour un centre de traitement de déchets aux Pays-Bas [Van Zanten, 2020]. D'un autre côté, les métaheuristiques peuvent tenir compte de l'aspect stochastique et proposent des algorithmes pouvant répondre à plusieurs objectifs simultanément. Par exemple, la méthode du recuit simulé est utilisée pour trouver la meilleure solution globale pour une usine possédant des centres inégaux [Hanagund & al., 2021]. Les différentes méthodes heuristiques sont plus flexibles que les méthodes exactes. Cependant, leur limite réside dans le fait qu'elles permettent d'obtenir uniquement des conceptions suboptimales pour une usine donnée.

Les approches intelligentes ont quant à elles, la capacité de résoudre des problèmes d'optimisation combinatoire difficiles. Ces approches ont la possibilité de résoudre le problème de disposition d'installations (FLP) malgré une discontinuité de l'espace, l'intégration de l'aspect stochastique ou la résolution d'une fonction *objectif* non linéaire. Par exemple, la logique floue [Dixit & al., 2020] et l'apprentissage machine [Tayal et al., 2019] ont été utilisés pour générer ou évaluer des conceptions afin de résoudre le FLP. Les métaheuristiques sont aussi parfois incluses dans les approches intelligentes [Moslemipour et al., 2011]. La flexibilité de ces approches n'est pas réservée à la conception d'usine. Divers algorithmes reposant sur les fondements des méthodes intelligentes ont permis de modéliser et d'améliorer différents aspects des conditions culturelles dans des serres. Par exemple, les réseaux de neurones ont été exploités pour mieux gérer la qualité et la productivité de fruits dans une serre hydroponique de tomates [Palacios & al., 2021]. La logique floue a été utilisée pour automatiser certains procédés cultureux tels que le contrôle de l'environnement et l'apport en eau dans une serre [Kurniawan, 2021]. Cependant, malgré le potentiel des méthodes décrites, aucune approche systématique ne semble avoir été utilisée pour résoudre le problème d'aménagement d'une pépinière forestière.

3. APPROCHES DE CONCEPTION ADAPTEES AU MILIEU DES PEPINIÈRES

Les travaux réalisés sur la conception des pépinières permettent de mettre en évidence quelques pistes d'amélioration. Les concepts généraux tels que des principes de conception de pépinière [Landis, 1995b] ou l'amélioration d'une opération telle que le classement des plants selon leur hauteur [Gabriel-Santos et al., 2015], permettent de poser des bases solides et confirment l'intérêt pour la conception revisitée de pépinières. Par ailleurs, aucune méthode de conception d'usines ne semble adaptée aux exigences spécifiques de la production d'un produit vivant générant des contraintes supplémentaires. Un certain vide méthodologique est constaté dans l'évaluation de la conception des pépinières forestières. Les méthodes traditionnellement utilisées dans les sciences managériales sont encore très peu appliquées à ce domaine d'activités [Eben-Chaïme et al., 2011]. Contrairement à une usine, le produit n'est

pas déplacé vers des machines pour y subir des opérations. Ce sont plutôt les ressources (eau, nutriments, personnels) qui sont déplacées vers les plants. Aussi, les opérations courtes nuisent à la mécanisation puisqu'il est coûteux d'investir dans des machines qui fonctionneront seulement quelques semaines à au plus deux mois par année. En plus, la responsabilité de certaines activités productives est détenue par l'environnement. Ce qui n'est pas le cas pour une usine traditionnelle. Même si les méthodes ne sont pas adaptées à la réalité des pépinières, des catégories d'approches peuvent être mises de l'avant en tant que point de départ d'une solution appropriée.

Pour débiter, des données sont nécessaires à la réalisation d'une conception d'aménagement. Les informations reliées à la demande et la disposition actuelle des centres doivent être recueillies. Les liens entre les différentes activités menant à la réalisation du produit de même que les différents systèmes de manutention permettent de déterminer les flux du système selon son état actuel. Pour une pépinière forestière, il est nécessaire d'ajouter à ces données la compréhension des processus physiologiques des végétaux en culture. Ensuite, il faut définir l'objectif d'amélioration, c'est-à-dire ce que le pépiniériste souhaite atteindre par l'aménagement ou le réaménagement de son site. La définition de cet objectif permet de choisir à quelle échelle il serait pertinent d'œuvrer pour l'atteindre. Par exemple, il est possible d'utiliser des méthodes de conception plus flexibles, mais plus complexes pour une conception intégrale. Inversement, des méthodes plus rigides et plus simples sont plus appropriées pour des cas tels que l'amélioration d'une opération. Ces diverses échelles d'amélioration peuvent être étudiées d'abord de façon isolée en les regroupant selon des catégories [Hyer et Wemmerlov, 2001], pour ensuite être intégrée à un tout pour atteindre des objectifs plus ambitieux.

À la figure 2, l'échelle d'amélioration est présentée de façon hiérarchique pour représenter la complexité croissante en fonction de la quantité d'opérations touchées. L'amélioration de

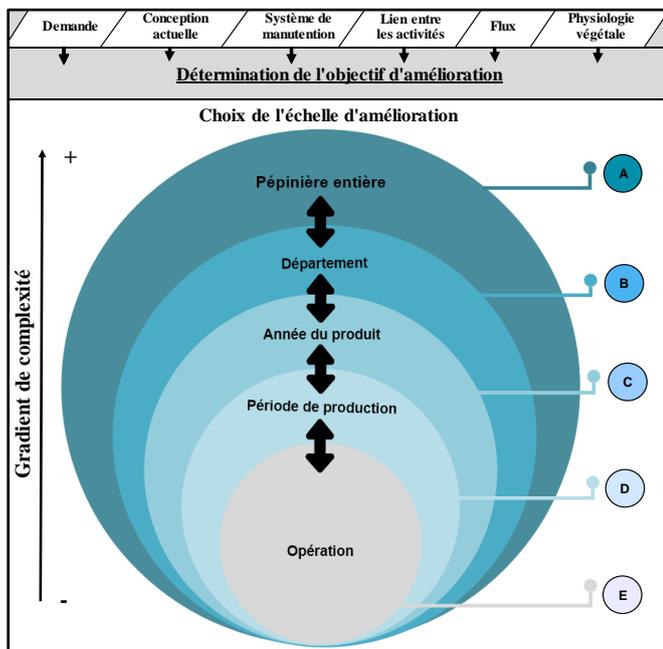


Figure 2. Représentation des données requises sous forme d'intrants et d'une l'échelle d'amélioration souhaitée (A, B, C, D et E)

la conception de la pépinière entière (A) représente le niveau de complexité le plus élevé. Il est possible d'améliorer la pépinière dans son ensemble en considérant l'intégralité des opérations des trois années régissant la production des plants. L'amélioration d'un département (B) représente le deuxième échelon en complexité puisque ce dernier a souvent des responsabilités dans chacune des années du produit et possède donc souvent des opérations qui peuvent se chevaucher. L'amélioration d'une année du produit (C) tel que l'ensemble des opérations encourues sur un plant d'un an (1-0) présente aussi un caractère dynamique affecté par de courtes opérations. Ces trois échelons de complexité doivent intégrer la nature stochastique du problème lié à la variabilité de la demande, à la probabilité de mort et de détérioration des plants, de même qu'à la variabilité au début et à la fin de plusieurs opérations. En effet, plusieurs départements et opérations sont touchés par les conditions météorologiques. Ceci concerne notamment le département responsable des soins cultureux comme l'irrigation et la fertilisation puisque la météo régit le besoin en eau des plants ainsi que les moments opportuns pour l'application de fertilisants permettant de stimuler la croissance. Certaines opérations sont conditionnées par l'état du produit aussi appelé stade phénologique. Par exemple, l'éclaircie (retrait des plants multiples dans une cavité d'un récipient) peut uniquement avoir lieu lorsque l'état de germination est terminé pour une culture donnée. Ainsi lorsque plusieurs opérations sont analysées conjointement, il faut inévitablement considérer leur caractère incertain.

La période de production (D) peut être considérée comme une façon de se concentrer sur une séquence de production d'une pépinière régie par les différents stades phénologiques pour une année de production donnée. Ces stades phénologiques peuvent être détaillés et sont reliés à la durée du jour et aux températures [Hamilton et al., 2016]. Pour certaines essences comme l'épinette blanche produite abondamment dans les pépinières québécoises, des calendriers détaillés existent et permettent de considérer davantage de catégories et différents stades phénologiques [Owens et al., 1977]. Cependant, ce sont plutôt les stades phénologiques globaux qui affectent les activités dans une pépinière. Ces derniers sont représentés au tableau 1, qui illustre les différentes périodes pouvant être considérées lors de l'élaboration de solutions de conception propre aux pépinières pour les années 1 à 3. Cette échelle d'analyse permet d'isoler les conditions générées par l'état du produit et de considérer des flux statiques pendant ces périodes puisqu'ils ne changent pas ou ils ne changent que très peu. Finalement, il est aussi possible de se concentrer sur l'amélioration d'une seule opération (E) de façon isolée. Il s'agit du degré de complexité le moins élevé.

Tableau 1. Tableau de différentes périodes régies par les stades phénologiques et les années de production.

Année	Stades phénologiques			
1 an (1-0)	Germination	Croissance (1-0)	Aoûtement (1-0)	Dormance (1-0)
2 ans (2-0)	Débourrement (2-0)	Croissance (2-0)	Aoûtement (2-0)	Dormance (2-0)
3 ans (3-0)	Débourrement (3-0)	Croissance (3-0)	Aoûtement (3-0)	-

L'étape suivante pour réaliser la conception d'une pépinière est la caractérisation de l'espace à modéliser en fonction de l'échelle d'amélioration choisie. À titre illustratif, la figure 3 présente le schéma de conception de l'ensemble du site d'une pépinière forestière de l'est du Québec. Il est possible d'y constater que la pépinière est séparée par une route publique et une voie ferrée, ce qui restreint le terrain disponible en plus de briser sa continuité. Le terrain est aussi adjacent à une rivière située du côté nord de la propriété. La partie sud de la propriété présente un sol rocailleux de faible profondeur, en plus d'une pente forte (non détectable sur la carte). La pépinière présentée comporte un total de 62 tunnels de production, 69 aires de cultures, huit bâtiments ainsi que des tunnels et des lieux dédiés à l'entreposage dispersés à différents endroits sur la propriété. Ce nombre de centres est trop élevé pour l'utilisation d'approches exactes sans l'utilisation de regroupements.

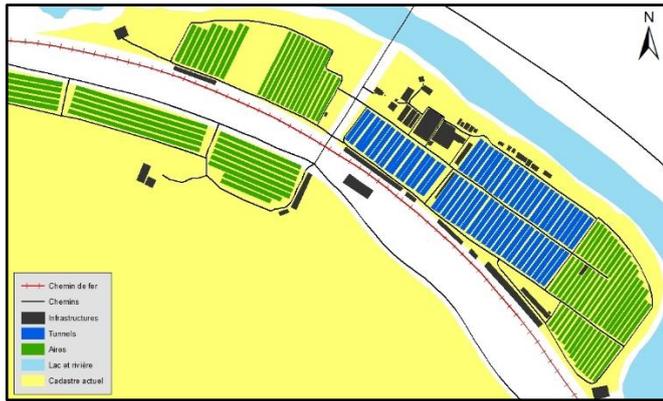


Figure 3. Schéma de conception actuel d'une pépinière forestière de l'est du Québec.

Le terrain appartenant à la pépinière est irrégulier, ainsi lorsque plus d'un centre est considéré, il devient donc préférable de modéliser le terrain sous la forme d'un espace continu, tel qu'exprimée par une formulation MIP [Montreuil, 1991]. Ce type de formulation permet d'éviter la division de l'espace sous forme discrète, soit en une multitude de carrés pour y positionner des centres dont les unités de surface représentent un nombre entier de ces carrés. La disposition des différents centres sur un terrain appartenant à la pépinière est plus difficilement représentable par ce type d'espace. Cependant lorsque des centres similaires sont considérés, tels que l'ensemble des tunnels de production, il demeure possible de représenter le terrain sous une forme discrète, qui est plus simple.

D'un autre côté, la forme des aires d'activités n'est pas égale. Par exemple, les aires de cultures reposent sur une largeur fixe entre 5,22 mètres et 6,96 mètres, tout en possédant une longueur variable limitée par la topographie du terrain et les limites de la propriété. Les tunnels de cultures possèdent une largeur fixe, mais une longueur pouvant varier d'une pépinière à l'autre par l'ajout ou le retrait d'arches, qui sont les composantes de la structure [Landis, 1995c]. Aussi, les lieux d'entreposages sont souvent à l'extérieur et n'ont pas une forme dédiée. Un modèle considérant plusieurs centres doit donc pouvoir être adapté pour représenter des centres ayant des tailles et des dimensions variables.

Les types de flux et leurs comportements doivent aussi être considérés lors de la réalisation d'une conception. Les flux peuvent être statiques, c'est-à-dire qu'ils demeurent les mêmes

tout au long de l'année puisque les opérations courantes ne changent pas énormément [Nordin et Lee, 2016], ou encore dynamiques. Dans ce cas, les opérations varient énormément selon la période de l'année [Nordin et Lee, 2016]. Idéalement, lorsque l'échelle d'amélioration cible l'ensemble de la pépinière, les départements ou une année complète de production, les flux devraient être considérés sous leur forme dynamique. Certaines opérations reliées à la production de plants ne peuvent être réalisées que dans une courte plage de temps. Les flux varieront énormément selon les opérations et les différentes méthodes de manutention à un moment donné du calendrier [Landis, 1995c]. En se basant sur les éléments caractérisant l'espace de la pépinière, il est possible de générer une caractérisation adéquate pour chacune des échelles d'amélioration telle que présentée au tableau 2.

Tableau 2. Caractéristique de la pépinière à concevoir selon l'échelle d'amélioration considérée (A, B, C, D et E).

Échelle d'amélioration	Caractéristique de l'usine à concevoir (pépinière)					
	Forme du terrain à modéliser	Type d'espace à privilégier	Forme des centres	Type de flux considérés	Comportement des flux	Complexité
A	Irrégulier	Continu	Non-égaux	Dynamique	Stochastique	Critique
B	Irrégulier	Continu	Non-égaux	Dynamique	Stochastique	Très élevée
C	Irrégulier	Continu	Non-égaux	Dynamique	Stochastique	Élevé
D	Irrégulier	Continu	Non-égaux	Statique	Déterministe	Modéré
E	Régulier	Discret	Souvent un seul	Statique	Déterministe	Faible

En se basant sur les méthodes existantes de design d'usines et l'approche globale utilisée lors de la résolution d'un problème de conception [Singh et Sharma, 2005], il est possible de générer une méthodologie de sélection d'une catégorie de méthode de conception selon la caractérisation de l'environnement précédemment effectué. La figure 4 présente une cartographie des flux de données menant vers l'élaboration, la sélection et l'adaptation d'une méthode propice à la résolution d'une conception de pépinière. La cartographie propose d'abord les échelons d'améliorations les moins complexes puisqu'ils sont généralement moins coûteux et plus faciles à implanter. Des intrants s'ajoutent à partir de certaines échelles d'amélioration pour représenter des aspects importants à considérer à ce niveau.

Lorsque l'amélioration cible une opération ⑤ de la pépinière et que cette dernière ne requiert pas la considération de plus de 20 centres [Anjos, 2021], l'utilisation d'une approche exacte donnant la meilleure solution possible est recommandée. L'utilisation de la conception axiomatique telle qu'utilisée pour une opération de classement [Gabriel-Santos, 2015] est aussi une avenue envisageable. L'intégration de nouvelles technologies est directement reliée aux opérations et c'est pourquoi cet intrant s'insère à cette étape. En divisant le problème de conception en période de production ④, il demeure possible d'éviter l'utilisation de flux dynamiques et d'utiliser des approches heuristiques pour proposer des conceptions sous optimales.

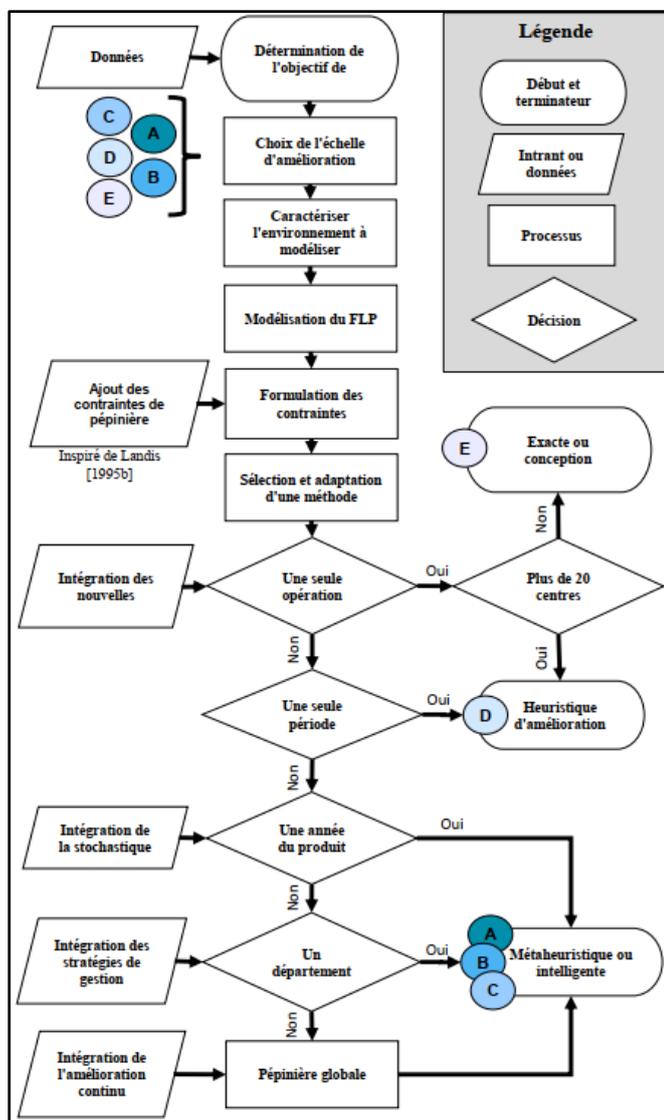


Figure 4. Cartographie des flux de donnée guidant la sélection d'une catégorie de méthode selon les échelles d'amélioration (A, B, C, D et E)

Finalement, pour représenter l'espace irrégulier, la nature stochastique de cette production et la considération de flux dynamiques variant selon les périodes de l'année, les approches métaheuristiques pourraient s'avérer prometteuses [Nordin et Lee, 2016]. C'est pourquoi l'amélioration de la pépinière entière (A), d'un département (B) ou d'une année de production (C) repose sur cette catégorie d'approche. Des approches métaheuristiques et intelligentes intégrant l'aspect stochastique existent [Moslemipour et Lee, 2011]. De plus, parmi les différentes approches permettant de résoudre un problème d'aménagement dynamique, les métaheuristiques sont les plus utilisées [Sharma et Singhal, 2016]. Même si la catégorie de méthode est la même, il faut s'attendre à un gradient de complexité. L'amélioration globale d'une pépinière reposera sur des algorithmes plus complexes intégrant plus d'éléments que l'amélioration d'un département. Aussi, les départements possèdent des responsabilités sur les trois années de production en cours. Leur amélioration sera plus sophistiquée qu'une séquence d'opération considérée dans une seule année de production.

4. CONCLUSION

Les pépinières sont des centres permettant la production d'arbres vivants, qui sont de nature complexe. La conception de tels centres s'est avérée peu étudiée dans la littérature. Le but de cet article était donc de combler ce vide en réalisant une revue sur les méthodes de conceptions existantes pour les pépinières et en créant des liens avec des méthodes de conception d'usines plus largement documentées. Pour ce faire, différentes bases de données ont été consultées pour y déceler quelques méthodes qualitatives et quantitatives mises de l'avant pour améliorer les pépinières. L'analyse de méthodes de conception d'usines a permis de faire ressortir des points communs avec les approches dédiées aux pépinières, mais aussi, des pistes d'adaptation. Les articles examinés semblent suggérer que la méthode de conception utilisée dépend énormément de la complexité du cas considéré. En unifiant les deux domaines, il est possible de proposer un cadre théorique de sélection d'une méthode en introduisant une échelle d'amélioration qui doit être sélectionnée pour atteindre un objectif de conception spécifique. Cette échelle d'amélioration comporte un gradient de complexité pouvant influencer la sélection d'une méthode de conception appropriée. La synthèse entre les caractéristiques d'une pépinière reliées aux catégories de méthodes issues de la conception d'usines permettra notamment aux pépiniéristes de mettre le doigt sur une méthode de conception idéale pour atteindre leurs objectifs de production. L'application de ce cadre théorique pour améliorer la productivité et la performance d'une pépinière fera l'objet d'études futures notamment pour la résolution du cas d'une pépinière québécoise.

5. REFERENCES

- Anjos M. F. & Viera. 2021. Mathematical optimization approach for facility layout on several rows. *Optimization Letters* 2021. 15. Pp. 9-23. doi: <https://doi.org/10.1007/s11590-020-01621-z>.
- Armour G. C., Buffa E. S. (1963). A heuristic algorithm and simulation approach to relative location of facilities, *Management science*, 9(2), 294-309.
- Boell S. K. & Cecez-Kecmanovic, D. (2010). Literature Reviews and the Hermeneutic Circle. *Australian Academic & Research Libraries*, 41(2), 129-144. doi : 10.1080/00048623.2010.10721450
- Brotherton E. (2003). Modélisation et optimisation d'aménagement d'usines avec structuration de l'espace par zones, Essai (MBA), Université Laval. 120 p.
- Colas F., Gagnon J., Gingras B-M., Labbé L., Lamhamedi M. S., Thiffault N., Tousignant D., Tremblay F., Bettez M. & et Tourigny M. (2003). La filière de production de plants au Québec : de la semence à la plantation, *Le Naturaliste Canadien*, 127(2), 77-84.
- Dixit, V., Verma, P. & Raj, P. 2020. Leveraging tacit knowledge for shipyard facility layout selection using fuzzy set theory, *Expert Systems with applications* (158). doi <https://doi.org/10.1016/j.eswa.2020.113423>.
- Don Taylor G. (2007). *Logistic Engineering Handbook* 1st Edition. CRC Press. ISBN: 084933053X. 640p.
- Eben-Chaïme M., Bechar A. & Baron A. (2011). Economical evaluation of greenhouse layout design, *Department of Industrial Engineering and Management*, University of Ben Gurion, Sheva, Israël.
- Gabriel-Santos A., Gonçalves-Coelhob A-M., Santosc P., Fradinhod J. & Mourãoe A. (2015). Redesigning a

- manufacturing system based on functional independence: the case of a tree nursery, *Procedia CIRP*, (34), 206 – 211.
- Gagnon, J. & DeBlois, J. (2017). Effects of foliar urea fertilization on nitrogen concentrations of containerized 2+0 Jack pine seedlings produced in forest nurseries. *Tree Planters' Notes*, 60(1), 44-50.
- Hamilton, J. A., El Kayal, W., Hart, A. T., Runcie, D. E., Arango-Velez, A., & Cooke, J. E. K. (2016). The joint influence of photoperiod and temperature during growth cessation and development of dormancy in white spruce (*Picea glauca*). *Tree Physiology*, 39, 1432-1448. doi:10.1093/treephys/tpw061.
- Hunagund I.B., Pillai V.M., Kempaiah U.N. (2021) Solving Unequal Area Facility Layout Problems with Flexible Bay Structure by Simulated Annealing Algorithm. *Advances in Production and Industrial Engineering (2021)*, 75-101, doi: https://doi.org/10.1007/978-981-15-5519-0_7.
- Hyer N., Wemmerlov U. (2001). Reorganizing the Factory: Competing Through Cellular Manufacturing, *Productivity Press*, ISBN: 1563272288, 874p.
- Ingram D. L. & Vanek S. J. (2013). [En ligne]. Sustainable Production Systems: Efficient Wholesale Nursery Layout, disponible à: <http://www2.ca.uky.edu/agcomm/pubs/HO/HO109/HO109.pdf> Consulté le 10 février 2020.
- Kurniawan D. & Witanti A. (2021). Prototype of Control and Monitor System with Fuzzy Logic Method for Smart Greenhouse. *Indonesian journal of information. Systems*, 3(2), 116-127. doi: <https://doi.org/10.24002/ijis.v3i2.4067>.
- Lamhamedi, M.S., Renaud, M., Auger, I. & Fortin J.A. 2020. Granular calcite simulates natural mycorrhization and growth of white spruce seedlings in peat-based substrates in forest nursery. *Microorganism* 8 (7), 19p. doi: doi:10.3390/microorganisms8071088.
- Landis T. (1995). (a). Ch 1: Initial planning and feasibility assessments in The Container Tree Nursery Manual: Volume 1, Nursery planning, development, and management. *U.S.D.A. Forest Service, Agriculture Handbook*, 674p.
- Landis T. (1995). (b). Ch 2: Nursery site selection in The Container Tree Nursery Manual: Volume 1, Nursery planning, development, and management. *U.S.D.A. Forest Service, Agriculture Handbook*, 674p.
- Landis T. (1995). (c). Ch 3: Nursery Design and Layout in The Container Tree Nursery Manual: Volume 1, Nursery planning, development, and management. *U.S.D.A. Forest Service, Agriculture Handbook*, 674p.
- Liao C., Luo, Y. Fang C., & Li B. (2010). Ecosystem Carbon Stock Influenced by Plantation Practice: Implications for Planting Forests as a Measure of Climate Change Mitigation, *PLoS ONE*, 5(5), e10867. doi: 10.1371/journal.pone.0010867.
- Mazzanatti T., Calzavara A. K., Pimenta J. A., Oliveira H. C., Stolf-Moreira R. & Bianchini, E. (2015). Light acclimation in nursery: morphoanatomy and ecophysiology of seedlings of three light-demanding neotropical tree species, *Brazilian Journal of Botany*, 39(1), 19–28, doi:10.1007/s40415-015-0203-5.
- Meller R.D., Narayanan V. & Vance P.H. (1999). Optimal Facility Layout Design, *Operations Research Letters*. (23), 117-127, doi: [https://doi.org/10.1016/S0167-6377\(98\)00024-8](https://doi.org/10.1016/S0167-6377(98)00024-8).
- Montreuil B. (1991). A modelling framework for integrating layout design and flow network design, *Progress in Material Handling and Logistics*, Springer-Verlag, 95-116, doi: https://doi.org/10.1007/978-3-642-84356-3_8.
- Montreuil B. & Lefrançois P. (1996). Organizing Factories as Responsibility Networks. *Progress in Material handling*, R. J. Graves & al., Material Handling Institute, Braun-Brumfield Inc., Ann Arbor, Michigan, U.S.A., 375-411.
- Montreuil B. (2006). Facilities network design: a recursive modular protomodel based approach. *Chapter in Progress in material handling research*.
- Moslemipour, G., & Lee, T. S. (2011). Intelligent design of a dynamic machine layout in uncertain environment of flexible manufacturing systems. *Journal of Intelligent Manufacturing*, 23(5), 1849–1860. doi:10.1007/s10845-010-0499-8
- Moslemipour G., Lee T. S. & Rilling, D. (2011). A review of intelligent approaches for designing dynamic and robust layouts in flexible manufacturing systems, *The International Journal of Advanced Manufacturing Technology*, 60(1-4), 11-27. doi:10.1007/s00170-011-3614-x.
- Muther R. (1973). Systematic layout planning, *Van Nost. Reinhold*, ISBN: 0442264097, 360 p.
- Nordin N. N., Lee L.S., (2016). [En ligne]. Heuristics and metaheuristics approach for facility layout problems: a survey. Disponible à: https://www.researchgate.net/publication/327592010_Heuristics_and_metaheuristics_approaches_for_facility_layout_problems_a_survey.
- Owens, J. N., Molder, M., & Langer, H. (1977). Bud development in *Picea glauca*. I. Annual growth cycle of vegetative buds and shoot elongation as they relate to date and temperature sums. *Canadian Journal of Botany*, 55(21), 2728–2745. doi:10.1139/b77-312.
- Palacios D., Arzamendia M., Gregor D., Cikel K., Leon R. & Villagra M. (2021). Definition and application of a computational parameter for the quantitative production of hydroponic tomatoes based on artificial neural networks and digital image processing. *AgriEngineering*, 3(6), 1-18. doi: <https://doi.org/10.3390/agriengineering3010001>.
- Sharma, P., & Singhal, S. (2016). A review on implementation of meta-heuristic approaches for layout problems in dynamic business environment. *International Journal of Multivariate Data Analysis*, 1(1), 6. doi:10.1504/ijmda.2016.081076.
- Singh, S. P., & Sharma, R. R. K. (2005). A review of different approaches to the facility layout problems, *The International Journal of Advanced Manufacturing Technology*, 30(5-6), 425–433. doi:10.1007/s00170-005-0087-9.
- Tayal, A., Kose, U., Solanki, A., Nayyar, A. & Sucedo, J.A.M. 2019. Efficiency analysis for stochastic dynamic facility layout problem using meta-heuristic data envelopment analysis and machine learning, *Computational intelligence* 2019 (1). p. 1-31. doi: 10.1111/coin.12251
- Tompkins, J.A., White, J.A., Bozer, Y.A. and Tanchoco J.M.A. (2010). Facilities Planning, 4th Edition, John Wiley & Sons, ISBN: 0470444045. 864p.
- Van Zanten, G.J. (2020). Facility design to improve the internal logistics at Van Gerrevink. Mémoire, Maitrise en génie industriel et gestion, Université de Twente, Pays-Bas.
- Vanek S.J. & Ingram D.L. (2014). Principles and Approaches for Optimizing Efficiency in Nursery and Landscape Businesses. Disponible à https://uknowledge.uky.edu/anr_reports/159. Consulté le 10 février 2020.
- Zhu, T., Balakrishnan, J., & Cheng, C. H. (2017). Recent advances in dynamic facility layout research, *INFOR: Information Systems and Operational Research*, 1–29, doi:10.1080/03155986.2017.1363591.

Factors influencing Drug Consumption and Prediction Methods

DENIS KOALA, ZAKARIA YAHOUNI, GÜLGÜN ALPAN, YANNICK FREIN

Univ. Grenoble Alpes, CNRS, Grenoble INP¹, G-SCOP, 38000 Grenoble, France

{denis.koala; zakarya.yahouni; gulgun.alpan; yannick.frein}@grenoble-inp.fr

Résumé

Les problèmes de logistique hospitalière, de gestions de stocks et d'estimations des besoins de produits de santé se posent encore de nos jours dans les centres hospitaliers qui doivent conjuguer budgets restreints et satisfactions de patients. De nouvelles perspectives devraient s'imposer pour résoudre les problèmes en amont, surtout la prédiction des besoins. Ce travail se plonge dans la littérature afin de mettre en évidence les méthodes de quantification et d'estimation des besoins des produits de santé dans les établissements de santé. Un second objectif est de dresser une liste des facteurs qui impactent la consommation des médicaments, notamment les facteurs qui sont souvent utilisés dans les méthodes de quantification. Cette revue de la littérature identifie six méthodes utilisées par les praticiens des quatre coins du monde en tenant compte de certains prérequis et de type de données. De même, 34 facteurs sont identifiés et regroupés en trois catégories. Ces résultats devraient permettre de mettre sur pied de nouveaux outils de prédiction des besoins en médicaments, de faciliter en amont le dimensionnement de nouveaux entrepôts pharmaceutiques et de résoudre des problèmes de logistique hospitalière.

Mots clés - Consommation de médicament ; facteur ; quantification ; prédiction ; logistique hospitalière.

Abstract

Estimating the needs of healthcare products and inventory management are still challenging issues in hospitals nowadays. Centers are supposed to cope with tight budgets and patient satisfaction at the same time. Some issues can be tackled in advance, especially regarding the prediction of drug consumption needs. This work delves into the literature in order to highlight existing methods of quantifying and estimating the needs for drugs in health facilities. A second objective is to draw up a list of factors that impact drug consumption in particular, factors that are used in these prediction methods. Following this literature review, it appears that six sustainable methods are being used by practitioners around the world, taking into account certain prerequisites and types of data. Thirty-four factors are identified as well and grouped into three categories. These results should participate in setting up new tools for predicting the need of drugs, to facilitate the upstream dimensioning of new pharmaceutical warehouses and to solve some hospital logistics issues.

Keywords: Drug consumption; factor; quantification; prediction; hospital logistic

1 INTRODUCTION

In France, as in other developed countries, the healthcare sector is constantly transforming to adapt to a complex and changing environment with increasingly tight budgets. OECD² reports [OECD, 2019; OECD, 2015; OECD, 2013] reveal that over the period 2000-2008, average per capita health expenditure per year increased by 4.1%, 1% from 2008 to 2013 and 2.4% from 2013 to 2018. In such a context, hospital centers are obliged to optimize their management. They have launched several series of investment programs in order to achieve this, while guaranteeing the safety and quality of the care provided to patients. The Hospital Plan from 2007 to 2012 in France [Safon, 2017] is a striking example. This has shown limits, as [Bernardini-Perinciolo et al, 2018] states. Hospital expenditures represent about 40% of healthcare expenditures [OECD, 2013], of which 30 to 46% are related to logistics activities, pharmaceutical logistics in particular [Nachtmann et al, 2009; Poulin, 2003], making hospital logistics the second largest component of expenditure after personnel expenses [Poulin, 2003; Volland et al, 2017]. Pharmaceuticals account for 70-

80% of supply costs [Rego and al 2014; Kelle, 2012]. There is therefore an interesting potential for gains.

Compared to other sectors such as industry, hospital logistics has not been considered a priority in the research world in the past. Several reasons have been pointed out such as legal aspects, the high complexity of supply chains, the efficiency of patient treatment, the stochastic and unpredictable nature of product demands as well as the importance of the human factor, etc. [Beier, 1995 ; Almarsdóttir et al, 2005 ; Moons et al 2019 ; Romero, 2013 ; Volland et al, 2017]. The research potential is therefore interesting. Lack of innovation in hospital logistics leads to problematic situations in hospital departments regarding inventory management, unjustified forecasting techniques, lack of IT support, etc. [Rachmania et al 2013; Romero, 2013].

Over the last twenty years, logistics has been identified as a key lever to manage healthcare costs [Dacosta-Claro et al 2002; De Vries, 2011]. Fruitful research has been conducted and methods have been implemented to optimize and curb certain deficiencies, inventory management problems in particular (chronic and generalized shortages, surpluses or overstocking,

¹ Institute of Engineering Univ. Grenoble Alpes

² OECD: Organization for Economic Co-operation and Development

etc.), supply chain problems, etc. [De Vries, 2011 ; Jurado et al 2016 ; Battersby 1993 ; Kelle et al, 2012 ; Jurado et al 2016 ; Maestre, 2018].

Some research estimates that through effective logistics management, approximately half of the costs associated with hospital logistics can be eliminated [Poulin, 2003; Romero, 2013; Volland et al, 2017] although the literature on this issue is still rare [Kharroubi, 2019]. A prerequisite for an efficient stock management is to have a correct dimension of the stock. This is a strategic problem and needs to be decided when the pharmaceutical warehouses are to be built.

In fact, setting up a new logistics warehouse for a hospital pharmacy can take several months, with corrective actions in case of incorrect dimensioning. To ensure the most accurate dimensioning and to guarantee the performance of the pharmacy supply chain within a hospital structure, it would be necessary to implement decision support tools that are both fast and efficient. The dimensioning and inventory management are strongly influenced by the consumption behavior of pharmaceutical products, which in turn is spurred by several factors [Kharroubi, 2019]. This work is positioned in this perspective.

In this work, we highlight the existing literature methods that aim to facilitate the quantification and estimation of drug needs from the 1980s to the present day. This cannot be done without taking into account the causal links and factors that impact consumption. The way in which health product quantification is carried out has serious implications for the health system of a country, a region, even the operation of a health facility. If projected estimates are lower than real needs, the impact can seriously hamper the effective delivery of needed health care: similarly, if the estimated quantities are excessive, the result is a waste of scarce resources. Wasteful or irrational drug use can be perpetuated by simply continuing to order products based on information from the historical use of each drug [WHO, 1988; MSH, 2012]. Significant shortages and overstocking can occur by ordering on the basis of theoretically determined quantities that have not been sufficiently tested. It is obvious that the process of estimating drug needs should be done by considering a certain number of indicators and factors of consumption. Additional complexity arises when estimating drug needs when dealing with exceptional situations such as disasters, epidemics or pandemics like COVID-19. These phenomena are difficult to model and predict. They are considered in certain estimation methods by adjustment [MSH, 2012]. However, the effectiveness is not very accurate, especially when the availability of health products is not the sole responsibility of hospitals (eg. : Masks in COVID-19 situation). In this work we do not consider such extreme and rare situations.

Section 2 and 3 are respectively dedicated to the research methodology, followed by a description of the quantification methods found in the literature with a brief comparison. Section 4 addresses the factors influencing drug consumption. As concluding remarks, a research perspective is issued, regarding hospital and pharmaceutical logistics in particular.

³ WHO: World Health Organization
<https://www.who.int/home/search>

⁴ MSH: Management Sciences for Health
<https://www.msh.org/search>

2 METHODOLOGY

2.1 Identification of publications

To identify relevant literature, we proceeded in several steps. The first step consisted in using the google scholar and direct science platforms to identify appropriate keywords for successful searches. Based on that, three classes of keywords were drawn up as shown in Table 1. The first-class groups together terminologies related to the quantitative and predictive aspect of needs, while the second class is an addition to the first class with a view to contextualizing and centralizing research. Depending on linguistic affinity, terms such as "drug" or "medicine" can be found in the literature but are terms to designate the same entity. The last class of keywords consists in highlighting consumption indicators and the work that is carried out in the context of stock sizing and management. A second search on the same databases was conducted for relevant keywords, which defined our second step. Obviously at this stage some keywords were excluded like "hospital", "influence", "stock inventory", "stock capacity" which for some are too broad and others a bad combination. In the third step a cross-referencing of the keywords in the platforms, google scholar, direct science, web of science and IEEE was carried out. Although the results were more or less satisfactory, we used some databases such as WHO³ and MSH⁴. The search is limited to documents in English and French over the period 1980 to 2019. Although we initially focused solely on articles, the request for official reports proved to be a plus, particularly on quantification methods.

Once the papers were grouped together, the fourth step was to eliminate duplicates. Finally, in the last step we eliminated papers that do not address our problem. These different steps resulted in 81 papers related to quantification methods and 59 related to drug consumption factors.

Class 1	Class 2	Class 3
"Estimating"	"Drug utilization"	"Inventory management"
"Forecasting"	"Essential drugs"	"Stock Inventory"
"Quantification"	"Pharmaceutical requirements"	"Hospital logistic"
"Prediction"	"Medicine consumption"	"Stock capacity"
"Quantifying"	"Drug consumption"	"Capacity planning"
"Forecast demand"	"Hospital pharmacy"	"Warehouse management"
	"Hospital"	"Influence"
		"Factor"

Table 1: Classes of Keywords

2.2 Scope

As a reminder, pharmaceuticals include items directly related to patient care, essential drugs, medical consumables, labile blood products, food, linen, sterile items, etc. In France, healthcare products are grouped into 20 groups⁵. The paper focuses on

⁵ French Website of health products categories
https://www.legifrance.gouv.fr/affichCodeArticle.do?jsessionid=8569371CC949E72C401CFBC44FDFFF32.ipdila16v_3?idArticle=LEGIARTI000033897163&cidTexte=LEGITEXT000006072665&categorieLien=id&dateText

essential drugs and the factors that influence their consumption in hospitals.

3 METHODS FOR DRUG NEEDS QUANTIFICATION

The effective introduction of an action plan for the consumption of essential drugs and vaccines requires a good supply management system. This system is composed of selection, quantification, procurement, distribution and use [WHO, 1988; MSH, 2012]. The selection consists of determining which essential drugs are needed. It is a process that begins by defining a list of common diseases for each level of health care. It has a considerable impact on the quality of care and the cost of treatment, so it is one of the pillars of an effective health system. Quantification is the process of estimating the quantities and costs of drugs and health products needed for a specific time period and determining when shipments of these products must be delivered to ensure an optimal and uninterrupted supply.

Procurement is the process of choosing suppliers, placing and monitoring orders, controlling the quantity and quality of medicines and paying suppliers. Then comes distribution, which includes reception, storage, inventory, transport and record keeping, i.e. information collection and control. Prescribing, dispensing and use of medicines; patient compliance concerns the use phase of the system.

In the context of our research, we are interested in the quantification process. Even though some of the methods that result from it include selection, based on the principle that the drugs must first be determined by the health institution or organization concerned.

Needs are estimated according to a given context, and the analysis must include contextual factors. It should be noted that, according to WHO, the effectiveness of quantification depends on the quality of available information and resources and accurate data on morbidity and drug use.

According to [MSH, 2012], quantification is not a simple calculation, it is the first step in the procurement process. There are many aspects to quantification such as: Calculation of estimated order quantities, costs and delivery dates of shipments; Planning, mobilizing and securing financial resources; Estimation of storage requirements; Assessing the rational use of products; Facilitate coordination of procurement and logistics with donors, suppliers, health facilities and other stakeholders; Inform manufacturers and suppliers about future demand for manufacturing, procurement and logistics management decisions, etc.

The quantification process is normally applied for: calculation of quantities ordered for public procurement, estimate budgetary requirements, develop purchasing quantities for new programs, expand procurement quantities for scale-up programs, etc. (see [WHO,1988] and [MSH, 2012] for more details).

Since common sense does not allow us to deal with certain complexities in quantification, methods have been developed since the 1980s, the identified ones are:

- Consumption method
- Morbidity method
- Proxy consumption methods
- Service-level projection of budget requirement
- Hospital Request method
- Population-based method

The methods most encountered in the literature are the first three in the list [Soeters, 1988; WHO, 1988; Osore, 1989; MSH,

2012]. Drugs could be quantified by using one or a combination of these approaches. Those methods are normally used to quantify needs for an annual or semiannual procurement. We will briefly explain each one of them in the following paragraphs.

3.1 Consumption method

This method is based on a facility's existing drug consumption. In fact, a list of all drugs eligible to be ordered or purchased is prepared and a more accurate inventory of past consumption, usually over a recent period of six to twelve months, is used to calculate the quantities required for each drug for each facility. An analysis of facility consumption is necessary and whenever it appears abnormal for a given drug, a correction is made by adjusting it upwards or downwards until it reaches an adequate level. The adjustment is already made according to stock shortages in order to obtain the average monthly consumption. Then, the average monthly consumption is multiplied by the number of months to be covered by the purchase. Safety stock levels and delivery times (in months) are also multiplied by the projected monthly consumption. These three figures are added together to obtain the gross requirements for the period, subtracting the usable stock on hand and any stock on order from the gross estimate, to deduct the quantity to be purchased. A further adjustment is then made for losses. And if cost is a factor in the quantification then the expected unit cost for each is multiplied by the number of units to be purchased to obtain the expected purchase value for the total quantity which is an addition of the purchase values of each drug on the list. Finally, a final adjustment is made if this cost is higher than the budget. This is why the method is also called the adjusted consumption method [WHO, 1988]. In the 1980s, WHO proposed three main steps to take into account the approach to quantify the drugs to be made available over the forecast period. In 2012, the MSH group developed 11 steps that included estimating the total cost of forecasting with the Quantimed tool (see [MSH, 2012] for more details).

In order to use this approach, preconditions must be met [WHO, 1988]. Accurate and reliable consumption data that are available or can be obtained relatively easily should be available. The same applies to essential drug lists with their packaging and prices. Losses due to expiration, damage and theft should not be excessive. Also, the supply of medicines in "typical" facilities should be enough (in practice, there has been no shortage of essential medicines for more than three months of the year). There should also be a record of suppliers' delivery times and budgets to be allocated.

While this method of quantification is inadequate for new hospitals, it is generally preferable for stable programs where funding, pharmaceutical management and prescribing are reasonably satisfactory. It is also easier to use in environments such as hospitals with many health problems and complex treatments.

3.2 Morbidity method

Based on the number of people suffering from a given disease at a given time in a given population, according to prevalence or incidence, the so-called morbidity method was established. This method uses data on patient health facility attendance and morbidity to predict the need for drugs based on assumptions about how problems will be treated. In this sense, it requires data sets: a list of common health problems, a list of essential

medicines that includes therapy for the problems, and a standard treatment package for quantification either on the basis of standard treatment regimens agreed upon for each defined health problem or on the basis of current practice.

Although very demanding in terms of data, this method is generally preferable for new or rapidly changing services or when services are radically reorganized. It is also preferable if prescribing practices are costly and irrational, as it provides a systematic basis for improvement.

To carry out its application and to obtain the most accurate estimate of the quantity of products and their costs, [MSH, 2012] presents this method with 12 steps in the process from specifying lists to comparing costs and budget to making corrections.

3.3 Proxy consumption method

Previous approaches require data that is often very precise, and some systems are faced with information deficits. In this way they are not feasible, and the proxy consumption method presents itself as an alternative. In this method, consumption data are extrapolated from a database in another health system or another region with similarities. For this reason, it is used for new facilities that may lack data.

This method can be based on two approaches. The first population-based approach is to define drug use per 1,000 inhabitants and the second service-based approach is to establish drug use by specific cases or by inpatient admissions of patients. It is also advisable, for a complete quantification, to use a combination of the two approaches.

For its application, 9 steps are established for successful quantification, from the selection of the standard system for comparison and extrapolation, the definition of product lists to the estimation of the costs of each and total drug [MSH, 2012]. It is a method that is quite concerned with the indicators and factors of consumption. Also, standard facilities should closely resemble the region or country for which the estimate is made as a main condition to get a good prediction. The resemblance should be in terms of geography and climate, patient population served, morbidity patterns, prescribing practices, standard treatment guidelines, essential medicines lists and pharmaceutical supply status.

3.4 Service-level projection of budget requirement

The approach outlined here is generally used to estimate financial needs, not specific quantities of drugs, for the purchase of pharmaceuticals on the basis of costs per patient treated at different levels of the same health system or on the basis of data from other health systems.

Like the proxy method of consumption, this method is an extrapolation method that leads to rough estimates given the large variations that may exist between the target health system and the system used as the standard data source. Sources of error encountered include: prescribers in the target system using a different drug combination than in the source system, variability in disease frequency and patient visits per facility, and differences in the effectiveness of procurement and financial management systems in the two settings. It does not anticipate specific drug needs but provides a clear and logical rationale for pharmaceutical funding requirements [MSH, 2012]. For its application, the main requirement of this method is a fairly reliable estimate of the average cost of medicines per patient and the average number of visits at different levels of the standard health system. The data that must be compiled are: the average number of curative and non-curative outpatient visits and bed days and/or other types of patient contact for each type of facility in the original health care system; and the average cost per curative and non-curative outpatient visit and bed day and/or other type of patient contact in each type of facility in the original health care system.

3.5 Hospital request method

This method is considered one of the oldest approaches for quantifying and estimating needs. It is a method of calculating drug requirements that is based on hospital requests for national, regional or warehouse pharmacies. The free choice of drugs to individual prescribers could be mentioned as an advantage of this method [Soeters, 1988; WHO, 2014]. We can also add the reduction of planning needs and guarantee a set of products will be available but are limited in the type and quantity of medicines and lead to significant waste and stock-outs. However, problems may include: hospital requests for more drugs than necessary, inadequate hospital requests (taking into account the cost/therapeutic ratio) or unnecessary drugs through personal preferences or the influence of pharmaceutical companies, health care staff often do less than optimal prescribing practices (polypharmacy, over-prescribing and unnecessary prescribing) [Soeters, 1988].

3.6 Population-based method

The last method that is elucidated in this paper is the population-based method which is based on the prevalence of various diseases in the population. The population-based method depends on total need that is different from demand. Information that is needed includes demographic and disease monitoring data e.g. number of new patients, number of continuing patients, those needing to change to a different treatment, in addition to disease prevalence rates among the target population. A formula is required for each product which can be unrealistic for medicines with multiple uses [WHO, 2014]. However, as can be expected, this theoretically interesting method is not used in practice because the required data are very difficult to obtain, implying very few references in the literature [Soeters, 1988]. Among the data we can note socio-economic data such as age, sex, marital status, co-morbidity, geographical position of patients. These difficulties were even more real in the first years of application of the approach because computer tools were lacking in health structures and data tools were not yet on the policy agenda.

3.7 Comparison of methods

Based on these findings, a comparison has been made in this document by highlighting the advantages and limitations of each method summarized in Table 2.

Among the six methods highlighted in the literature, the most widely used are the consumption, morbidity and proxy consumption methods. The reasons that can be put forward are the availability and specificity of the data that fall within the scope of the application of quantification and the factors that can have an impact on consumption in health facilities.

The method considered in the literature to be the most accurate for quantifying the use of pharmaceuticals is the consumption-based approach. To apply this method, data should be complete, accurate, and correctly adjusted for out-of-stock periods and anticipated changes in demand and use. However, the main problem with this method is that it does not normally consider the relevance of past consumption patterns, which may or may not correspond to public health priorities and needs regarding morbidity. This can lead to irrational drug use. If stock-outs have been widespread for long periods of time, it may not be possible to apply this method accurately, which is why capturing actual demand is essential for the most accurate approach.

Quantification based on morbidity is the most complex and time-consuming method to apply, according to what is said in

	Advantages	Disadvantages
<i>Consumption method</i>	<ul style="list-style-type: none"> – Detailed morbidity data and standard treatment regimens are not required. – It requires less detailed calculations. – It is useful for institutions such as hospitals, where there are many problems and complex drug treatments. – It is reliable if consumption is well recorded and stable and is not likely to differ significantly from the current supply. – It identifies inventory management problems and encourages improvements. – Easy method, requiring only one study by a few people. – Improves the availability of essential medicines and improves cost-effectiveness 	<ul style="list-style-type: none"> – Must have accurate consumption data. – Reliable data on pharmaceutical consumption can be difficult to obtain, especially in new or rapidly changing services. – It does not provide a detailed or systematic basis for reviewing drug use and improving prescribing; if the prescribing pattern is unsatisfactory and is not corrected, this method risks perpetuating it. – It is unreliable if there have been long stock-outs (more than three months) or significant loss or waste of medicines. – It is not conducive to good morbidity recording. – Does not interact with the existing information system. Previous drug shortages can lead to biased results. – Existing (non-optimal) prescribing practices are accepted. – Changes in demand and use including irrational use of medicines
<i>Morbidity method</i>	<ul style="list-style-type: none"> – No need for pharmaceutical consumption data; the method can be used for new services that do not have this data. – Based on a rational prescribing system, it provides a systematic basis for reviewing drug use and prescribing, particularly at the primary care level where drug treatments are less frequent and simpler. – Improves prescribing habits and interacts with the existing information system. Improves the availability of essential drugs, reduces the consumption of non-essential drugs and improves cost-effectiveness. – Promotes reliable recording of morbidity. 	<ul style="list-style-type: none"> – Problems can arise with detailed morbidity data and agreed-upon standard treatment regimens. – It requires more detailed calculations. – Results may differ significantly from the actual drug supply. – Supply will not match utilization if standard treatments are not followed. – Assesses only the quantities needed to treat patients; losses and wastage must be considered separately. – Exact hospitalizations are difficult to predict
<i>Proxy consumption method</i>	<ul style="list-style-type: none"> – Not demanding in terms of the amount of data on morbidity and pharmaceutical consumption 	<ul style="list-style-type: none"> – Comparability of facilities, morbidity patterns and treatment practices between sites/countries, – Incomplete or inaccurate consumption data from the reference facility.
<i>Service-level projection of budget requirement</i>	<ul style="list-style-type: none"> – Reliable estimate of average medicine cost per patient – No need for consumption details and morbidity data 	<ul style="list-style-type: none"> – Does not forecast needs for specific medicines, – Variable facility use, – Attendance treatment patterns, – Supply system efficiency.
<i>Hospital Request method</i>	<ul style="list-style-type: none"> – The free choice of drugs by prescribers, especially for centralized health systems (countries, regions, etc.). 	<ul style="list-style-type: none"> – Hospitals are requesting more medication than necessary, knowing that the quantities will be reduced anyway. – Hospitals request drugs that are inappropriate (cost/therapeutic) or unnecessary because of personal preferences or the influence of pharmaceutical companies. – Health care staff often adopt suboptimal prescribing practices – Expensive, non-essential drugs are required. – Non-optimal prescribing practices are accepted. – Prescribers are sensitive to the marketing practices of pharmaceutical companies and consumer pressure.
<i>Population-based method</i>	<ul style="list-style-type: none"> – Interesting for drugs that do not have multiple uses 	<ul style="list-style-type: none"> – Data are very difficult to obtain, especially patient socioeconomic information – Difficult to quantifying medicines with multiple uses

Table 2: Methods advantages and limits

the literature and the type of data input. Indeed, in many countries, it is very difficult to collect valid morbidity data for several diseases; therefore, some needs will be neglected in quantification. Data on patient consultations are often incomplete and inaccurate, and it is difficult to predict what percentage of prescribers will actually follow the standard treatment regimens used for quantification. Despite these constraints, this method may still be the best alternative for procurement planning or for estimating budgetary needs in a system or health facility. The method is also often used to estimate the cost of drugs in a large number of settings, such as

a small primary care system, a specialty hospital, or a new hospital.

Proxy consumption is the method generally used when neither the consumption-based nor the morbidity-based method is feasible. This method is more likely to provide accurate projections when used to extrapolate one set of facilities from another set of facilities serving the same type of population in the same type of geographic and climatic environment. If the method is applied using standard data from another country, the results will only be rough estimates. Even when both target and health facilities are similar, it is very difficult to assume that

disease incidence, patterns of use, and prescribing habits will be essentially the same in both cases. Nevertheless, this method may be the best alternative in the absence of appropriate data required for the consumption or morbidity-based method. The surrogate consumption method is also useful for cross-checking projections made with other methods. All of the methods reported are based on a number of factors and indicators related

to drug use. Therefore, a determination of the factors that influence consumption should contribute to both a judicious choice of method and the most accurate quantification.

4 FACTORS

Categories of factors	N°	Factors	Sources
Socio-demographic and Socioeconomic	1	Gender (Sex)	[Pappa et al, 2008 ; Jiménez-Rubio et al, 2010 ; Rocío et al, 2018 ; Mishuk et al, 2018 ; Howard et al, 2018 ; Henricson et al, 1998 ; Sarwar et al, 2017 ; Borrell et al, 2010]
	2	Age	[Pappa et al 2008 ; Jiménez-Rubio et al, 2010 ; Rocío et al, 2018 ; Mishuk et al, 2018 ; Howard et al, 2018 ; May et al, 1974 ; Henricson et al, 1998 ; Pokela et al, 2014]
	3	Race/ethnicity/country of birth	[Mishuk et al, 2018 ; Howard et al, 2018 ; Henricson et al, 1998]
	4	Income(annual)	[Jiménez-Rubio et al, 2010 ; Matin et al, 2015 ; Mishuk et al, 2018 ; Howard et al, 2018 ; Henricson et al, 1998]
	5	Marital/Civil status	[Pappa et al 2008 ; Rocío et al, 2018 ; Sarwar et al, 2017]
	6	Education	[Pappa et al 2008 ; Jiménez-Rubio,2010 ; Henricson et al, 1998 ; Pokela et al,2014]
	7	Residence(rural/urban)	[Pappa et al 2008 ; Matin et al, 2015 ; Henricson et al, 1998]
	8	Employment or Activity status (employed, retired, unemployed, inactive)	[Jiménez-Rubio, 2010 ; Henricson et al, 1998 ; Sarwar et al, 2017]
	9	Social class	[Rocío et al, 2018 ; Borrell et al, 2010]
Health-related	10	Health insurance	[Jiménez-Rubio, 2010 ; Howard et al, 2018]
	11	Health risk or Lifestyle (smoking; alcohol consumption, obesity; Body Mass)	[Rocío et al, 2018]
	12	Chronic conditions/chronic diseases	[Pappa et al 2008 ; Jiménez-Rubio et al, 2010 ; Howard et al, 2018]
	13	Comorbidity	[Howard et al, 2018]
	14	Prior experience with drug use	[Howard et al, 2018]
Facility related & staff member	15	The drug price index	[Berndt, 2002 ; Pappa et al 2008]
	16	Number of physicians for consultations and hospitalization	[Pappa et al 2008;]
	17	The number of hospital beds	[Pappa et al 2008 ; Matin et al, 2015]
	18	The number of medical visits	[Matin et al, 2015]
	19	Morbidity patterns (frequency of health problem)	[Matin et al, 2015;]
	20	Number of doses of each medicine per day	[Muhia et al, 2017 ; Henricson et al, 1998 ; Sarwar et al, 2017]
	21	Geographic area (region, district)	[WHO, 1988 ; Alkan et al, 2015]
	22	Facilities geographic position	[WHO, 1988; Alkan et al, 2015]
	23	Patient contact for each category of facility	[WHO, 1988]
	24	Service levels (Services (bed- day))	[Muhia et al, 2017]
	25	Number and type of health facilities	[Muhia et al, 2017]
	26	Seasonal factors (and replacing an older medicine)	[MSH, 2012]
	27	Periods (lead time, time to review)	[MSH, 2012]
	28	Medical Specialist (health workers characteristics)	[Mishuk et al, 2018; MSH, 2012]
	29	Storage point (capacity)	[Julius et al, 2018 ; Rocío et al, 2018]
	30	Procurement processes	[MSH, 2012]
	31	Budgeting processes	[Julius et al, 2018]
	32	Legal requirements	[Julius et al, 2018]
	33	New programs, expansion of existing services	[Julius et al, 2018]
	34	Technologies of facilities	[Howard et al, 2018]

Table 3: Factors affecting drug consumption in hospitals

In the literature, there are several variations of factors that are considered in the quantification of pharmaceutical needs. A number of papers [Pappa et al 2008 ; Jiménez-Rubio et al, 2010; Rocío et al, 2018; Mishuk et al, 2018 ; Howard et al, 2018 ; May et al, 1974; Henricson et al, 1998; Pokela et al, 2014] refer to factors related to patients, to the population to be taken into account in the healthcare system, generally socio-economic factors related to the medical system. Various studies have been carried out, some for specific products and others for product categories, taking into account the classification of the 29 WHO categories, in all four corners of the world. To summarize, these factors have been subdivided into three categories and presented in Table 3.

The first category includes socio-economic factors such as: gender, age, origin of the patient, annual income, social class, and so on. Some studies conducted in the USA on the correlation between drug use and race and origin are not applicable in all countries. In the second category comes the factors related to the health of the patient. These include factors such as co-morbidity, health insurance, lifestyle (smoking; alcohol consumption, obesity: Body Mass, etc.).

The last category includes factors related to health care institutions and personnel.

All of these factors could be parameters for estimating drug requirements for a new hospital structure, or for good inventory management. It shall be noted that some factors are not applicable in all health systems, for example, race or country of birth, which are considered to be impacting factors in the United States [Mishuk et al, 2018], may not be significant in another country. Some studies related to consumption factors found in the literature were less generalized. For instance, some factors focus on specific drugs and types of products such as generic drugs [Howard et al, 2018], ARVs, tuberculosis, etc. Also, among the factors in this list some may not be possible to integrate in a prediction model if the related data is not available.

5 CONCLUSION AND PERSPECTIVES

Looking forward to the well-being of patients, hospitals are facing several challenges. Among these challenges, reducing the cost of logistics, which represents the second most important expense, and reducing the risks of poor pharmaceutical inventory management seem to be the priorities. Of course, over the last twenty years, solutions have been proposed to curb this phenomenon as optimally as possible, but beyond the logistics and inventory management aspect, the question arises about dimensioning new warehouses with new services related to product management. This dimensioning process can take several months and lead to unfortunate incidents that can affect the quality of care, a luxury that organizations cannot afford. Determining dimensioning parameters implies knowledge of existing methods in terms of quantifying and estimating pharmaceutical needs, as well as the factors that influence these needs.

The objective of this paper has been to review these methods and the factors influencing the consumption of pharmaceutical products. Six quantification methods have emerged, three of which are the most widespread, such as consumption and morbidity methods, from the 1980s to the present day, proposed by academics as well as practitioners in the health sector who have recognized the potential of their applications. The application of these approaches is based on a set of preconditions. In this sense, the determination of consumption

factors was essential. A set of 34 factors were therefore listed and grouped into three categories.

As this paper is the first in a series, the determination of these factors will allow future work to contextualize them in the French health system. The resulting factors should be studied in order to prove the effectiveness and correlations with the consumption of drugs and health products in general. Robust parameters for the dimensioning and management of stocks should result from this perspective. Moreover, the estimation methods found in the literature do not make use of Big Data techniques. It would be appropriate to propose tools for predicting the needs of hospital centers, both by functional unit and by health establishment, along the lines of artificial intelligence and big data techniques. Such are the prospects of our future work.

6 ACKNOWLEDGMENTS

This work has been partially supported by MIAI@Grenoble Alpes, (ANR-19-P3IA-0003).

7 REFERENCES

- Alkan, Ö., & Elmali, K. (2015). The Determination of the Factors Affecting Pharmaceutical Consumption in Respect to Provinces in Turkey. *Mediterranean Journal of Social Sciences*, 6(6), 356.
- Almarsdóttir, A. B., & Traulsen, J. M. (2005). Cost-containment as part of pharmaceutical policy. *Pharmacy world and science*, 27(3), 144-148.
- Battersby, A., & Garnett, A. (1993). *How to estimate warehouse space for drugs* (No. WHO/DAP/93.3. Unpublished). World Health Organization.
- Beier, F. J. (1995). The management of the supply chain for hospital pharmacies: a focus on inventory management practices. *Journal of Business Logistics*, 16(2), 153.
- Bernardini-Perinciolo, J., & Bertolucci, M. (2018). Quand la sur-rationalité provoque la sous-rationalité. *Revue internationale de psychosociologie et de gestion des comportements organisationnels*, 24(58), 91-109.
- Berndt, E. R. (2002). Pharmaceuticals in US health care: determinants of quantity and price. *Journal of Economic perspectives*, 16(4), 45-66.
- Borrell, C., Muntaner, C., Gil-González, D., Artazcoz, L., Rodríguez-Sanz, M., Rohlfs, I., ... & Álvarez-Dardet, C. (2010). Perceived discrimination and health by gender, social class, and country of birth in a Southern European country. *Preventive medicine*, 50(1-2), 86-92.
- Dacosta-Claro, I. (2002). The performance of material management in health care organizations. *The International journal of health planning and management*, 17(1), 69-85.
- De Vries, J. (2011). The shaping of inventory systems in health services: A stakeholder analysis. *International Journal of Production Economics*, 133(1), 60-69.
- Henricson, K., Stenberg, P., Rametsteiner, G., Ranstam, J., Hanson, B. S., & Melander, A. (1998). Socioeconomic factors, morbidity and drug utilization an ecological study. *Pharmacoepidemiology and Drug Safety*, 7(4), 261-267.
- Howard, J. N., Harris, I., Frank, G., Kiptanui, Z., Qian, J., & Hansen, R. (2018). Influencers of generic drug

- utilization: a systematic review. *Research in Social and Administrative Pharmacy*, 14(7), 619-627.
- Jiménez-Rubio, D., & Hernández-Quevedo, C. (2010). Explaining the demand for pharmaceuticals in Spain: Are there differences in drug consumption between foreigners and the Spanish population? *Health Policy*, 97(2-3), 217-224.
- Julius, K., Tenambergen, W., & Mapesa, J. (2018). Factors Influencing Management of Pharmaceutical Supplies at Meru Teaching and Referral Hospital, Kenya.
- Jurado, I., Maestre, J. M., Velarde, P., Ocampo-Martínez, C., Fernández, I., Tejera, B. I., & del Prado, J. R. (2016). Stock management in hospital pharmacy using chance-constrained model predictive control. *Computers in biology and medicine*, 72, 248-255.
- Kharroubi, Y. (2019). *Étude de la consommation des médicaments dans une unité hospitalière en s'appuyant sur les techniques du Big Data*. Grenoble : Bibliothèque de l'Institut National Polytechnique de Grenoble.
- Kelle, P., Woosley, J., & Schneider, H. (2012). Pharmaceutical supply chain specifics and inventory solutions for a hospital case. *Operations Research for Health Care*, 1(2-3), 54-63.
- Maestre, J. M., Fernández, M. I., & Jurado, I. J. C. E. P. (2018). An application of economic model predictive control to inventory management in hospitals. *Control Engineering Practice*, 71, 120-128.
- Matin, B. K., Azami, S. R., Mahmoudi, S., Rezaei, S., Shaahmadi, F., & Karyani, A. K. (2015). Determinants of pharmaceutical expenditures of urban households: A time series study in Kermanshah province (Iran). *Electronic physician*, 7(7), 1470.
- May, F. E., Stewart, R. B., & Cluff, L. E. (1974). Drug use in the hospital: Evaluation of determinants. *Clinical Pharmacology & Therapeutics*, 16(5part1), 834-845. doi:10.1002/cpt1974165part1834.
- Mishuk, A. U., Qian, J., Howard, J. N., Harris, I., Frank, G., Kiptanui, Z., & Hansen, R. (2018). The association between patient sociodemographic characteristics and generic drug use: a systematic review and meta-analysis. *Journal of managed care & specialty pharmacy*, 24(3), 252-264.
- Moons, K., Waeyenbergh, G., & Pintelon, L. (2019). Measuring the logistics performance of internal hospital supply chains—a literature study. *Omega*, 82, 205-217.
- MSH. (2012). *MDS-3: Managing Access to Medicines and Health Technologies*. Arlington, VA: Management Sciences for Health.
- Muhia, J., Waithera, L., & Songole, R. (2017). Factors affecting the procurement of pharmaceutical drugs: A case study of Narok County Referral Hospital, Kenya. *Med Clin Rev*, 3(4), 20.
- Nachtmann, H., & Pohl, E. (2009). The state of healthcare logistics. *Cost and quality improvement opportunities*. Center for innovation in healthcare logistics, Arkansas.
- OECD (2019), *Health at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, <https://doi.org/10.1787/4dd50c09-en>.
- OECD (2015), *Fiscal Sustainability of Health Systems : Bridging Health and Finance Perspectives*, Éditions OCDE, Paris, <https://doi.org/10.1787/9789264233386-en>.
- OECD (2013), *Health at a Glance 2013: OECD Indicators*, OECD Publishing, http://dx.doi.org/10.1787/health_glance-2013-en
- Osores, H. (1989). Estimating drug requirements using morbidity data-based method: cumulative country experience. *Tropical doctor*, 19(2), 90-94.
- Pappa, E., Kontodimopoulos, N., Papadopoulos, A. A., Tountas, Y., & Niakas, D. (2008). Pmp15 Determinants Of Pharmaceutical Consumption In A General Population. *Value in Health*, 11(6), A366-A367.
- Pokela, N., Bell, J. S., Lihavainen, K., Sulkava, R., & Hartikainen, S. (2010). Analgesic use among community-dwelling people aged 75 years and older: a population-based interview study. *The American journal of geriatric pharmacotherapy*, 8(3), 233-244.
- Poulin, É.(2003). Benchmarking the hospital logistics process: a potential cure for the ailing health care sector. *CMA Management*; Vol. 77, No.1 :20-23.
- Rachmania, I. N., & Basri, M. H. (2013). Pharmaceutical inventory management issues in hospital supply chains. *Management*, 3(1), 1-5.
- Rego, N., Claro, J., & de Sousa, J. P. (2014). A hybrid approach for integrated healthcare cooperative purchasing and supply chain configuration. *Health care management science*, 17(4), 303-320.
- Rocío, D. R. D., Valentín, H. B., Isabel, J. T., & Pilar, C. G. (2018). Factors Associated to Medication Consumption Among the Immigrant Population Residing in Spain. *Journal of Immigrant and Minority Health*, 20(4), 909-919.
- Romero, A. (2013). Managing medicines in the hospital pharmacy: logistics inefficiencies. In *proceedings of the world congress on engineering and computer science* (Vol. 2, pp. 1-6).
- Safon, M. O. (2017). Les réformes hospitalières en France. *Aspects historiques et réglementaires, Synthèse documentaire, IRDES*, 57.
- Soeters, R., & Bannenberg, W. (1988). Computerized calculation of essential drug requirements. *Social Science & Medicine*, 27(9), 955-970.
- Sarwar, M. R., Atif, M., Scahill, S., Saqib, A., Qamar-uz-Zaman, M., & Babar, Z. (2017). Drug utilization patterns among elderly hospitalized patients on polypharmacy in Punjab, Pakistan. *Journal of pharmaceutical policy and practice*, 10(1), 23
- Volland, J., Fügener, A., Schoenfelder, J., & Brunner, J. O. (2017). Material logistics in hospitals: a literature review. *Omega*, 69, 82-101.
- WHO. (1988). *Estimating drug requirements: A practical manual* (No. WHO/DAP/88.2. Unpublished). World Health Organization.
- WHO. (2014). *Regional workshop on strengthening quantification and procurement of essential medicines* (No. SEA-DRUGS-165). WHO Regional Office for South-East Asia.
- WHO. (1988). *Estimating drug requirements: A practical manual* (No. WHO/DAP/88.2. Unpublished). World Health Organization.

Conversational Bots with Reinforcement Learning for Multi-criteria Suppliers Selection Focused on post-COVID-19 Risk: Case of French Automobile Manufacturer

SAMIA CHEHBI GAMOURA^{*1}, DAVID DAMAND^{*2}, MARC BARTH^{*3}, HALIL IBRAHIM KORUCA⁴

^{*} EM Strasbourg Business School, Strasbourg University, HuManiS (UR 7308), France

samia.gamoura@em-strasbourg.eu

damand@unistra.fr

marc.barth@unistra.fr

⁴Suleyman Demirel University, Isparta, Turkey

halilkoruca@sdu.edu.tr

Abstract – The supplier selection problem has always been under the spotlight of research interests from academics and industrials. After the COVID-19 outbreak, the vulnerability of supply has hugely affected the partnerships between the chain's stakeholders. The risk and volatility of suppliers have massively increased accordingly. This paper proposes a combination of the Multi-Criteria Decision Analytics (MCDA) method founded on ranking through the conversational bots technology. Reinforcement Learning is also used to bias the reward system's risk factor and improve the selection by reducing the risk impact. The approach associates the MCDA ranking methods' merits in evaluating the benefits of bots' collective intelligence through adaptation to changes. Validation experiments are conducted by comparing our model with an existing traditional approach. An industrial case study of a French automobile manufacturer serves for a preliminary scenario for further experimentations. The pilot results demonstrate the outperformance of the proposed approach.

Keywords – Supplier Selection, Multi-criteria Decision Analysis, Conversational bots, Reinforcement Learning.

1 INTRODUCTION

Supply Chain (SC) is a form of networked relationships involving the movement of flows between all the stakeholders from the source to the consumers (Chehbi-Gamoura 2019). These relationships are often more complicated than normal relations (Chen, et al. 2020), including the complex process of supplier selection. They often involve challenging situations that need advanced tools (Kucharavy et al., 2020). The literature also reveals that the supplier selection problem is one of the most challenging research topics discussed extensively in supply chains (Chien, Lin and Lin 2020). However, the common motives shared by these studies are often to look for success and sustainability, such as (Bai et al. 2019), (Durmic 2019), (Javad, et al. 2020), and (Gupta, Soni et Kumar 2019).

The examination of the existing research exposes the quality of goods and services as the unique indicator through which the model of selection focuses on the customers' satisfaction (Vijayakumar, et al. 2019). However, the quality is just one among a range of criteria that interact and impact the selection. The selection process has different aspects that may reverse the final decision's influence (Liu, Gao et Ma 2019). For example, the supplier's reputation is whether he has a 'good' or 'bad' service and relationships (Arioğlu, et al. 2020) historically. On the other hand, the literature indicates that only rare publications have considered the changing effect of these indicators, the risk impact, and the volatility (Kaya, et al. 2019).

The supplier selection process is considered a multi-criterial problem, which mainly depends on the customer assessment, which, in turn, biases the preferences and priorities (Bai et al.

2019). Moreover, in any partnership, both partners (supplier and customer) base their decisions on multiple criteria, such as quality, price, lead-time, brand, reputation, marketplace, etc. (Chien, Lin and Lin 2020). Before giving any analysis, we have to keep in mind that the supplier-customer partnerships and the supplier selection process are strongly associated with the suppliers' risk (Bai et al. 2019). Customers in supply chains, particularly in manufacturing, consider the risk assessment priority, mainly where suppliers' risk is too high, as the manufacturing in heavy industries and healthcare supply chains.

The COVID-19 outbreak is causing, until today, a critical vulnerability in the relationships between suppliers and customers (Sharma, et al. 2020). The risk assessment impacts all the chain, including the supplier selection processes being under the focus (Ivanov et Dolgui 2020). The lockdown politic and the closing of borders between countries have handicapped the fulfillment of goods and services and hugely affect the suppliers, causing instability (Ghadge, et al. 2020).

According to the motives above, the proposed approach focuses on the risk extensively increased due to such instability.

The additional challenging issue is that customers' priorities in the decision process are not stable; they are subject to continuous changes. The continuous changes engender, in fact, an additional complexity in the decision process for both vendors and buyers (Hadian, et al. 2020).

To face such challenges, we propose an interface based on conversational bots using the reinforcement-learning algorithm to trace and adapt intelligently to changes and volatility.

Under the name of chatbots, conversational bots invade our smartphones, web applications, and enterprise systems in today's modern society (Kulkarni, et al. 2017). Examples of use

cover applications from entertainment and gaming to technical online selling and customer management. However, their use in the context of the supply chain remains very limited. The rest of the paper is structured as follows: Section 2 presents the relevant literature and the research gap that needs to be addressed. Section 3 focuses on the proposed approach and methodology. Section 4 concludes this paper with a summary and a discussion about the limits and open views.

2 BACKGROUND AND RESEARCH GAP

2.1 Suppliers Selection Process in Supply Chains

A supply chain network involves many complex flows and several channels to deliver the orders to sales points. These channels involve complex linkages of suppliers-customer partnerships (Bai, et al. 2019). The most challenging decision-making process associated with these linkages is the supplier selection process (Arioğlu, Sarkis and Dhavale 2020) (Figure 1).

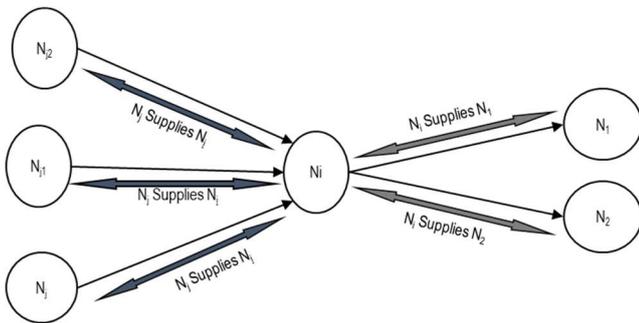


Figure.1. Simplified illustration of the supplier-customer selection

Substantial efforts have been spent in developing decision-making solutions (Wang et al. 2019). They often use conventional approaches, including mathematical programming (Bai, et al. 2019), simulation (Chien, Lin and Lin 2020), probabilistic and stochastic techniques (Kulkarni, et al. 2017), and so forth.

The research works on supplier selection are plenty, and the literature has well investigated the issue. However, most of these works rely on traditional approaches through mathematics and operational research models to find a more beneficial supplier (Chien, Lin and Lin 2020). Only a limited number of studies have proposed adaptive systems from Artificial Intelligence (AI) such as AHP (Hosseini et Al Khaled 2019) and fuzzy systems (Wang et al. 2019) (Wang, et al., 2019)).

2.2 The growing risk of suppliers because of the COVID-19

By nature, partnerships in the supply chain are open to an assortment of risks involving geopolitical, commercial, technological, and environmental disruptions (Sharma, et al. 2020). However, the focus on the suppliers' risk took a vast concern since the financial crisis of 2008 that caused damages (Singh, et al. 2020). The supplier risk is an umbrella term wrapping several root points such as foreign exchange stoppage, price fluctuation, stock out, and inflations (Ivanov et Dolgui 2020).

Today and for several months, we are living in a similar situation because of the COVID-19 outbreak. Recent research published in (Liu, et al. 2020) predicted a disruption in the

global economy to cause a considerable decline, especially in hand-tasks activities like manufacturing (Ghadge, et al. 2020). The breakdown of relationships causing problems with orders and cuts in services and deliveries in industries during confinement affects the suppliers.

Two main factors are involved in the selection process of suppliers: the criteria's values and the weights of the criteria estimated by the client. With the COVID19 crisis, the risk associated with the weights (prioritization) has been altered and increased due to the pandemic. For example, a car manufacturer has changed the focus on the risk linked to the "Geographical location" criterion during the first wave of the COVID-19 crisis (April-May 2020) because of the lockdown and the complete stop of import/export transports, and closure of borders

3 PROPOSED APPROACH

Before giving the proposed model, we introduce the foundation concepts of the solution in the following sub-sections.

3.1 Multi-Criteria Selection Methods for Supplier Selection

The selection of suppliers is factually a multiple-criteria decision-making (MCDM) that needs a Multi-Criteria Decision Analysis (MCDA) (Gupta, Soni et Kumar 2019). It is one of the most complex decision-making processes containing several challenges like the heterogeneity in quantitative and qualitative criteria, the changing values in time, the implicit preferences, the complex assessment, and so forth. The complexity in the multicriteria model increases according to the criteria (Bai, et al. 2019). The number of criteria involved in the evaluation process makes the feasibility of some approaches complex, such as the traditional multicriteria methods, because of the combinatory explosion. Therefore, the research works are more using AI heuristics (Bai, et al. 2019) (Figure 2).

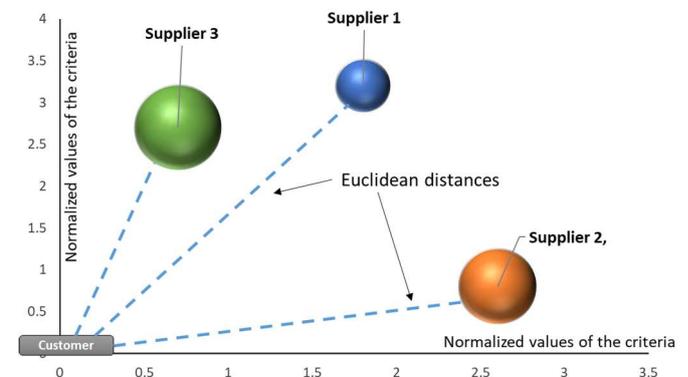


Figure.2. Projection of the multicriteria supplier selection on the 2D plan with Euclidean distances

Several algorithms exist and are already in use for supplier selection, such as Analytic Hierarchy Process (AHP), ELECTRE, MOORA, MACBETH (BANA Consulting 2020), etc. Table 1 proposes a macro-synthesis of the main approaches from (Memari, et al. 2019) following the classification (Zimmer, Fröhling and Schultmann 2016).

Classification	Approach / meta-approach	Relevant references
Mathematical Analytics (MA)	Analytic Hierarchy Process (AHP)	(Hosseini et Al Khaled 2019)
	Analytic Network Process (ANP)	(Hsu et Hu 2009)
	Fuzzy TOPSIS	(Memari, et al. 2019)
	Data Envelopment Analysis (DEA)	(Dobos et Vörösmarty 2019)
	PROMETHEE	(Azadfallah 2017)
	ELECTRE	(Tavassoli, Darestani and Tavassoli 2018)
	Fuzzy Entropy Weighted (FEW)	(Wood, 2016)
	Matrix-based analysis	(Gelderman et Semeijn, 2006)
Qualitative Analysis (QA)	DELPHI	(Gama, Alves et Oliveira 2020)
	Quality Function Deployment (QFD)	(Tavana, Yazdani et Di Caprio 2017)
Artificial Intelligence (AI)	K-NN and Linear Regression (LR) algorithms (Machine learning)	(Cavalcante, et al. 2019)
	Particle Swarm Optimization (PSO) algorithm	(Li, et al. 2018)
	Genetic Algorithm (GA)	(Ben Jouida et Krichen 2020)
	SWOT & AI-based Approach	(Aggarwal 2019)

Table 1. Synthesis of some relevant approaches in MNDM inspired by (Memari, et al. 2019) and the classification of (Zimmer, Fröhling and Schultmann 2016)

In our approach, we chose to use the most straightforward and the most usually applied, the Weighted Sum Method (WSM) (also used in (Rani et Mishra 2020), (Tusnial, et al. 2020), and many other research works).

3.2 Conversational Bots

Conversational bots are digital entities that can express explicit dialog with humans or between each other by behaving autonomously or semi-autonomously (Van Pinxteren, Pluymaekers et Lemmink 2020). Conversational bots imitate human behavior in reasoning and "thinking" by involving the dialog capacity (communication). A bot can be embedded in physical support or hardware (such as Amazon Echo® embedding Alexa® software). The dialog support can be through the vocal canal, textual canal, or both (Figure 3).

Also called 'Chatbots,' conversational bots are technically at the intersection between the agent's system (autonomous software) and the Human-to-Machine (HM) system (Cuayáhuitl et al., 2019, 2). To summarize, in addition to the agent's features, the chatbot has the particularity of conversation with humans (one or multiple humans) and the use of the human-machine communication layer to ensure this conversation feature. In a nutshell, a chatbot must have an understanding ability of natural

human language. Therefore, technically, chatbots must embed a Natural Language Processing (NLP) based technique (Ricciardelli et Biswas 2019). Besides, the chatbot has to master the written and / or the vocal reply to the human. Moreover, chatbots' aim is focused on the interactions with humans, and a system can include one bot without basing the processing on the collective intelligence, which can be the purpose of multi-agent systems (Van Pinxteren, Pluymaekers et Lemmink 2020).

A bot is meant and designed to learn from the human answers, which is not mandatory in multi-agent systems where the learning can be through self-organization and evolution from the environment (Vaidyam, et al. 2019). Another difference is that multi-agent systems can be completely independent of human interactions and without natural speech/text abilities (AbuShawar and Atwell 2015).

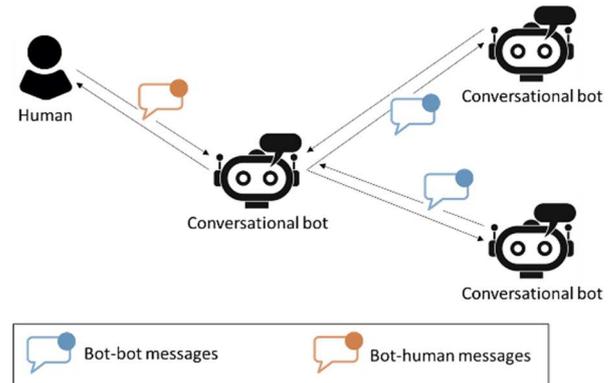


Figure.3. Representative model of conversational bots

Conversational bots are widely used in medical studies because of their benefit in assisting physicians and talk to patients (Vaidyam, et al. 2019). However, as far as we know, no research proposed an approach based on the conversational bots for the supplier selection decision-making process, and the few works we found about the use of chatbots in the supply chain domain are in the majority oriented to the customer service field like the marketing automation and the interactive marketing (Pradana, Goh et Kumar 2018).

The partners involved in the exchanges assimilated to software conversational bots can represent negotiations in selecting suppliers (1 customer-n supplier) (Bai, et al. 2019) where the traditional models may be too heavy to assimilate changes.

3.3 Reinforcement Learning

Reinforcement learning (RL) is a class of algorithms in Machine Learning (ML) that can learn the best action to select by maximizing a reward from the environment (Chien, Lin and Lin 2020). That is why reinforcement learning is primarily used in the complex changing environment where it is a matter of selection action in decision making. There are, indeed, standard analog features between the selection of actions in Reinforcement Learning and the selection of suppliers. Despite this analogy, the research works that employ reinforcement learning for supplier selection are non-existent. We think that maybe a serious research path to explore.

On the other hand, only a few works proposed reinforcement learning for chat-bots: (Serban, et al. 2017). Authors in (Cuayáhuitl, et al., 2019) proposed human-likeness assessment as rewards for the algorithm. Reinforcement learning applications in the supply chain field seem to be oriented to self-improvement, scheduling, and forecasting more than predictive

tasks. Examples, among others, are (Chien, Lin and Lin 2020), (Wang, et al. 2020), and (Hwang et Jang 2020).

Defined as a prevailing algorithm in Reinforcement Learning, Q-learning is probably one of the most used in academic literature and industrial applications (Azar, et al. 2013). The algorithm offers innate learning from the time-series of actions (states).

In this research, the bot interrelates with the environment (including the other bots) continuously and tries to optimize (maximize) the rewards received from the environment (Figure 4).

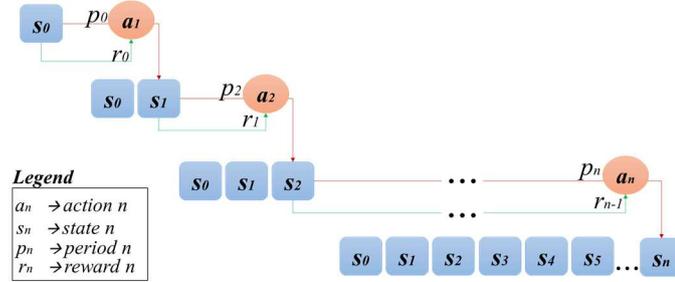


Figure 4. Reinforcement learning mechanism

3.4 Model

3.4.1 Description

In the proposed model, each customer tries to determine the distribution of quantities to order to a selected set of suppliers. Given the varying number of suppliers and the criteria priorities, each supplier's selection and risk assessment operation continuously changes. The criteria set includes parameters with sequential time-oriented values, including prices, inventories, qualities, product quality, contract mode, etc. At the first phase of the process, the customer must rank and place the orders associated with the quantities after evaluating and ranking the suppliers.

The basic scenario is described the following: a customer-bot makes an order to the set of potential suppliers-bots. Then, the suppliers-bots reply with their proposals. Afterward, the customer then collects the proposals and evaluates the criteria according to the priorities using the MCDA-engine that activates the RL-engine. The MCDA-engine uses the Weighted Sum Method (WSM) approach to rank the suppliers and affect the quantities mentioned in the previous section. On the other hand, the RL-engine provides a rewarding assessment and generates the bias coefficients associated with the risk factor (Figure 5).

3.4.2 Formulation

Suppose for a time-period p , the customer-bot (C^p) has to order a set of d_i^p quantities to n suppliers-bots $\{S_i^p\}_{i \leq n}$.

$$C^p \rightarrow \{S_i^p\}_{i \leq n}$$

Suppose the criteria set includes two subsets: the set of the product criteria and the suppliers' offer criteria. The following list of criteria is an example of the permanent criteria we can find in almost all models, but this list is not exhaustive:

Product quality: the quantitative criterion of the quality of the

i^{th} supplier product Q_i^p

$$Q_i^p \in \{1, 2, 3, \dots, Q^p(\max)\}$$

Supplier Reputation: the qualitative criterion related to the position of the i^{th} supplier in the market R_i^p

$$R_i^p = \{1, 2, 3, \dots, R^p(\max)\}$$

Stockout recovering mode: When the stock out occurs and the contract still running, it is usual that suppliers practice recovering through two ways:

- **Product substitution with a different price:** The product is unavailable, and the suppliers propose an alternative one BS_i^p with a different price,
- **Product substitution with the same price:** The product is not available, and the seller offers a comparable or superior product at the price BS_i^p .

Supply contract mode: There are two modes

- **Long-Term contract:** the supplier requires the customer a maximum order following the stock capacity level K , based on the pricing function $f_k(d_i^p)$. We define hence

$\varepsilon_k(d^p)$ is the capacity reservation penalty.

$$\varepsilon_k(d_i^p) = f_k(d_i^p) - f_k(K)$$

The total cost of the goods $C_k^{lrc}(d)$ is then:

$$C_i^p(d_i^p) = \varepsilon_k(d_i^p) \times d_i^p = (f_k(d_i^p) - f_k(K)) \times d_i^p$$

- **Short-term contract:** This model describes the spot markets where the price function is directly derivative from the demand where the unit price $f(d_i^p)$ would be a random variable depending on the ordered quantity d . Hence, the total cost of goods would be:

$$C_i^p = f(d_i^p) \times d_i^p$$

The model for the customer is to minimize the cost, and the price is defined as:

$$\text{Minimize } \sum_{i=1}^n C_i^p(d_i^p) \text{ and } \sum_{i=1}^n f_i(d_i^p)$$

The transportation cost and the other costs are not considered in the approach. We assume that the evaluated costs are only those generated by the supplier involved in the price evaluation.

Service level: We assume the service level L_i^p offered by the supplier i :

$$L_i^p = \{1, 2, 3, \dots, L^p(\max)\}$$

The customer-bot attributes the non-negative weights α_i^p that are associated with the set of preferences for each criterion:

$$\alpha_i^p = [\alpha_1^p \quad \alpha_2^p \quad \alpha_3^p \quad \dots \quad \alpha_n^p]^T$$

Accordingly, the model is defined as follows:

Minimize the objective function:

$$F^p(d_i^p) = \sum_{i=1}^n Q_i^p \cdot \alpha_1^p + \sum_{i=1}^n R_i^p \cdot \alpha_2^p + \sum_{i=1}^n q_i^p \cdot BS_i^p \cdot \alpha_3^p + \sum_{i=1}^n (1 - q_i^p) \cdot BS_i^p \cdot \alpha_3^p + \sum_{i=1}^n C_i^p(d_i^p) \cdot \alpha_4^p + \sum_{i=1}^n f_i^p(d_i^p) \cdot \alpha_5^p + \sum_{i=1}^n L_i^p \cdot \alpha_6^p$$

With:

$$\begin{bmatrix} Q_1 & Q_2 & Q_3 & \dots & Q_n \\ R_1 & R_2 & R_3 & \dots & R_n \\ q_1 \cdot BS_1 & q_2 \cdot BS_2 & q_3 \cdot BS_3 & \dots & q_n \cdot BS_n \\ (1 - q_1) \cdot SS_1 & (1 - q_2) \cdot SS_2 & (1 - q_3) \cdot SS_3 & \dots & (1 - q_n) \cdot SS_n \\ C_1(d_1) & C_2(d_2) & C_3(d_3) & \dots & C_n(d_n) \\ f_1(d_1) & f_2(d_2) & f_3(d_3) & \dots & f_n(d_n) \\ L_1 & L_2 & L_3 & \dots & L_n \end{bmatrix}^p \times [\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6]^p$$

Suppose the risk factor associated with the i^{th} supplier is w_i^p . We associate the risk to the reward function θ_i^p in Q-Learning algorithm. As the reward must be maximized, we reverse the function of risk. In other words, the higher the risk, the less the supplier would be selected:

$$\theta_i^p = 1/w_i^p$$

The Q values of Q-Learning are computed as the following:

$$Q_{new} \leftarrow Q_{old} + \alpha \cdot (1/w_i^p + \gamma \cdot \max_{history} (Q) - Q_{old})$$

Where:

$\alpha \in [-1,1]$ is the learning rate,

$\gamma \in [-1,1]$ is the discount factor of rewards,

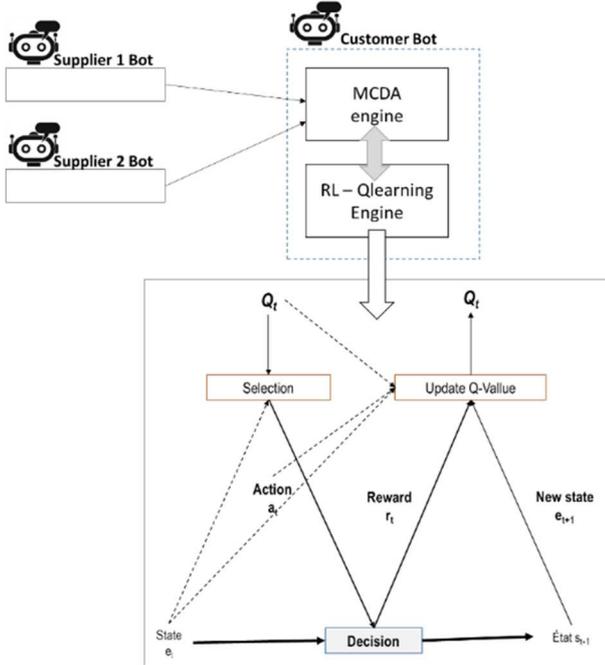


Figure 5. The simplified model of the proposed approach

4 EXPERIMENTS AND INDUSTRIAL VALIDATION

4.1 Case Study of Automobile Manufacturer

Supplier selection is one of the most critical challenges in the automobile manufacturing industry because of the high sensibility to stock-outs and the suppliers' dependence to make the production lines operating non-stop in the plants. The complex flows of sourcing requiring certified expert suppliers are also causes (Yoon, et al. 2018).

Since the beginning of restrictions caused by the COVID-19 crisis, the automobile manufacturers intensified the fears about the suppliers' risks and tried to find an efficient solution for the supplier's selection with the core-focus on the risk primarily.

Table 2 summarizes the two views for ranking according to Weber (Weber Current et Benton 1991) and Dickson (Dickson 1966) from the extensive literature study of (El-Makrini 2020).

Code	Criteria	Ranking of Dickson	Ranking of Weber
[C1]	Quality	1	3
[C2]	Delivery	2	2
[C3]	Past performance	3	9
[C4]	Warranty policy	4	14
[C5]	Production capacity	5	4
[C6]	Price	6	1
[C7]	Technical capacity	7	6
[C8]	Financial situation	8	9
[C9]	Process compliance	9	12
[C10]	Communication system	10	12
[C11]	Reputation in the industry	11	8
[C12]	Desire to do business	12	13
[C13]	Management and organization	13	7
[C14]	Control of operations	14	11
[C15]	Repair service	15	9
[C16]	Attitude	16	10
[C17]	Impression	17	12
[C18]	Wrapping skill	18	11
[C19]	Social relations	19	12
[C20]	Geographical location	20	5
[C22]	Purchasing volumes in the past	21	13
[C23]	Training and support	22	12
[C24]	Reciprocity of the relationship	23	12

Table 2. Classification of the 24 criteria according to Weber (Weber Current et Benton 1991) and Dickson (Dickson 1966)

4.2 Scenario of validation and preliminary results

We started the validation by a preliminary simulation scenario by using the Dataset of 10 suppliers from Kaggle® jointly with the list of the objective criteria provided in Table 2.

To simulate the case in this preliminary scenario, we implemented our new model and algorithm with R language in RStudio® and reused a method of MCDA named MACBETH for the benchmarking. MACBETH is a primarily used method available on (BANA Consulting 2020).

The outstanding result is shown as followed in Figure 6.

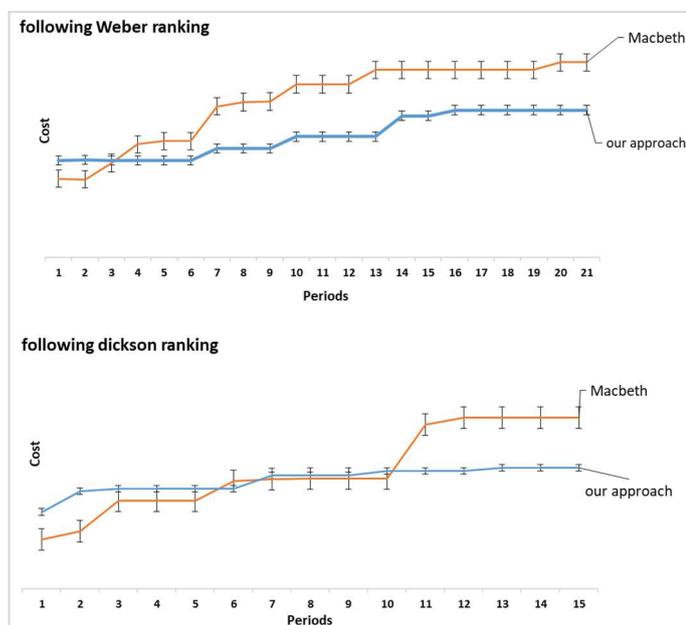


Figure.6. Results of the preliminary scenario

Figure 6 (following Weber ranking) shows that the orange and blue curves represent the cost related to MACBETH and our approach, respectively. The orange curve started with a cost lesser than our approach; however, after the 10th period, the curve continued increasing mainly higher than the other one. Likewise, our model still performs better than MACBETH does from the other comparative curves (following Dickson ranking). Even in the initial stage, the blue curve was slightly higher but has rapidly achieved stability. Hence, our approach provides promising results in this preliminary scenario and seems more robust than the traditional MCDA method.

5 CONCLUSION, LIMITATIONS, AND OPEN VIEWS

This paper's main motive is to propose a practical approach for supplier selection with high consideration for the risk related to the suppliers (because of the COVID-19 effect). The proposed approach is based on interactive exchanges, and dynamic ranking endorsed through conversational bots with the Reinforcement Learning (RL) algorithm.

The pilot results obtained are promising, and the work continues to deploy on the real Dataset. However, despite the evident observed increasing suppliers risk caused by the COVID-19 outbreak, the research on the topic is still premature and limited because we still do not have enough hindsight after only a few months since the beginning of the pandemic.

Despite the limitation of research to have good feedback about the COVID-19 related risks, the model we propose may serve as an innovative pilot and make a clear contribution to improve and automate the supplier selection process.

6 REFERENCES

Aggarwal, A. K., Davè, D. S. "Supplier Selection in Supply Chains: A SWOT Guided Artificial Intelligence Approach." *Journal of Supply Chain Management*, 2019: 16(3), 27-35.

Arıoğlu, M. Ö., J. Sarkis, and D. G. Dhavale. "Selection of suppliers using Bayesian estimators: a case of concrete ring suppliers to Eurasia Tunnel of Turkey."

International Journal of Production Research, 2020: 1-12.

Azadfallah, M. «Multi criteria supplier selection using PROMETHEE outranking procedures. *Journal of Supply Chain Management Systems*.» 2017: 6(1), 24.

Azar, M. G., R. Munos, M. Ghavamzadeh, and H. J. Kappen. "Speedy Q-Learning: A computationally efficient Reinforcement Learning algorithm with a near-optimal rate of convergence." *Journal of Machine Learning Research*, 2013: 1-26.

Bai, C., S. Kusi-Sarpong, H. Badri Ahmadi, et J. Sarkis. «Social sustainable supplier evaluation and selection: a group decision-support approach.» *International Journal of Production Research*, 2019: 57(22), 7046-7067.

BANA Consulting. *MACBETH*. 10 12 2020. <http://m-macbeth.com/>.

Ben Jouida, S., et S. Krichen. «A genetic algorithm for supplier selection problem under collaboration opportunities.» *Journal of Experimental & Theoretical Artificial Intelligence*, 2020: 1-27.

Cavalcante, I. M., E. M. Frazzon, F. A. Forcellini, and D. Ivanov. "A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing." *International Journal of Information Management*, 2019: 49, 86-9.

Chehbi-Gamoura, S. «A Cloud-Based Approach for Cross-Management of Disaster Plans: Managing Risk in Networked Enterprises.» Dans *Critical Research on Scalability and Security Issues in Virtual Cloud Environments*, de S. Aljawarneh et M. Malhotra, 240-268. New York USA: IGI Global, 2019.

Chen, Z., X. Ming, T. Zhou, and Y. Chang. "Sustainable supplier selection for smart supply chain considering internal and external uncertainty: An integrated rough-fuzzy approach." *Applied Soft Computing*, 2020: 87, 106004.

Chien, C. F., Y. S. Lin, and S. K. Lin. "Deep reinforcement learning for selecting demand forecast models to empower Industry 3.5 and an empirical study for a semiconductor component distributor." *International Journal of Production Research*, 2020: 58(9).

Cuayáhuítl, H., D. Lee, S. Ryu, et Y., Choi, S., Indurthi, S., ... & Kim, J. Cho. «Ensemble-based deep reinforcement learning for chatbots. » *Neurocomputing*, 2019: 366, 118-130.

Cuayáhuítl, H., D. Lee, S. Ryu, S. Choi, I. Hwang, and J. Kim. "Deep reinforcement learning for chatbots using clustered actions and human-likeness rewards. ." *In Proceedings of International Joint Conference on Neural Networks*, 2019: (pp. 1-8). IEEE.

Dickson. «An analysis of vendor selection systems and decisions.» *Journal of Purchasing*, 1966: 5-17.

Dobos, I., et G. Vörösmarty. «Inventory-related costs in green supplier selection problems with Data Envelopment Analysis (DEA).» *International Journal of Production Economics*, 2019: 209, 374-380.

Durmic, E. «Evaluation of criteria for sustainable supplier selection using FUCOM method.» *Operational*

- El-Makrini, Jamila. *Évaluation multicritère des performances qualité et logistique des fournisseurs : le cas PSA groupe [Confidential]*. Master 2 Thesis, under supervision of Samia Chehbi GAMOURA and Vincent BARUTHIO, Strasbourg, France: EM Strasbourg Business School, 2020.
- Gama, N., C. A. Alves, et P. S. Oliveira. «Suppliers Selection in Restaurants: Application of Delphi and Fuzzy AHP Methods.» *Journal of Hospitality*, 2020: 2(3-4), 94-106.
- Ghadge, A., S. K. Jena, S. Kamble, D. Misra, and M. K. Tiwari. "Impact of financial risk on supply chains: a manufacturer-supplier relational perspective." *International Journal of Production Research*, 2020: 1-16.
- Gupta, S., U. Soni, et G. Kumar. «Green supplier selection using multi-criterion decision making under fuzzy environment: A case study in automotive industry.» *Computers & Industrial Engineering*, 2019: 136, 663-680.
- Hadian, H., S. Chahardoli, A. M. Golmohammadi, and A. Mostafaeipour. "A practical framework for supplier selection decisions with an application to the automotive sector." *International Journal of Production Research*, 2020: 58(10), 2997-3014.
- Hosseini, S., et A. Al Khaled. «A hybrid ensemble and AHP approach for resilient supplier selection.» *Journal of Intelligent Manufacturing*, 2019: 30(1), 207-228.
- Hsu, C. W., et A. H. Hu. «Applying hazardous substance management to supplier selection using analytic network process.» *Journal of cleaner production*, 2009: 17(2), 255-264.
- Huang, Z., W. M. van der Aalst, X. Lu, and H. Duan. "Reinforcement learning based resource allocation in business process management." *Data & Knowledge Engineering* 70, no. 1 (2011): 127-145.
- Hwang, I., et Y. J. Jang. «Q (λ) learning-based dynamic route guidance algorithm for overhead hoist transport systems in semiconductor fabs.» *International Journal of Production Research*, 2020: 58(4), 1199-1221.
- Ivanov, D., et A. Dolgui. «Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak.» *International Journal of Production Research*, 2020: 58(10), 2904-2915.
- Javad, M. O. M., M. Darvishi, et A. O. M. Javad. «Green supplier selection for the steel industry using BWM and fuzzy TOPSIS: A case study of Khouzestan steel company.» *Sustainable Futures*, 2020: 2, 100012.
- Kaya, Ü., H. İ. Koruca, et S. Chehbi-Gamoura. «Development of a Flexible Software for Disassembly Line Balancing with Heuristic Algorithms.» Dans *Artificial Intelligence and Applied Mathematics in Engineering Problems*, de D. Jude et H. U. Kose, 1014-1026. Cham, Switzerland: Springer, 2019.
- Kucharavy, D., D. Damand, S. Chehbi-Gamoura, M. Barth, and S. Mornay. "Entrepôt du futur: le concept de cartographie de contradictions." *Logistique & Management*, 2020: 28(1), 48-56.
- Kulkarni, C. S., A. U. Bhavsar, S. R. Pingale, et S. S. Kumbhar. «BANK CHAT BOT—An Intelligent Assistant System Using NLP and Machine Learning.» *International Research Journal of Engineering and Technology*, 2017: 4(05).
- Li, Y., K. Ding, L. Wang, W. Zheng, Z. Peng, et S. Guo. «An optimizing model for solving outsourcing supplier selecting problem based on particle swarm algorithm.» *Journal of Industrial and Production Engineering*, 2018: 35(8), 526-534.
- Liu, M., Z. Liu, F. Chu, F. Zheng, et C. Chu. «A new robust dynamic Bayesian network approach for disruption risk assessment under the supply chain ripple effect.» *International Journal of Production Research*, 2020: 1-21.
- Liu, P., H. Gao, et J. Ma. «Novel green supplier selection method by combining quality function deployment with partitioned Bonferroni mean operator in interval type-2 fuzzy environment.» *Information Sciences*, 2019: 490, 292-316.
- Memari, A., A. Dargi, M. R. A. Jokar, R. Ahmad, et A. R. A. Rahim. «Sustainable supplier selection: A multicriteria intuitionistic fuzzy TOPSIS method.» *Journal of Manufacturing Systems*, 2019: 50, 9-24.
- Pradana, A. D. I. T. Y. A., O. S. Goh, et Y. J. Kumar. «Intelligent conversational bot for interactive marketing. » *Journal of Telecommunication, Electronic and Computer Engineering*, 2018: 10(1-7), 1-4.
- Rani, P., et A. R. Mishra. «Multi-criteria weighted aggregated sum product assessment framework for fuel technology selection using q-rung orthopair fuzzy sets.» *Sustainable Production and Consumption*, 2020: 24, 90-104.
- Ricciardelli, E., et D. Biswas. «Self-improving Chatbots based on Reinforcement Learning.» *In Proceeding of the 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making*, 2019.
- Samiri, M. Y., M. Najib, A. El Fazziki, and J. Boukachour. "Toward a Self-Adaptive Workflow Management System Through Learning and Prediction Models." *Arabian Journal for Science and Engineering* 42, no. 2 (2017): 897-912.
- Scrimieri, D., N. Antzoulatos, E. Castro, and S. M. Ratchev. "Automated experience-based learning for plug and produce assembly systems." *International Journal of Production Research* 55, no. 13 (2017): 3674-3685.
- Scuotto, V., F. Caputo, M. Villasalero, et M. Del Giudice. «A multiple buyer–supplier relationship in the context of SMEs’ digital supply chain management.» *Production Planning & Control*, 2017: 28(16), 1378-1388.
- Serban, I. V., et al. «A deep reinforcement learning chatbot.» 2017: arXiv preprint arXiv:1709.02349.

- Sharma, M., S. Luthra, S. Joshi, et A. Kumar. «Developing a framework for enhancing survivability of sustainable supply chains during and post-COVID-19 pandemic.» *International Journal of Logistics Research and Applications*, 2020: 1-21.
- Singh, S., R. Kumar, R. Panchal, et M. K. Tiwari. «Impact of COVID-19 on logistics systems and disruptions in food supply chain.» *International Journal of Production Research*, 2020: 1-16.
- Tavana, M., M. Yazdani, et D. Di Caprio. «An application of an integrated ANP–QFD framework for sustainable supplier selection.» *International Journal of Logistics Research and Applications*, 2017: 20(3), 254-275.
- Tavassoli, M. A., S. A. Darestani, and S. A. Tavassoli. "Supplier selection and evaluation using QFD and ELECTRE in quality management system environment (case study: Faravari & Sakht Company)." *International Journal of Productivity and Quality Management*, 2018: 24(1), 84-100.
- Tusnial, A., S. K. Sharma, P. Dhingra, et S. Routroy. «Supplier selection using hybrid multicriteria decision-making methods.» *International Journal of Productivity and Performance Management*, 2020.
- Vaidyam, A. N., H. Wisniewski, J. D. Halamka, M. S. Kashavan, et J. B. Torous. «Chatbots and conversational agents in mental health: a review of the psychiatric landscape. .» *The Canadian Journal of Psychiatry*, 2019: 64(7), 456-464.
- Van Pinxteren, M. M., M. Pluymaekers, et J. G. Lemmink. «Human-like communication in conversational agents: a literature review and research agenda.» *Journal of Service Management*, 2020.
- Vijayakumar, Y., S. A. Rahim, A. Ahmi, and N. A. Rahman. "Investigation of supplier selection criteria that leads to buyer-supplier long term relationship for semiconductor industry. ." *International Journal of Supply Chain Management*, 2019: 8(3), 982-993.
- Wang, C. N., C. Y. Yang, and H. C. Cheng. "A fuzzy multicriteria decision-making (MCDM) model for sustainable supplier evaluation and selection based on triple bottom line approaches in the garment industry." *Processes*, 2019: 7(7), 400.
- Wang, C. N., C. Y. Yang, et H. C. Cheng. «Fuzzy multicriteria decision-making model for supplier evaluation and selection in a wind power plant project. .» *Mathematics*, 2019: 7(5), 417.
- Wang, H., B. R. Sarker, J. Li, et J. Li. «Adaptive scheduling for assembly job shop with uncertain assembly times based on dual Q-learning.» *International Journal of Production Research*, 2020: 1-17.
- Weber Current et Benton. «Vendor Selection criteria and methods .» *European Journal of Operational Research*, 50, 1991: 2-18.
- Yoon, J., S. Talluri, H. Yildiz, et W. Ho. «Models for supplier selection and risk mitigation: a holistic approach.» *International Journal of Production Research*, 2018: 56(10), 3636-3661.
- Zimmer, K., M. Fröhling, and F. Schultmann. "Sustainable supplier management—a review of models supporting sustainable supplier selection, monitoring and development." *International Journal of Production Research*, 2016: 54(5): 1412-1442.

Session 1.5 : Design et Pilotage des systèmes de production de biens et de services

Chairman : Eric Ballot - Gülgün Alpan

Coordination mechanisms in supply chains dealing with supply uncertainty: A state of the art

EMILIE LACHANCE, VANESSA SIMARD, NADIA LEHOUX, CAROLINE CLOUTIER

DEPARTMENT OF MECHANICAL ENGINEERING, UNIVERSITÉ LAVAL
2325 Rue de l'Université, Québec, QC, Canada

emilie.lachance.4@ulaval.ca, vanessa.simard.4@ulaval.ca, nadia.lehoux@gmc.ulaval.ca, caroline.cloutier@gmc.ulaval.ca

Résumé

La coordination de la chaîne d'approvisionnement est un défi pour beaucoup d'entreprises dues aux multiples sources d'incertitudes auxquelles elles sont confrontées. Peu importe s'il s'agit de satisfaire la demande ou de satisfaire le calendrier de production, le recours à des mécanismes de coordination peut faciliter la prise en compte de cette incertitude et la coordination de la chaîne d'approvisionnement. Cet article vise donc à présenter, à partir d'une revue de la littérature, un ensemble de mécanismes permettant de mieux faire face à l'incertitude lorsque celle-ci est associée à la matière première dans une chaîne d'approvisionnement. 25 articles ont répondu aux critères de sélection définis et permis de dégager de tels mécanismes de coordination classifiés en neuf catégories : contrats et pratiques économiques, marché aux enchères, coordination verticale, coordination horizontale, confiance, collaboration, partage d'information, pratiques d'évaluation et pratiques de gouvernance. L'étude et l'analyse des mécanismes ont également permis de synthétiser leur application dans le cadre d'un processus de prise de décision.

Abstract

Supply chain coordination is a challenge in many firms due to the multiple sources of uncertainty they face. Whether it is to satisfy the demand or to satisfy the production schedule, the use of coordination mechanisms can help to address this uncertainty and facilitate supply chain coordination. This paper therefore aims to present, based on a literature review, a set of mechanisms to better deal with uncertainty when it is associated with raw materials in a supply chain. 25 articles met the selection criteria and identified coordination mechanisms that were classified into nine categories: contractual and economic practices, spot markets, vertical coordination, horizontal coordination, trust, collaboration, information sharing, assessment practices and governance practices. The study and analysis of the mechanisms also made it possible to synthesize their application in a decision-making process.

Keywords – Supply uncertainty, coordination mechanisms, literature review, supply chains

1 INTRODUCTION

A supply chain represents the flow of information and material exchanged by all processes needed to answer the customer's demand [Lehoux et al., 2012]. The role of decision makers is to use the tools and methods at their disposal to forecast future events so the supply chain can function properly. However, several sources of uncertainty have an effect on these forecasts and thus the efficiency of the chain [Arshinder et al., 2011, Adhikari et al., 2020]. In a demand-driven supply chain, sales forecasts are subject to information that is difficult to predict with certainty as they depend on customer demand or market price. Production processes can also be a source of uncertainty, as it is not always possible to control production time and quality of the final product. As for supply activities, although all the information may be known at the time of ordering, unforeseeable events can affect the actual delivery or the availability of materials. The uncertainty may typically be present in every decision made in a supply chain.

Compared to the manufacturing sector, industries like agriculture, forestry, fishing, and textile are particularly affected by supply uncertainty [Williams, 2011]. Since the ensuing transformation process usually highly depends on resource supplies, any changes in the predicted resource quantity or quality influence the productivity of the entire chain. This type of raw material is also subject to multiple uncontrollable sources

of change such as weather, variable yield, natural disasters, perishability or even disease [Rönnqvist et al. 2015] and can be especially difficult to foresee. Although more apparent during the exploitation and first transformation phases, supply uncertainty affects all of the supply chain. Because of the bullwhip effect, if the exploitation cannot deliver what was ordered, it can cause shortages in all following processes [Lehoux et al. 2012].

Supply uncertainty may be difficult to predict, but a coordinated supply chain can be prepared. One of the easiest ways to prepare for this type of uncertainty is maintaining high inventory. Then, if supply becomes limited, there is no shortage [Dolgui & Prodhon, 2007]. Keeping high inventory can solve many problems, but it also comes at a high cost. Numerous coordination mechanisms exist to reduce the negative effects of supply uncertainty in a supply chain more efficiently [Moyaux et al., 2007, Rönnqvist et al., 2015].

The aim of this article is to focus on coordination mechanisms that reduce the negative effects caused by an uncertain supply. A literature review was therefore conducted to look at the industries subject to multiple uncertain factors related to their supply activities. Based on the articles found, coordination mechanisms were then extracted and classified into nine categories as well as four types of governance. Further analysis of these coordination mechanisms showed that some appear to be more useful during certain periods. Hence, a diagram illustrating which mechanisms should be considered during four

phases of the decision-making process for a typical supply chain with uncertain supply was developed.

Most articles found in the literature focus on reducing the effect of demand uncertainty [Adhikari *et al.*, 2020, Dolgui & Prodhon, 2007]. To the best of our knowledge, coordination mechanisms to support supply chains dealing with supply uncertainty has not been extensively covered in the literature. This paper presents the current state of the art on such coordination mechanisms. The following section summarizes the methodology of the review. Then, section 3 describes the types of supply uncertainty and section 4 presents the coordination mechanisms identified in the 25 relevant articles. A research synthesis is then presented in section 5 followed by a short conclusion in section 6. Section 7 presents an appendix.

2 METHODOLOGY

In order to identify what types of coordination mechanisms could facilitate supply uncertainty management in a supply chain, a literature review was conducted. This section presents the search strategy, the assessment and selection criteria and the analytical results of the literature review.

2.1 Search Strategy

In order to find the right combination of keywords, a few attempts were made on the Engineering Village database. Three main keywords were established in the search field subject/title/abstract. The search strategy followed three phases. The keyword “supply chain” was in all trials of the three phases. The two other main keywords were “uncertainties” and “coordination”, with different variations in the three phases. The search strategy of the third phase was extended to two other databases: Web of Science and ProQuest.

In the first phase of trials, along with the three main keywords, the following keywords were added in the search field “all fields”, one after the other as different trials: “forestry”, “agriculture”, “forest industry”, “fish industry”, “textile industry”, “supply-driven industry”, “petroleum industry” and “biofuels”. These keywords helped narrow down the search to industries that could deal with supply uncertainty.

In a second phase, the first phase was repeated by changing the keyword “coordination” for “coordination mechanism” and adding one other keyword to the search field “all fields” which was “renewable energies”. This search only provided one more record than in the first phase. Two other trials were done in this phase. The first try was done by changing the main keyword “uncertainties” for “raw materials” and combining it with the keyword “industry” in the search field “all fields”. Finally, a last trial was done by changing the keyword “raw materials” for “uncertain environment” and the keyword “industry” for “supply uncertainty”.

In these two phases, only 10 articles out of 83 were found relevant in regards to the subject of research. Most of the articles considered only demand uncertainty or did not mention any coordination mechanisms.

In the third phase, the search strategy was modified in order to find more relevant articles. The following search criteria were specified in the subject/title/abstract field:

- 1) supply chain AND
- 2) uncertainty, uncertain, variability, stochastic, variation or unknown AND
- 3) coordination, coordinate, cooperation, synchronize, collaboration or partnership AND
- 4) industry AND

5) stochastic supply, variable supply, uncertain supply, supply uncertainty, supply variations, supply variability, uncertainty in supply or variation in supply. However, it was specified that these following keywords should not appear in the subject/title/abstract: “demand uncertainty”, “uncertain demand”, “demand variations”, “variation in demand”, “stochastic demand”, “variable demand” or “demand variability”. Without this filter, many articles would only propose coordination mechanisms to reduce demand uncertainties. Although demand uncertainty can cause uncoordinated supply chains, the purpose of this review was to concentrate on the uncertainties in supply.

The third phase was repeated for the search engines Web of Science and ProQuest. In addition, as most of the relevant articles in the first phase came from the combination of the three main initial keywords and the keyword « agriculture », this combination was also searched in these two search engines. With this search strategy, 175 articles were found.

2.2 Assessment and Selection Criteria

Of the 175 articles, 25 were found relevant for this research. It is important to note that some of the articles were neither available or written in English. To ensure the validity of these articles, the abstracts were first analyzed to assess if the articles mentioned coordination mechanisms dealing with supply uncertainty. Thereafter, the introduction and conclusion were evaluated in order to ensure the relevance of each article. For the assessment, at least one of the following keywords from each category had to be mentioned in the article.

- (1) Supply uncertainty, uncertain supply, stochastic supply, variable supply, yield uncertainty, variable yield, stochastic yield;
- (2) Coordination, mechanism, collaboration, governance, partnership, cooperation, structure, integration;
- (3) Agriculture, renewable energy, industry, agri-food, chemical, oil/bioethanol industry, forest, textiles, coffee, wine, beer, meat (or its derivatives), fish, wheat.

2.3 Analytical results

The literature review was conducted between May 11th and June 6th 2020. During the research, it was found that only a few articles focussed on coordination mechanisms for coordinating uncertain supply and random yield. Of the 175 articles found, most covered demand uncertainty or coordination mechanisms between the manufacturer and the customer. Actually, Adhikari *et al.* [2020] who studied a five-level textile supply chain, mentioned in their literature review that very few authors have ventured to consider the supplier’s uncertainty. Thus, it was necessary to use very specific keywords during the search to identify them. Table A in the Appendix presents an overall view on the coordination mechanisms’ categories, their description, the actors involved and the percentage of time it was found in the 25 articles. The next graph presents the number of articles per type of industry. It is interesting to note that most articles are related to the agriculture industry (fruits and vegetables) and to the bioethanol industry.

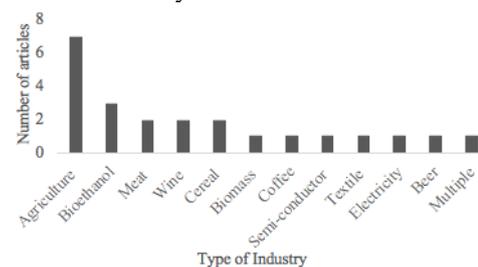


Figure 1. Number of articles per industry

The next figure shows the number of relevant articles published per year. Most research papers were published in recent years, as of 2007. A slight increase in this research area can be observed from 2015 to 2020, which represents 68% of the reviewed articles. Note that since the review includes articles up to June 2020, more articles could be published in 2020.

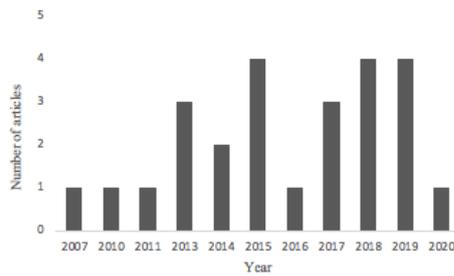


Figure 2. Number of articles based on years of publication
The articles reviewed emanate from 14 countries. The following figure demonstrates that 36% of the articles originate from China. The rest of the countries have one or two articles each.

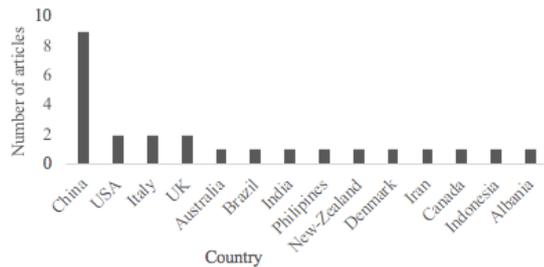


Figure 3. Reviewed articles based on the first author's country affiliation.

The next section presents the factors causing supply uncertainty. They were identified in the reviewed articles by following a content analysis methodology as proposed by Erlingsson and Brysiewicz [2017]. The content analysis mainly consisted of repeated lectures, writing down of keywords, grouping of the ideas which were similar to then extract and interpret the content of each article. The same methodology was followed to classify the coordination mechanisms as presented in section 4.

3 UNCERTAINTIES FOR SUPPLY-DRIVEN CHAINS MENTIONED IN THE LITERATURE

To begin with, the factors related to supply uncertainty found in the articles were classified into categories. Figure 4 shows the percentage of articles mentioning at least one category of uncertainty factor.

The most common category was **weather uncertainty (F1)**, which considers extreme weather, unpredictable and variable weather, humidity exposure, adverse climate and climatic conditions. This uncertainty factor was mentioned in 68% of the reviewed articles. It seems to be the main issue related to other uncertainty factors that will be discussed here. The second category was **uncertain yield (F2)**, which includes volatile crop yield, variable or low yield, unpredictable harvest, uncertain performance of the crops, and was found in 48% of the articles. The third category was **natural and environmental risks (F3)**, which include pest damage, disease, environmental accidents, fire, earthquakes, drought, hurricanes, floods, technological accidents, safety, and was found in 48% of the articles. The fourth was **market uncertainty (F4)**, with price volatility, uncertain prediction of prices in the long run and uncertain future prices which was also found in 44% of the articles.

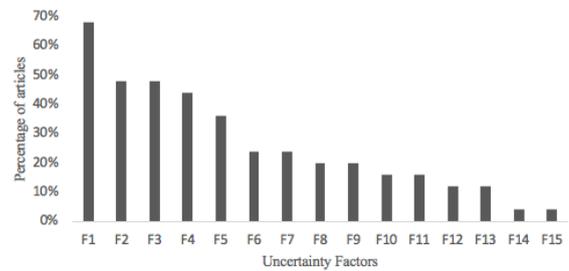


Figure 4. Uncertainty factors present in the articles

Technological uncertainty (F5) corresponding to low use of mechanical planting, low levels of access to technology, policy changes, production technology, technology change and machine breakdown was found 36% of the time. The **variability in quality (F6)** and **in lead-time (F7)**, i.e. maturation of the crops, were both found in 24% of the articles. The next categories were **low capital (F8)** and **perishability (F9)** of the supply, both present in 20% of the articles. The categories of **lack of labor (F10)** and **geographical dispersion and scattered cultivation (F11)** were only mentioned 16% of the time. A total of three articles mentioned supply uncertainty without specifying any type of factors (**F12**). Three articles mentioned **variation in demand (F13)** although the focus of the articles was supply uncertainties. As explained in the search strategy, none of the articles only considered an uncertain demand. **Contamination of materials (F14)** as well as **limited capacity (F15)** were two factors only acknowledged once each. These unpredictable uncertainties represent well why any supply-driven chain with uncertain raw material or yield can be disorganized or dysfunctional. The next section proposes a synthesis of the coordination mechanisms found in the literature review to coordinate these supply chains with stochastic yield.

4 SYNTHESIS OF THE COORDINATION MECHANISMS

The aim of this research is to find coordination mechanisms for supply chains where supply uncertainty is one of the biggest concerns. Upon analysis, contractual arrangements were a dominant form of coordination mechanisms which appeared as different types of contract in almost all articles. Collaboration and trust were also mentioned as coordination mechanisms in many relevant articles. The coordination mechanisms were divided into nine categories, which were also grouped into types of governance:

Table 1. Mechanism categories and types of governance

Type of Governance	Coordination Mechanism Categories
Contractual Governance	1) Contracts & Economic Practices; 2) Spot Markets
Governance Structure	3) Horizontal Coordination; 4) Vertical Coordination
Relational Governance	5) Trust; 6) Collaboration; 7) Technology & Information Sharing
Preventive Governance	8) Assessment Practices; 9) Governance Practices

Five of the nine mechanism categories proposed were previously studied by Wilson et al. [2015] and Mirkovski et al. [2016] who both studied the wine industry's supply chain. Inspired by Han et al. [2011] and Mirkovski et al. [2016], the nine coordination mechanism categories were also divided into

four types of governance. Each of the categories and their respective mechanisms are presented below.

4.1 Contractual Governance

This category includes all mechanisms using transactions and other economic practices. The two mechanisms are presented below.

Contracts and Economic Practices (CEP), of various types, are mentioned in 92% of the articles and seem to be the main coordinating mechanism used to govern a supply chain with uncertainty. Contracts are defined as economic transactions between two parties and are based on written agreements [Mirkovski, 2016] which may involve costs, risks and benefits sharing between supply chain partners. According to Han et al. [2011], contracts should indicate price, quantity, quality, delivery time, rights and obligations of each member of the supply chain. Contracts are mainly used between a supplier and a buyer but can involve more parties if necessary. Depending on the situation, different types of contracts are proposed. The double discount contract is a contract offering discounts by the supplier and reduced prices paid by the buyer. Anderson and Monjardino [2019] used this contract in their research and concluded that for a supply chain in the agriculture industry, the discount should be increased by the level of risk encountered by the grower. The revenue-sharing contract was mentioned in a few articles [Gong et al., 2010; Chaudhuri et al., 2014; Ye et al., 2017; Zhu et al., 2019]. In this type of contract, the retailer or buyer should return a portion of its revenue to the supplier. It usually implies that the parties involved choose a wholesale price and a revenue-sharing percentage. Call option contracts are when the retailer reserves a certain amount of products from the supplier by paying a reservation fee pre-settled or pre-negotiated by the supplier [Wang and Chen, 2017]. Put option contracts are described by Luo et al. [2018], as the ability for a manufacturer to buy the missing yield production on spot markets if needed and the supplier to sell the excess yield produced. Risk-sharing contracts are also mentioned a few times [Ye et al., 2018; Zare et al., 2019; Adhikari et al., 2020] and are contracts that encourage the farmer to produce more, as he will be compensated by the buyer if there is an over production, but penalized for the units of supply not produced in case of under production. Other contracts were stated briefly such as quantity discounts, wholesale prices, hybrids, production cost-sharing, advance, basis, delayed price, and minimum price contracts. Incentive alignments are classified as contracts since they are a way for two or more parties involved to share the costs, benefits and risks [Chaudhuri et al., 2014; Mamillo, Wilson et al., 2015; Ye et al., 2017; Biggeri et al., Ye et al., 2018; Oliveira et al., 2019]. De Rada [2015] mentioned transaction economics such as auctions and developed a linear programming model to coordinate the supply chain in the agriculture industry with yield uncertainty. The mechanism iteratively adjusts the prices which are publicly announced by a consolidation facility until an optimal solution is found. Contractual mechanisms need to be implemented before the growing and harvesting season and tend to reduce risks and uncertainties.

At first glance, **Spot Market (SP)** was considered to be a contract, however as this mechanism usually occurs at a different moment in the decision-process timeline, it was classified as its own mechanism. Spot markets act as residual markets when contracted amounts cannot match the desired demand or production. They occur as instant transactions [Han et al., 2011; Weseen et al., 2014]. This mechanism is mentioned in 28% of the articles. Some argue that spot markets are good when the degree of ownership of a firm and transaction specific

investments are both low [Han et al., 2011; Weseen et al., 2014; Wilson et al., 2015]. Mendelson and Tunca [2007] stated that spot markets are an indirect platform for exchanging information about market prices, demand and new information. Other authors such as Luo et al. [2018] proposed a contract mixed with spot trading. A research paper on the bioethanol industry in Canada also proposed mixed contracts and spot markets which allows the simple exchange of a product without long-term commitment [Weseen et al., 2014]. One can conclude that spot markets are a complementary coordination mechanism to contracts. They involve no long-term commitments and act as a way to exchange market information. Spot markets occur during the selling season.

4.2 Governance Structure

This type of category combines two mechanisms that govern the way transactions occur in a supply chain. The structure can either be horizontal or vertical.

Horizontal Coordination (HC) is one possible structure to govern a supply chain. This means that contracts and spot markets can occur in a horizontal organizational structure. While many firms belong to different supply chains, competitors can pool their common resources of production in order to achieve a better coordination of their own supply chain [Retno et al., 2013]. Moreover, horizontal coordination can be described as decentralized decision-making. For instance, Illy Café, a high-quality coffee supply chain studied by Oliveira et al. [2019], have applied horizontal coordination by giving their suppliers decision rights in order to achieve their goals. The article concluded that the mechanism was efficient enough to coordinate the supply chain and could replace the need for vertical coordination. In order to be a supplier for Illy Café, some requirements were put into place such as adopting post-harvest equipment, a manual workforce structure and production of gourmet coffee. These requirements are to ensure that the horizontal coordination mechanism works. 32% of the articles considered decentralized settings for the coordination of supply uncertainty but not all concluded that this kind of setting is adequate [Ye et al., 2018]. Participative leadership is a type of horizontal coordination, with shared decision-making powers, which has shown to be suitable for industries such as dairy, meat, vegetables and fruit [Akhtar and Khan, 2015].

Vertical Coordination (VC) is defined as the degree of ownership of a firm with respect to its integrated activities along the supply chain [Retno et al., 2013]. This mechanism is present in 44% of the articles. Vertical coordination can also be defined as a centralized decision-making setting. Wilson et al. [2015] mentioned that vertical integration is usually correlated with the firm's age and size. Retno et al. [2013] argued that this mechanism reduces costs, disruption of supply and disruption of quality. The higher the transaction costs, the more one would tend to vertically integrate [Han et al., 2011; Weseen et al., 2014; Wilson et al., 2015]. However, Oliveira et al. [2019] showed evidence that vertical integration makes it almost impossible to achieve high quality grains in the coffee supply chain due to the fact that only 5 to 10% of the farmer's production achieves the desired quality due to weather and climate uncertainties. Similar to the coffee industry, grape growers in the wine industry have also opted for contracts and spot markets rather than vertical integration [Wilson et al., 2015]. Both of these industries have small-scale farmers. Weseen et al. [2014] showed that three firms from the bioethanol industry of Canada have all opted for contracts and spot markets in the first transaction between the grain grower and the ethanol plant. The second transaction (which occurs

between the ethanol plant and the livestock farmer) has one company that chose spot markets, the other one chose vertical integration and the third one chose contracts. Vertical integration seems to be implemented when the specific investments are high.

Figure 6 is an interpretation of the correlation between the mechanisms used and the average degree of transaction specific investment (TSI) based on the articles found. The diagram shows that spot markets are used when the degree of TSI of a firm in another company is low. Horizontal coordination is more of a participative leadership, which requires a higher degree of investment than spot markets. Since contracts require more organization, rules and paperwork, it is considered to have a higher average of TSI than horizontal coordination. Vertical integration is used when the assets are high and there is a need for tight control over the supply chain. Hence, the average degree of TSI is the highest among the four types of coordination mechanisms.

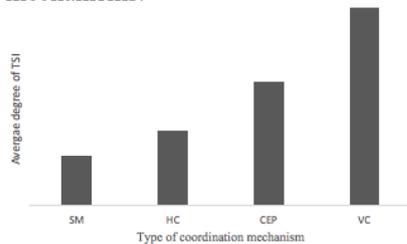


Figure 6. Mechanisms used in relation to the transaction specific investment (TSI)

4.3 Relational Governance

Relational governance is the category where the mechanisms of trust, collaboration and information sharing are found. These mechanisms are reoccurring since they are based on long-term and collaborative relationships [Han et al., 2011]. Relational governance has led to positive impacts in environments with uncertainty due to the fact that it has a certain amount of flexibility compared to contractual governance [Mirkovski et al., 2016]. These three mechanisms all seem to reduce transaction costs [Go Yap, 2017].

Trust (T) is an interesting coordination mechanism as it is neither a contract, a transaction or even quantifiable. Nevertheless, 40% of the articles directly mentioned trust as an efficient mechanism to coordinate a supply chain with uncertainty [Han et al., 2011; Fischer, 2013; Retno et al., 2014; Weseen, 2014; Wilson et al., 2015; Akhtar and Khan, 2015; Mamillo, 2015; Mirkovski et al., 2016; Go Yap, 2017; Oliveira et al., 2019]. Fisher [2013] defined trust as a safeguard mechanism in which trust in itself can reduce the need for contracts. To one extent, it can be described as the willingness of the partners involved in taking risks. Trust is also a tool that can help respect a contract [Mirkovski et al., 2016]. Wilson et al. [2015] studied the supply chain of 585 wineries in South Africa and 580 wineries in New Zealand. With surveys and regression models, the authors concluded that trust can be an effective legal system for contractual arrangements. Although the legal systems are not the same in both countries, both countries had the same results regarding trust as a coordination mechanism. Some examples of trust given by Mirkovski et al. [2016] include a supplier's benevolence, competence, integrity, free information sharing, and high autonomy. As for Fisher [2013], he characterized trust as one's honesty, integrity, sincerity, competence and reliability. He indicated that some research used trust in an empirical model, representing it by action indicators, and found that investing on relationship assets tends to build trust among partners. He also added that a lack of

trust can have high cost implications, and therefore trust reduces transaction costs. Trust can be a good alternative in developing countries, where legally enforced systems are not always accessible to the local growers. In the Philippines, Go Yap [2017] studied the coordination mechanisms in the fresh vegetable supply chain with a qualitative research. Since farmers and buyers do not have access to courts in order to settle the contracts, the relational governance is based on trust to reduce transaction costs. Trust is then an efficient coordination mechanism for those suppliers. However, Fisher [2013] demonstrated through structural equation modelling that trust alone as a mechanism cannot be the solution.

Collaboration (C) is also a type of relational governance. It was presented as a coordination mechanism in 44% of the articles. Collaboration is defined as "joint work between two or more supply chain partners who plan or execute operations together" [Mirkovski et al., 2016]. Mamillo [2015] defined collaboration as joint objectives, intellectual agility, trust, respect and commitment. There can be horizontal collaboration, which implies collaboration between suppliers themselves or competitors, and vertical collaboration between the actors of the supply chain. Sharing views on innovation and production performance exchange are examples of horizontal collaboration [Leat and Revoredo-Giha, 2013]. In their research on supply chains involving farmers in remote areas of the Philippines, Go Yap [2017] presented collaboration as one of the main mechanisms used between supply chain members. Go Yap [2017] argued that there are three elements to collaboration: who you want to collaborate with, to what extent you are going to collaborate (decision synchronization) and on what activities will there be collaboration. Collaboration has been known to enhance trust and vice-versa, which results in reduced transaction costs and time [Fisher, 2013; Mirkovski et al., 2016].

Technology and Information Sharing (TIS) is a type of relational governance strongly related to trust and collaboration [Mirkovski et al., 2016]. Information sharing can be seen as the frequency of communication and the quality of the given information [Fisher, 2013] and is mentioned as a coordination mechanism in 40% of the articles. However, De Rada [2015] claimed that information sharing is necessary but not sufficient enough to coordinate the supply chain alone. The technology is the mean used to transmit the information such as emails, EDI (electronic data interchange), web conferencing, bar codes or radio-frequency identification [Mirkovski et al., 2016]. This same author, in their study of the supply chain of the wine industry in Macedonia and America, explained that information communication technology (ICT) reduces risk costs due to efficient platforms for searching, contact and negotiation. The purpose of the author's research was to determine how contracts and social bonds vary between developing and developed countries. His research concluded that relational governance such as TIS had a positive impact on the supply chain.

The three coordination mechanisms of relational governance are strongly related to one another. Figure 7, inspired from Mirkovski et al. [2016], demonstrates the links between them. It can be observed that one influences the other and vice-versa.

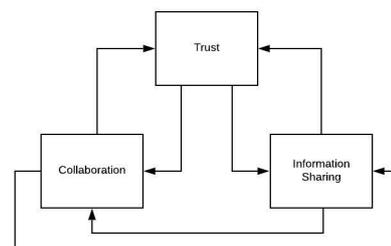


Figure 7. Relational Governance

4.4 Preventive governance

Preventive mechanisms tend to reduce uncertainty in a supply chain, before or during the processes.

Assessment practices (AP), such as monitoring and traceability, are only mentioned in 20% of the articles as complementary mechanisms. In the regions of New Zealand and South Africa, monitoring is one of the two main coordination mechanisms for grape quality [Wilson et al., 2015]. Monitoring, quality control, quality assurance schemes are all assessment practices occurring during the production process. Chaudhuri et al. [2014] discussed coping strategies for the meat industry and proposed monitoring as a coordination mechanism.

Governance practices (GP) are a type of preventive governance that occurs before the production process. In the reviewed articles, it was suggested that governments should establish policies with favorable taxes and subsidies to give incentives to suppliers and buyers to join contract farming [Ye et al., 2017]. In the hybrid power market, government subsidy is proposed to both the retailer and the energy supplier so both can enter a revenue-sharing contract agreement [Zhu et al., 2019]. Subsequently, governance practices act as a complementary mechanism and are proposed in 20% of the articles. In the biofuel industry in China, government incentive is also proposed as a complementary mechanism to risk-sharing contracts [Ye et al., 2018].

5 RESEARCH SYNTHESIS

Based on the literature review, contractual governance, governance structure, relational governance and preventive governance should complement one another. Contracts can initiate future collaboration and trust building. For example, such mechanisms can be used between two or more companies starting to work together. The governance structure mechanisms determine the way the other mechanisms are going to be

implemented. For instance, if a firm is vertically integrated, it may not need contracts or collaboration with other suppliers. On the contrary, if a firm is horizontally integrated, it may want to have a good collaboration or enter in a contract with the other actors of the supply chain. Relational mechanisms tend to reduce the need for other types of mechanisms. These mechanisms are well suited to businesses sharing an existing relationship. Preventive mechanisms are mentioned in few articles, most of the time as a recommendation and an addition to the other proposed mechanisms. These mechanisms may be used by two parties, regardless of the companies' relationship. The categorization by type of governance helped identify the different mechanisms used in supply chains dealing with supply uncertainties. At least 80% of the articles presented more than one mechanism and 64% presented at least four coordination mechanisms.

In order to highlight when specific mechanisms should be used, Figure 8 presents the coordination mechanisms as a function of time in relation to the decision-making phases in a typical supply chain. The x-axis represents the four time periods in the timeline. In each period, the decisions made to ensure the efficiency of the supply chain are different. The construction of the phases was inspired from the agricultural sector, where state of raw material changes according to the seasons. In the first period, which is the planning phase, there is the planting season, pre-harvest season or pre-production season. The second period is the production phase, and can be represented by the growing season, the manufacturing process, the transformation of raw materials or the lead-time to harvest. The third period is the selling phase, which can be the selling season, the harvesting time and the final product/produce. The fourth period is the post-selling phase, which is the post-harvest or post-production season. It is a functional way to show when the different mechanisms seem to be most useful.

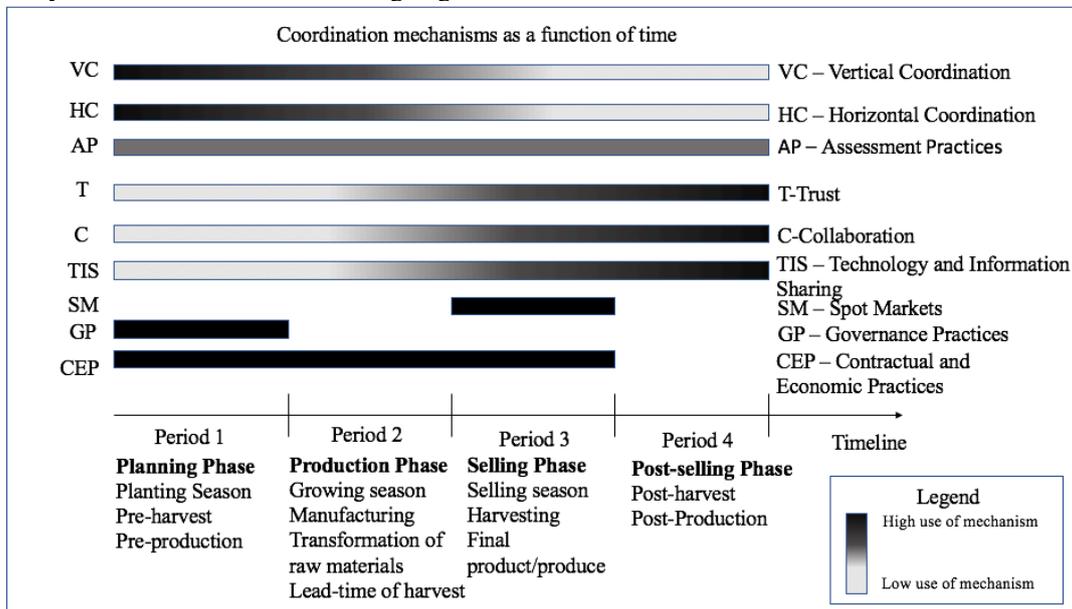


Figure 8. Coordination mechanisms as a function of time

As they are integrated in the supply chain early on, vertical and horizontal coordination tend to be implemented at the beginning of the first period. As time goes by, both mechanisms are still present but the effort needed for implementation is lower. This is represented by a fading bar. Assessment practices are a preventive mechanism that need to be implemented for the four periods since monitoring and traceability are very important aspects of produce quality. Trust, collaboration and technology

and information sharing are all present throughout the four time periods. However, as time goes by, the presence of trust, collaboration and information sharing increases, represented as a darkening bar in the figure. Spot markets occur as instant transactions during the selling phase [Han, 2011]. It is represented as a constant black bar during the third time period. Governance practices are also presented as a constant black bar in the first period since policies, tax reductions and subsidies are

usually determined before the planning phase. Contractual and economic practices is a way of coordinating the supply chain during the first three periods. The wholesale price and type of contract are established during the planning phase [Ye et al., 2017]. Once the product has been yielded (period 2), a certain type of contract allows the supplier and the buyer to share the risks and production costs. During the selling phase (period 3), the buyer might share the revenues with the supplier. Once the selling phase is over, the contract ends and will begin again before the next planning phase.

Overall, it was understood from the articles analyzed that one mechanism is beneficial for another and vice-versa. For instance, government subsidy can influence two actors to enter contract farming. A contract can specify what type of collaboration, information sharing, and assessment practices will be established between the two actors of the supply chain. Collaboration and information sharing can lead to trust and therefore reduce transaction costs. Depending on the needs, a company can centralize or decentralize its decision-making. Finally, spot markets are there in case of under and over production due to uncertainties.

As stated in the beginning of the article, most research found today in the literature focus on the effect of demand uncertainty. Demand uncertainty may not affect the supply chain in the same way that other types of uncertainties do. Nothing guarantees that the mechanisms found in the literature are convenient for all types of uncertainties. The link between the type of uncertainty, the type of mechanism used and their impact on a supply chain has therefore not been seen in the reviewed articles. However, the review showed that supply uncertainty is just as common in the decisions of a typical supply chain and there are multiple solutions available.

6 CONCLUSION

The objective of this paper was to conduct a literature review focussing on the different coordination mechanisms that can reduce the negative effects of uncertain supply in the case of natural resource exploitation sectors. The search strategy for the review provided 175 papers and the selection criteria identified 25 relevant articles. From these articles, the factors influencing supply uncertainties were first explored. Then, the coordination mechanisms presented in the studies were classified into nine categories. The nine categories were also divided into four distinct types of governance. Finally, the coordination mechanisms were distributed in relation to the decision-making phases for a supply chain dealing with supply uncertainty. Most of the reviewed articles recommended the use of more than one coordination mechanism to face supply uncertainty. The review showed that the use of coordination mechanisms to address supply uncertainty is quite current. This is especially true in agricultural sectors. This overview of the literature could be used in further research to test coordination mechanisms in different primary product sectors. Moreover, it would be interesting to research the types of uncertainties in details in the context of supply chain coordination.

Références

Adhikari, A., Bisi, A., & Avittathur, B. (2020). Coordination mechanism, risk sharing, and risk aversion in a five-level textile supply chain under demand and supply uncertainty. *European Journal of Operational Research*, 282 (1), 93–107.

Akhtar P. A., & Khan, Z. (2015). The linkages between leadership approaches and coordination effectiveness. *British Food Journal*, 117 (1), 443–460.

Anderson, E., & Monjardino, M. (2019). Contract design in agriculture

supply chains with random yield. *European Journal of Operational Research*, 277 (3), 1072–1082.

Arshinder, K., Kanda, A. and Deshmukh, S.G., (2011). A review on supply chain coordination: coordination mechanisms, managing uncertainty and research directions. In *Supply chain coordination under uncertainty* (pp. 39-82). Springer Berlin Heidelberg.

Biggeri, M., Burchi, F., Ciani, F., & Herrmann, R. (2018). Linking small-scale farmers to the durum wheat value chain in Ethiopia: Assessing the effects on production and wellbeing. *Food Policy*, 79, 77–91.

Chaudhuri, A., Dukovska-Popovska, I., Damgaard, C.M. and Hvolby, H.H., (2014), September. Supply uncertainty in food processing supply chain: sources and coping strategies. In *IFIP Int. Conf. on Advances in Production Management Systems* (pp. 183-191). Springer, Berlin, Heidelberg.

De Rada, A. N. M. (2015). *Development of Horizontal Coordination Mechanisms for Planning Agricultural Production*. Arizona State University.

Dolgui, A. & Prodron, C. (2007). Supply planning under uncertainties in MRP environments: A state of the art. *Annual Reviews in Control*, 31 (2), 269–279.

Erlingsson, C., & Brysiewicz, P. (2017), “A hands-on guide to doing content analysis”, *African Journal of Emergency Medicine*, Vol. 7 No. 3, pp.93-99.

Fischer, C. (2013). Trust and communication in European agri-food chains. *Supply Chain Management*, 18 (2), 208–218.

Go Yap, E. (2017). A Model of Trust and Collaboration in a Fresh Vegetable Supply Chain in Central Philippines. *Int. J. Applied Industrial Engineering*, 4 (2), 47–57.

Gong, H., Zhang, Y., & LI, J. (2010). Coordination mechanism by option contract in the biomass supply chain organized by “Company and Farmer.” 2010 IEEE Int. Conf. on Automation and Logistics, Hong Kong.

Han, J., H. Trienekens, J., & Omta, S. W. F. Onno. (2011). Relationship and quality management in the Chinese pork supply chain. *Int. J. Prod. Economics*, 134 (2), 312–321.

Leat, P., & Revoredo-Giha, C. (2013). Risk and resilience in agri-food supply chains: The case of the ASDA PorkLink supply chain in Scotland. *Supply Chain Management*, 18 (2), 219–231.

Lehoux, N., Marier, P., D’Amours, S., Ouellet, D., & Beaulieu, J. (2012). *Le réseau de création de valeur de la fibre de bois canadienne*. Centre interuniversitaire de recherche sur les réseaux d’entreprise, la logistique et le transport, Québec (Québec). Rapport scientifique CIRRELT-2012-33. 248 p.

Luo, J., Zhang, X., & Wang, C. (2018). Using put option contracts in supply chains to manage demand and supply uncertainty. *Ind. Mgmt & Data Syst.*, 118 (7), 1477–1497.

Mamillio, D. (2015). Supply Chain Collaboration under Uncertainty in the Albanian Beer Market. *Management Dynamics in the Knowledge Economy J.*, 3 (1), 99–177.

Mendelson, H., & I. Tunca, T. (2007). Strategic spot trading in supply chains. *Management Science*, 53 (5).

Mirkovski, K., Benjamin Lowry, P., & Feng, B. (2016). Factors that influence interorganizational use of information and communications technology in relationship-based supply chains: Evidence from the Macedonian and American wine industries. *Supply Chain Management*, 21 (3), 334–351.

Moyaux, T., Chaib-Draa, B., & D’Amours, S. (2007). Information sharing as a coordination mechanism for reducing the bullwhip effect in a supply chain. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 37 (3), 396–409.

Oliveira, G. M. de, Zylbersztajn, D., & Saes, M. S. M. (2019). Can contracts substitute hierarchy? Evidence from high-quality coffee supply in Brazil. *British Food J.*, 121 (3), 787–802.

Retno, A., Marimin, M., Machfud, M., & Arkeman, Y. (2014). Risks and Risks Mitigations in the Supply Chain of Mangosteen: A Case Study. *Operations and Supply Chain Management: An International Journal*, 6 (1), 11–25.

Rönnqvist, M., D’Amours, S., Weintraub, A., Jofre, A., Gunn, E., Haight, R. G., ... Romero, C. (2015). Operations Research challenges in forestry: 33 open problems. *Annals of Operations Research*, 232 (1), 11–40.

- Wang, C., & Chen, X. (2017). Option pricing and coordination in the fresh produce supply chain with portfolio contracts. *Annals of Operations Research*, 248, 471–491.
- Weseen, S., Hobbs, J., & Kerr, W. (2014). Reducing Hold-up Risks in Ethanol Supply Chains: A Transaction Cost Perspective. *International Food and Agribusiness Management Review*, 17 (2), 83–106.
- Williams, B. K. (2011). Adaptive management of natural resources—framework and issues. *Journal of Environmental Management*, 92 (5), 1346–1353.
- Wilson, M. M. J., MacDonald, I. A., & M. Monnane, M. (2015). Governance of procurement relationships in the wine industry: The importance of trust and the legal system. *International Journal of Wine Business*, 27 (4), 299–311.
- Ye, F., Hou, G., Li, Y., & Fu, S. (2018). Managing bioethanol supply chain resiliency: A risk-sharing model to mitigate yield uncertainty risk. *Industrial Management & Data Systems*, 118 (7), 1510–1527.
- Ye, F., Li, Y., Lin, Q., & Zhan, Y. (2017). Modeling of China's cassava-based bioethanol supply chain operation and coordination. *Energy*, 120, 217–228.
- Ye, F., Li, Y., & Yang, Q. (2018). Designing coordination contract for biofuel supply chain in China. *Resources, Conservation and Recycling*, 128, 306–314.
- Zare, M., Esmacili, M., & He, Y. (2019). Implications of risk-sharing strategies on supply chains with multiple retailers and under random yield. *Int. J. Prod. Econ.*, 216, 413–424.
- Zhu, Z., Kong, L., Xie, J., Li, J., & Cao, B. (2019). Contract coordination optimization of a multi-power supplier-single dominant grid supply chain in hybrid electricity market. *Ind. Management & Data Systems*, 119 (9), 1861–1887.

7 APPENDIX

Table A. Summary of the Nine Coordination Mechanisms

Coordination mechanism category	Description	Examples	Actors involved	Type of governance	% out of 25 articles
Contracts and Economic Practices (CEP)	Economic transactions based on written agreements which involve the sharing of costs, benefits and risks between supply chain partners. Contracts usually indicate price, quantity, quality, delivery time, rights and obligations	Contracts (double discount, quantity discount, revenue-sharing, option, wholesale price, hybrid contracts, risk-sharing, production cost sharing, advance, basis, contract farming), incentives alignments, auctions	Producer enters an agreement with the buyer	Contractual governance	92%
Spot Markets (SM)	Residual market when contracted amounts are not suitable or not available. Supplier and buyer can bargain the price as well as exchange information	Instant transactions, simple exchange at current market price	Farmer enters an instant transaction with buyer (or manufacturer)	Contractual governance	28%
Horizontal Coordination (HC)	Competing suppliers pool their resources together to share a common resource of production or information	Decentralized decision-making, participative decision making	Actors at the same level of the supply chain coordinate their activities	Governance structure	32%
Vertical Coordination (VC)	Degree of ownership of a firm with respect to the vertically integrated activities	Centralized decision-making, joint ownership, joint decision-making, decision synchronization, cooperatives, backward integration	Actors at different levels of the supply chain coordinate their activities	Governance structure	44%
Trust (T)	Safeguard mechanism. Willingness to take risks. Effective substitute for an efficient legal system in governing contractual relations	Communication, dependency, social bonds, psychological contract	Trust can be among all the actors of the supply chain (farmer, supplier, manufacturer, buyer)	Relational governance	40%
Collaboration (C)	Information sharing, relational bonds, joint objectives, intellectual agility, trust, respect and commitment	Communication, sharing views on how to innovate, exchange production performance information, vertical collaboration, horizontal collaboration	Cooperatives collaborate with suppliers and buyers. Suppliers can collaborate with other competitive suppliers. Suppliers can collaborate with the buyers.	Relational governance	44%
Technology and information sharing (TIS)	Communication frequency and information quality.	Examples of technologies: efficient platform, EDI, email. Knowledge sharing	Information sharing among farmers, suppliers, buyers, cooperatives	Relational governance	40%
Assessment practices (AP)	Monitoring the supply chain for assuring more quality	Control, monitoring, quality assurance schemes, traceability	Third party, cooperatives, and the firm itself can do assessments	Preventive mechanism	20%
Governance Practices (GP)	Policies, favorable taxation, subsidies	Policies, favorable taxation, subsidies	The government is mostly responsible for favorable taxation, subsidies and policies. The company itself or buyers can integrate policies.	Preventive mechanism	20%

Planification et optimisation de la logistique de la construction dans les zones urbaines

IMEN JARROUDI, AMIN CHAABANE, MARC PAQUET

DÉPARTEMENT DE GÉNIE DES SYSTÈMES

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

MONTRÉAL, CANADA

IMEN.JARROUDI.1@ENS.ETSMTL.CA, AMIN.CHAABANE@ETSMTL.CA, MARC.PAQUET@ETSMTL.CA

Résumé – Une forte demande pour les projets de construction est attribuée en majeure partie à la croissance de la population en milieu urbain. Faisant face à des défis en logistique, de nombreux entrepreneurs voient leurs chaînes d'approvisionnement perdre coordination et efficacité, augmentant ainsi leurs frais d'exploitation. Cela est sans compter la nature même de l'industrie de la construction, cette dernière étant souvent décrite comme complexe, fragmentée et manquant d'intégration.

Nous étudions la chaîne d'approvisionnement des projets de construction en zones urbaines, et ce en la modélisant tel un problème MILP. Ainsi, nous nous concentrerons sur le rôle des centres de consolidation comme une solution logistique permettant d'optimiser les opérations logistiques et de minimiser les coûts en mettant l'accent sur la nécessité d'une collaboration entre les différents acteurs. De telles solutions demeurent encore relativement inexploitées dans le secteur de la construction, surtout par rapport à d'autres industries. En nous concentrant sur le cas d'une chaîne d'approvisionnement de construction multi période, multi ressources et multi projets à deux échelons, nous optimisons le processus de décision quant à la sélection de fournisseurs et l'allocation des centres de consolidation, ainsi que la planification d'un tel type de projet.

Mots clés - gestion de la chaîne d'approvisionnement de la construction, logistique de construction, modèle d'optimisation, centre de consolidation de la construction, sites encombrés

1 INTRODUCTION

Le rapprochement de la population vers le milieu urbain est une tendance qui ne fait qu'accroître au fil du temps. On prévoit que 68 % de la population mondiale résidera en zones urbaines d'ici à l'an 2050 [United Nations-Department of Economic and Social Affairs, 2018]. Dans la province de Québec, la population atteindrait 9 millions d'habitants en 2030, soit 0,6 million de plus que l'an 2019 [Gouvernement du Québec, ministère des Transports, 2019]. Alors que 80 % des Canadiens vivent en milieu urbain, 50 % de la population québécoise vivent dans les régions métropolitaines de Montréal [Nazarnia, Schwick et A.G.Jaeger, 2016].

En vue de cette tendance, le nombre de projets de construction augmente afin de répondre au besoin de logement des habitants. En effet, on a compté 111 projets de ce type à Montréal en 2018, dont 39 % étaient des immeubles résidentiels au coût de 30 milliards de dollars en investissements [Ville Montréal, 2018]. De plus, l'industrie de la construction représente une partie significative dans l'économie du Canada, soit 23,4 milliards de dollars en 2010 dont une grande partie est consacrée à l'achat de matériaux et aux opérations logistiques [Statistics Canada, 2018].

La logistique urbaine englobe la distribution et la gestion des marchandises depuis et vers la zone urbaine. Ce type de livraison comprend les biens consommables, la collecte des déchets et la logistique de construction [Boudoin, Morel et Gardat, 2013 ; Cardenas et al., 2017]. Les agglomérations

urbaines sont sujettes à des flux logistiques importants. En effet, on estime que 20 % à 30 % des véhicules HGV qui entrent dans les villes sont des camions transportant des matériaux de construction [CIVIC Project, 2018 ; Guerlain, Renault et Ferrero, 2019]. De plus, la densité de l'environnement urbain rend problématique l'exécution de ces activités. Par ailleurs, les projets d'envergure sont soumis à un espace de stockage limité sur le chantier, ce qui met les matériaux de construction à risque de dommages ou de perte. Une telle éventualité réduirait la productivité du chantier et augmenterait le risque pour les travailleurs qui s'y trouvent.

À cause de la fragmentation de l'industrie de la construction et la particularité de son produit, étant donné les différents types de matériaux et de méthodes de construction utilisés, les entreprises de construction opèrent généralement de manière indépendante et leurs projets font habituellement usage d'une chaîne d'approvisionnement distincte [Scott Wilson Ltd, 2010]. Les projets de construction à grande échelle se concrétisent à travers des partenariats entre plusieurs parties prenantes ayant chacun des intérêts différents. Pourtant, leurs collaborations demeurent limitées, leur rendant difficile d'atteindre une durabilité dans les opérations logistiques [Radziszewska-Zielina et Kania, 2017]. Néanmoins, la formation de ces partenariats peut réduire la fréquence des retards, les dépassements de budget des projets et éliminer les défauts moyens et majeurs des projets [Meng, 2012]. Les opérations logistiques dispersées engendrent généralement des frais excessifs pour les fournisseurs et les entrepreneurs, limitant

ainsi les possibilités d'établir une alliance à long terme entre les partenaires. De plus, les livraisons non consolidées de bien de construction aux taux d'occupation des véhicules plus faibles ont un impact négatif immédiat sur l'efficacité de l'infrastructure et des transports, entraînant des trajets plus longs et une plus grande contribuant à la congestion, à la pollution et aux bruits.

Au cours des dernières décennies, d'énormes efforts ont été déployés pour détourner les dysfonctionnements de l'industrie de la construction. L'utilisation de technologies avancées telles que BIM (Building information modeling) s'est avérée efficace pour encourager la collaboration entre les parties prenantes. En outre, une attention particulière a été accordée aux UCC (centres de consolidation urbains). Une approche prometteuse suggérant des bénéfices potentiels à long terme en améliorant l'efficacité et l'efficacité des processus logistiques tout en minimisant les conflits entre les parties prenantes

Dans ce travail, nous nous concentrons sur l'évaluation de l'impact d'une solution de collaboration logistique du point de vue de l'industrie de la construction. Nous mettons en œuvre une approche quantitative pour optimiser la chaîne d'approvisionnement d'un réseau de projets de construction. De plus, nous étudions l'impact de l'intégration des centres de consolidation de la construction dans la chaîne logistique dans le contexte urbain. Nous proposons de repenser la chaîne d'approvisionnement en y ajoutant un point focal : un centre de consolidation. Les centres de logistiques sont un sujet de recherche fort populaire au sein d'autres industries, telles la manufacture et la vente au détail. C'est un concept qui s'est avéré efficace pour optimiser les livraisons au dernier kilomètre. Les solutions logistiques telles que TPL et les centres de consolidation restent une nouvelle approche dans l'industrie de la construction. Par conséquent, nous nous concentrons sur le rôle que prendrait le centre de consolidation pour synchroniser les opérations du chantier de construction vis-à-vis des tâches principales d'une chaîne d'approvisionnement.

Notre principal objectif est d'optimiser la chaîne d'approvisionnement en construction urbaine en établissant un schéma de collaboration. Afin d'atteindre cet objectif, il faut élaborer un modèle de prise de décision qui présente les particularités du projet de construction urbaine :

- Stockage limité en chantier ;
- Multiples projets récurrents ;
- Variété de types de matériaux.

2 REVUE DE LA LITTÉRATURE

Dans d'autres industries, l'intégration de la chaîne d'approvisionnement a su amener des bénéfices permettant d'améliorer le processus de prise de décision dans la planification de la production et les achats [Akintoye, McIntosh et Fitzgerald, 2000]. Dans le secteur de la construction, la mise en place d'une chaîne d'approvisionnement intégrée est décrite tel un défi, d'autant plus que les entreprises de construction choisissent souvent de fonctionner de manière indépendante. Une façon de mettre en œuvre l'intégration est de gérer de manière conjointe le flux, le stockage et la distribution des matériaux [Stevens G., 1989 ; Stevens G., 2016]. Cela peut être très apprécié dans le secteur de la construction où les matériaux de construction sont consommés en grande quantité. Une solution consciente et économique telle que l'intégration de la chaîne d'approvisionnement peut être très valorisée et peut déclencher un changement de culture chez les entrepreneurs, les

transporteurs, les fournisseurs et les sous-traitants.

Tserng, Yin et Li (2006) ont développé une solution à deux échelons où le stock est géré par le vendeur pour les opérations de production et d'approvisionnement d'armature en acier. Le modèle de programmation par contraintes qu'ils ont développées a permis de minimiser les niveaux de stocks pour l'ensemble des membres de la chaîne d'approvisionnement et de réduire les coûts en intégrant les opérations de production et d'approvisionnement. Fang et Ng (2011) argumentent que la planification des flux de matériaux est un facteur important dans la réduction des coûts logistiques des matériaux. Les auteurs ont su identifier les différentes composantes du coût en appliquant la méthode ABC, afin d'analyser les opérations logistiques du béton préfabriqué. Faisant ainsi, ils ont pu générer un outil de gestion des coûts systémiques et valider leur travail en utilisant un algorithme générique pour établir la planification de l'approvisionnement en matériel.

Certaines études portaient spécifiquement sur le problème de la congestion et de l'espace limité sur le site dans les immeubles de grande hauteur. Said et El-Rayes (2011) et Said et El-Rayes (2013) ont étudié l'impact des limitations d'espace dans un projet à grande échelle sur les activités logistiques. Les auteurs soutiennent que le stockage des matériaux est lié aux activités d'approvisionnement, en particulier lorsque l'espace de stockage sur le chantier est limité. Par conséquent, l'intégration de la planification du stockage des matériaux est essentielle afin d'améliorer la productivité de la construction et réduire les coûts logistiques. Ces auteurs se concentrent sur les problèmes d'aménagement du site pour résoudre les problèmes d'espace de stockage limité dans les immeubles de grande hauteur. Liu, Xu et Zhang (2015) ont alors élaboré un modèle de chaîne d'approvisionnement en matériaux pour permettre aux acteurs du projet de contrôler les coûts opérationnels.

Dans leurs recherches Ma, Yan, Kang et Wei (2016) ont présenté une méthode de programmation à deux niveaux basés sur la théorie aléatoire floue pour décrire l'incertitude des coûts de transport. Le modèle permet de minimiser le coût de production et le coût de distribution. Ils ont résolu le problème en utilisant un algorithme génétique hybride à deux niveaux, puis ils ont appliqué leur modèle développé dans un projet hydroélectrique. En outre, Liu, Xu et Qin (2017) ont également utilisé la logique floue pour modéliser les incertitudes dans les retards d'approvisionnement, les commandes urgentes, et les changements de la demande. Ainsi, leur modèle intégré optimise à la fois les coûts du fabricant et les niveaux de services dans un projet hydroélectrique. Sanjay et Tindwani (2017) ont appliqué une approche quantitative pour démontrer la praticabilité d'une méthode d'optimisation dans le but d'obtenir des plans logistiques rentables au sein d'un projet de construction de route.

Dell'Amico et Novellani (2017) ont discuté de l'intégration de la chaîne d'approvisionnement de la construction en centralisant la chaîne d'approvisionnement à l'aide de centres de consolidation. Dans le cadre d'une initiative européenne visant à limiter l'impact négatif de la logistique de la construction en zone urbaine, le modèle à deux niveaux est résolu à l'aide de l'algorithme L-Shaped. Le modèle résout le problème de l'emplacement des installations et optimise les flux de matériaux et les stocks, tant en amont qu'en aval et au centre logistique. Lu et al., (2018) ont développé un modèle intégré de réapprovisionnement et d'allocation des stocks. Le modèle d'inventaire centralisé utilise des centres de consolidation afin de minimiser les coûts d'acquisition de stock et les coûts de retard du projet, ainsi que le coût des activités des projets. Le modèle intègre les activités sur le site

de construction avec les activités logistiques, en optimisant la politique d'allocation des matériaux et le niveau de stock de bases au centre de consolidation. Le modèle mathématique développé prend en considération les incertitudes dans le rendement de l'offre et la demande non stationnaire. Dans leurs études Jaskowski, Sobotka et Czarnigowska (2018) optimisent le coût total de gestion des stocks en utilisant un modèle de programmation linéaire. Le cadre proposé génère le plan d'approvisionnement optimal de deux types de matériaux différents et prend en considération le stockage limité sur site d'un projet de sous-bas routier. Hsua, Angeloudis et Aurisicchio (2018) se sont concentrés sur le problème de la fabrication modulaire, les auteurs ont développé un modèle de programmation stochastique qui établit les plans optimaux pour la production, le transport et la gestion des stocks dans la construction modulaire. Les auteurs étudient l'impact des variations de la demande sur la planification de construction.

Une étude récente de Fang et Ng (2019) fait appel aux algorithmes génétiques et à la méthode ABC pour élaborer des plans logistiques pour la production, l'approvisionnement et la consommation de matériaux de construction. Leurs recherches prennent en compte la planification des projets, la planification de la production ainsi que les délais de livraison des produits de construction préfabriqués. Deng, Gan, Das, Cheng et Anumba (2019) ont proposé un modèle intégré qui combine le Nous étudions le rôle qui pourrait être joué par les CCC une fois intégré comme point de stockage pour les matériaux en vrac. Ainsi, dans cet exemple, nous considérons un problème NP-complet à deux échelons, un problème d'optimisation MILP pour gérer le fonctionnement d'une CSC multi projets, multi ressources, multi fournisseurs et multi périodes qui est intégré au problème de l'emplacement des installations. GIS et le BIM en 4D, afin de déterminer la sélection optimale de plusieurs variables logistiques. Ces derniers sont la sélection des fournisseurs, l'identification de nombre optimal de livraisons et l'attribution de centres de consolidation dans les régions encombrées. Cette solution est générée tout en minimisant le coût logistique total de la construction. Plus récemment, Golpîra (2020) a utilisé la programmation linéaire en nombres entiers qui incorpore une stratégie d'inventaire gérée par les fournisseurs. La solution développée optimise des tâches de la chaîne d'approvisionnement comme les décisions relatives à l'emplacement des installations, le flux de matériaux dans chaque échelon, la sélection des fournisseurs et les tâches d'affectation. Le modèle MILP optimise la chaîne d'approvisionnement de construction multi projets, multi ressources et multi fournisseurs.

Dans notre travail, nous utilisons une solution quantitative pour étudier le schéma de coordination et son impact potentiel sur la chaîne d'approvisionnement de la construction. Nous mettons l'accent sur la planification de la chaîne d'approvisionnement ainsi que l'intégration de ses acteurs dans la planification de l'offre et de la demande. Nous partageons nos trouvailles dans le but d'améliorer la collaboration et la confiance entre les parties prenantes.

3 DESCRIPTION DE PROBLÈME

Dans notre étude, nous proposons une chaîne d'approvisionnement intégré avec des centres de consolidation de la construction tirant parti de la collaboration et de la coopération entre les différents acteurs de SC. En général, l'opérateur de CCC gère les opérations logistiques des sites de construction telles que l'approvisionnement en matériaux, la livraison, le stockage, et la coordination des livraisons entre les différents sites de construction. Les décisions sont

communiquées aux contractants par le biais de documents officiels. L'utilisation de la solution du centre de consolidation est facultative pour certaines livraisons de matériaux. L'opérateur de CCC est traité comme un décideur focal. En matière de logistique de construction, les décideurs focaux doivent identifier les coûts optimaux pour faire fonctionner le projet, y compris les coûts de commande, les coûts de contrôle, les coûts de transport et les coûts de stockage des matériaux [Fang et Ng 2019].

Les entrepreneurs travaillant sur les chantiers de construction passent des commandes pour leurs besoins en matériaux auprès des fournisseurs et demandent que la livraison des matériaux en vrac soit faite au CCC et pas au chantier de construction. Une fois les matériaux sont prêts, les entrepreneurs reçoivent un avis du CCC. Les entrepreneurs passent alors un ordre de livraison, avec le CCC pour les matériaux dont ils avaient besoin en juste à temps (Transport for London, 2008). L'entrepreneur principal gère plusieurs projets simultanés, de nature similaire, appartenant à la même entreprise de construction où la coordination de leurs opérations peut être bénéfique. Traditionnellement, chaque projet est géré individuellement. Nous proposons un schéma coordonné pour ces projets simultanés.

Par la suite l'approche d'optimisation choisie est utilisée pour étudier l'impact des centres de consolidation sur la chaîne d'approvisionnement de la construction dans le cadre d'une zone urbaine encombrée, la figure 1, ci-après montre plus en détail la structure de la chaîne d'approvisionnement étudié

3.1 Ensembles et indices

- S Ensemble de fournisseurs, indice $s \in S$
- M Ensemble de matériaux, indice $m \in M$
- M^1 Ensemble de matériaux de type 1 qui peuvent être transportés à l'entrepôt et /ou directement envoyé vers le chantier de construction, indice $m \in M^1$
- M^2 Ensemble de matériaux de type 2 qui sont envoyés directement vers le chantier, indice $m \in M^2$
- S^1 Ensemble de fournisseurs qui produisent les matériaux de type 1, indice $s \in S^1$
- S^2 Ensemble de fournisseurs qui produisent les matériaux de type 2, indice $s \in S^2$
- T Ensemble de périodes de planification, indice $t \in T$
- I Ensemble de sites de construction, indice $i \in I$
- Q Ensemble d'emplacements potentiels pour les CCC, indice $k \in Q$
- A_i Ensemble d'activités de chantier i , indice $j \in A_i$
- $P_{(i,j)}$ Ensemble d'activités antécédentes de l'activité $j \in A_i$ sur le chantier $i \in I$, indice $h \in P_{(i,j)}$

3.2 Paramètres

- T_{ms} Coût de transport par taille de lot de fournisseurs $s \in S^1$ aux CCC pour les matériaux $m \in M^1$
- I_m Coût de transport par unité pour les matériaux $m \in M^1$ de CCC au chantier de construction pour n'importe quelle période
- DC_{ms} Coût de transport direct par unité des fournisseurs $s \in S$ au chantier de construction pour les matériaux $m \in M$ pour toute période
- D_{ms} Prix d'achat réduit auprès des fournisseurs $s \in S^1$ pour les matériaux $m \in M^1$ pour toute période $t \in T$

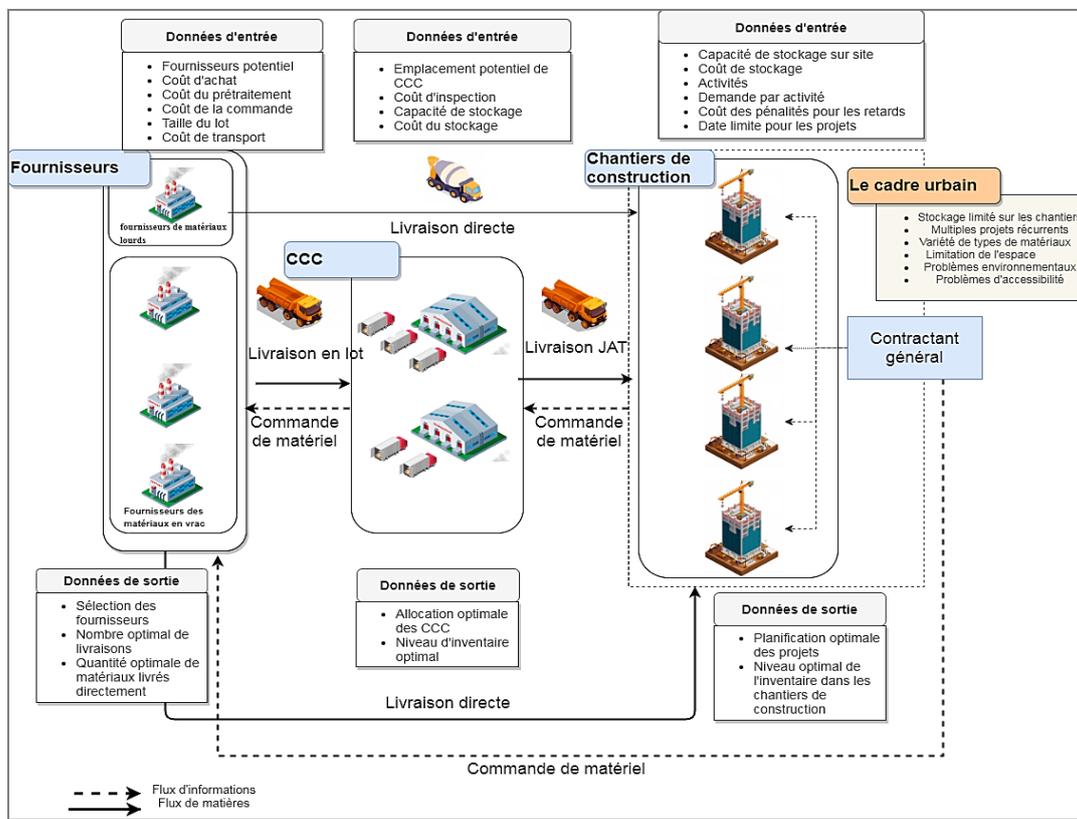


Figure 1. Chaîne d'approvisionnement avec les centres de consolidations

k_m	Coût de prétraitement par unité pour le chargement des matériaux $m \in M$ pour toute période	hc_{mi}	Coût unitaire de stockage des matériaux $m \in M$ sur le chantier $i \in I$ pour toute période
U_{sm}	Capacité du fournisseur $s \in S$ pour les matériaux $m \in M$	Inh_{in}	Stock initial pour matériaux $m \in M$ sur le chantier $i \in I$ à la période $t \in T$
FS_{ms}	Coût fixe : coût administratif/coût d'approvisionnement pour la commande des matériaux $m \in M$ auprès des fournisseurs $s \in S$	3.3 Variables de décisions	
Q_{ms}	Taille de lot fixe proposée par le fournisseur $s \in S^1$ pour les matériaux $m \in M^1$	a_{skmt}	Nombre de livraisons de camions pour les matériaux $m \in M^1$ des fournisseurs $s \in S^1$ au CCC, $k \in Q$ à la période $t \in T$
h_{mk}	Coût de stockage des matériaux $m \in M^1$ au CCC $k \in Q$ pour toute période	X_{simt}	Quantité de matériaux $m \in M^1$ des fournisseurs $s \in S^1$ envoyées au le chantier $i \in I$ à la période $t \in T$
V_m	Coût d'inspection des matériaux $m \in M^1$ au CCC $k \in Q$	O_{simt}	Quantité de matériaux $m \in M^2$ des fournisseurs $s \in S^2$ envoyées au chantier $i \in I$ à la période $t \in T$
F_K	Capacité du CCC $k \in Q$	y_{kimt}	Flux de matériaux $m \in M^1$ du CCC, $k \in Q$ au chantier $i \in I$ à la période $t \in T$
f_{wk}	Coût fixe de location et frais administratifs pour le CCC $k \in Q$	Iw_{kmt}	Niveau de stock des matériaux $m \in M^1$ dans le CCC $k \in Q$ à la période $t \in T$
Z_{im}	Capacité du chantier $i \in I$ pour les matériaux $m \in M$	Ib_{imt}	Niveau de stock des matériaux $m \in M^1$ dans le chantier $i \in I$ à la période $t \in T$
DD_i	Date d'échéance du chantier $i \in I$	Ih_{imt}	Niveau de stock des matériaux $m \in M^2$ dans le chantier $i \in I$ à la période $t \in T$
μ_{ij}	Durée de l'activité $j \in A_i$ du chantier $i \in I$	φ_{ijt}	Variable binaire qui est égale à 1 si l'activité $j \in A_i$ du chantier $i \in I$ commence à $t \in T$
δ_{ijm}	Demande de matériaux $m \in M^1$ d'activité $j \in A_j$ pour chantier $i \in I$	ϑ_{ijt}	Variable binaire qui est égale à 1 si l'activité $j \in A_i$ du chantier $i \in I$ se termine à $t \in T$
p_i	Pénalité d'avoir un chantier de construction $i \in I$ en retard après la date d'échéance DD_i et une période		

W_K Variable binaire qui est égale à 1 si CCC est situé à l'emplacement $k \in Q$

S_{sm} Variable binaire qui est égale à 1 si le fournisseur $s \in S$ est sélectionné pour le matériau $m \in M$

Sur la base des éléments définis ci-dessus, notre proposition de formulation mathématique de problème est la suivante :

3.4 Fonction objective

$$\begin{aligned} \text{Min} \sum_{m \in M^1 \cup M^2} \sum_{s \in S^1 \cup S^2} F S_{ms} S_{ms} + \sum_{s \in S^1} \sum_{k \in Q} \sum_{m \in M^1} \sum_{t \in T} (K_m + D_{ms}) Q_{ms} a_{skmt} + \sum_{s \in S^1} \sum_{i \in I} \sum_{m \in M^1} \sum_{t \in T} (K_m + P_m) \\ + \sum_{s \in S^1} \sum_{i \in I} \sum_{m \in M^1} \sum_{t \in T} (K_m + P_{ms}) O_{simt} + \sum_{s \in S^1} \sum_{k \in Q} \sum_{m \in M^1} \sum_{t \in T} T_{ms} a_{skmt} \\ + \sum_{s \in S^1} \sum_{i \in I} \sum_{m \in M^1} \sum_{t \in T} DC_{ms} X_{simt} + \sum_{s \in S^2} \sum_{i \in I} \sum_{m \in M^2} \sum_{t \in T} DC_{ms} O_{simt} + \sum_{k \in Q} \sum_{i \in I} \sum_{m \in M^1} \sum_{t \in T} I_m y_{ij} \\ + \sum_{k \in Q} f w_k W_k \\ + \sum_{s \in S^1} \sum_{k \in Q} \sum_{m \in M^1} \sum_{t \in T} V_m a_{skmt} Q_{ms} + \sum_{s \in S^1} \sum_{k \in Q} \sum_{m \in M^1} \sum_{t \in T} \left(\frac{I w_{km(t-1)} + a_{skmt} Q_{ms} + I w_{kmt}}{2} \right) \\ + \sum_{k \in Q} \sum_{i \in I} \sum_{s \in S^1} \sum_{m \in M^1} \sum_{t \in T} \left(\frac{I b_{im(t-1)} + y_{ikmt} + X_{msit} + I b_{imt}}{2} \right) h_{cim} \\ + \sum_{s \in S^1} \sum_{i \in I} \sum_{m \in M^1} \sum_{t \in T} \left(\frac{I h_{im(t-1)} + O_{simt} + I h_{imt}}{2} \right) h_{cim} + \sum_{i \in I} \sum_{t=DD_i+1}^{T} p_i (t - (DD_i + 1)) \theta_i \end{aligned} \quad (1)$$

3.5 Les contraintes

$$\sum_{k \in Q} a_{skmt} Q_{ms} + \sum_{i \in I} X_{simt} \leq U_S S_{sm} \quad \forall s \in S^1, \quad \forall m \in M^1, \forall t \in T \quad (2)$$

$$\sum_{i \in I} O_{simt} \leq U_S S_{sm} \quad \forall s \in S^2, \quad \forall m \in M^2, \forall t \in T \quad (3)$$

$$\sum_{m \in M^1} I w_{kmt} \leq F_K W_K \quad \forall k \in Q, \forall t \in T \quad (4)$$

$$I b_{imt} \leq Z_{im} \quad \forall i \in I, \forall m \in M^1, \forall t \in T \quad (5)$$

$$I h_{imt} \leq Z_{im} \quad \forall i \in I, \forall m \in M^2, \forall t \in T \quad (6)$$

$$I w_{kmt} = 0 \quad \forall m \in M^1, \forall k \in Q, \forall t = 0 \quad (7)$$

$$I w_{km(t-1)} + \sum_{s \in S^1} a_{mskt} Q_{ms} = I w_{kmt} + \sum_{i \in I} y_{ikmt} \quad \forall m \in M^1, \quad \forall k \in Q, \quad \forall t \in \{1, T_{N-1}\} \quad (8)$$

$$I w_{km(t-1)} + \sum_{s \in S^1} a_{mskt} Q_{ms} = \sum_{i \in I} y_{ikmt} \quad \begin{matrix} m \in M^1, \\ \forall k \in Q, \\ \forall t = T_N \end{matrix} \quad (9)$$

$$I b_{imt} = 0 \quad \forall m \in M^1, \forall i \in I, \forall t = 0 \quad (10)$$

$$I b_{im(t-1)} + \sum_{k \in Q} y_{ikmt} + \sum_{s \in S^1} X_{simt} = I b_{imt} + \sum_{j \in A_i} \delta_{ijm} \varphi_{ijt} \quad \begin{matrix} \forall m \in M^1, \\ \forall i \in I, \\ \forall t \in \{1, T_{N-1}\} \end{matrix} \quad (11)$$

$$I b_{im(t-1)} + \sum_{k \in Q} y_{ikmt} + \sum_{s \in S^1} X_{simt} = \sum_{j \in A_i} \delta_{ijm} \varphi_{ijt} \quad \begin{matrix} \forall m \in M^1, \\ \forall i \in I, \\ \forall t = T_N \end{matrix} \quad (12)$$

$$I h_{imt} = I N h_{imt} \quad \forall m \in M^2, \forall i \in I, \forall t = 0 \quad (13)$$

$$I h_{im(t-1)} \sum_{s \in S^2} O_{simt} = I h_{imt} + \sum_{j \in A_i} \delta_{ijm} \varphi_{ijt} \quad \begin{matrix} \forall m \in M^2, \\ \forall i \in I \\ \forall t \in \{1, T_{N-1}\} \end{matrix} \quad (14)$$

$$I h_{im(t-1)} + \sum_{s \in S^2} O_{simt} = \sum_{j \in A_i} \delta_{ijm} \varphi_{ijt} \quad \begin{matrix} \forall m \in M^2, \\ \forall i \in I, \forall t = T_N \end{matrix} \quad (15)$$

$$\sum_{s \in S} S_{sm} = 1 \quad \forall m \in M \quad (16)$$

$$\sum_{t \in T} t \varphi_{int} + \mu_{ih} \leq \sum_{t \in T} t \varphi_{ijt} \quad \forall i \in I, \forall j \in A_i \quad (17)$$

$$\sum_{t \in T} t \varphi_{ijt} + \mu_{ij} = \sum_{t \in T} t \vartheta_{ijt} \quad \forall i \in I, \forall j \in A_i \quad (18)$$

$$\sum_{t \in T} \varphi_{ijt} = 1 \quad \forall i \in I, \forall j \in A_i \quad (19)$$

$$\sum_{t \in T} \vartheta_{ijt} = 1 \quad \forall i \in I, \forall j \in A_i \quad (20)$$

$$\sum_{i \in I} \sum_{m \in M^1} y_{ikmt} \leq N W_K \quad \forall k \in Q, \forall t \in T \quad (21)$$

$$a_{skmt} \leq N W_K \quad \forall s \in S^1, \forall m \in M^1 \forall k \in Q, \forall t \in T \quad (22)$$

$$X_{simt} \geq 0 \quad \forall s \in S^1, \forall i \in I, \forall m \in M^1, \forall t \in T \quad (23)$$

$$a_{mskt} \geq 0 \quad \forall s \in S^1, \forall k \in Q, \forall m \in M^1, \forall t \in T \quad (24)$$

$$O_{simt} \geq 0 \quad \forall s \in S^2, \forall i \in I, \forall m \in M^2 \forall t \in T \quad (25)$$

$$I w_{kmt} \geq 0 \quad \forall k \in Q, m \in M^1, \forall t \in T \quad (26)$$

$$y_{ikmt} \geq 0 \quad \forall k \in Q, \forall i \in I, \forall m \in M^1, \forall t \in T \quad (27)$$

$$I b_{imt} \geq 0 \quad i \in I, \forall m \in M^1, \forall t \in T \quad (28)$$

$$I h_{imt} \geq 0 \quad i \in I, \forall m \in M^2, \forall t \in T \quad (29)$$

$$\vartheta_{ijt} \in \{0,1\} \quad \forall i \in I, \forall j \in A_i, \forall t \in T \quad (30)$$

$$\varphi_{ijt} \in \{0,1\} \quad \forall i \in I, \forall j \in A_i, \forall t \in T \quad (31)$$

$$S_{sm} \in \{0,1\} \quad \forall s \in S, \forall m \in M \quad (32)$$

$$W_K \in \{0,1\} \quad \forall k \in Q \quad (33)$$

La fonction objective optimise le coût logistique total de construction qui est composé de quatre éléments de coût. Le coût d'approvisionnement qui est la somme de la passation des commandes, du prétraitement des livraisons et le coût d'achat de matériaux. Nous insistons sur l'importance de l'escompte appliqué lors de l'achat de matériaux en grande taille de lot. La deuxième partie représente le coût du transport à la fois pour les livraisons directes et les livraisons par le centre de consolidation de la construction. La troisième partie représente le coût de stockage, ce coût est composé du coût fixe du centre de consolidation de la construction, les coûts variables pour le stockage des matériaux sur les sites de construction et le coût

de stockage dans consolidation et le coût d'inspection des matériaux arrivant au centre de consolidation de la construction. Nous adoptons la méthode d'inventaire moyen lors de la modélisation de l'inventaire. La dernière partie du coût logistique total de construction est le coût de la pénalité à appliquer lorsqu'on dépasse de l'échéance du projet de construction.

Les contraintes (2) et (3) imposent les seuils des quantités des matériaux disponibles chez les fournisseurs. La contrainte (4) a pour effet de limiter les capacités de stockage chez les CCC. Les contraintes (5) et (6) signifient que l'inventaire dans les sites de construction ne doit pas dépasser une certaine valeur. Les contraintes (10) jusqu'à (15) assurent l'initialisation d'inventaire et la satisfaction de demande des sites de construction pour les deux types de matériaux. La contrainte (16) impose la sélection d'un seul fournisseur par matériau. L'ensemble des contraintes (17), jusqu'au (20) définissent les contraintes de planification des projets de construction. La contrainte (17) définit la relation de préséance entre les activités de même projet. La contrainte (18) impose la date d'achèvement d'une activité, et les contraintes (19) et (20) imposent que chaque activité débute et termine à une date unique. Les contraintes (21) et (22) sont des contraintes logiques. Enfin les contraintes (23), jusqu'au (29) imposent que les variables de décisions relatives aux flux et aux stocks soient entières et positives et que les variables restantes soient binaires (30), (31), (32) et (33).

4 EXPERIMENTATION

Nous considérons un réseau de chaîne d'approvisionnement pour évaluer la performance de notre modèle représentée dans le chapitre précédent. Nous voulons démontrer l'aptitude de modèle à optimiser la chaîne d'approvisionnement et la planification logistique d'un réseau de projets de construction dans une zone urbaine encombrée.

La chaîne d'approvisionnement est constituée de douze sites, dont six fournisseurs, deux centres de consolidation potentiels et quatre sites de construction. La chaîne d'approvisionnement est simplifiée à cause de la complexité des problèmes NP-complet ce qui signifie que des instances relativement larges ne peuvent pas être résolues de manière optimale par un algorithme de temps polynomial. On considère quatre matériaux en totalité, dont un m1 fait partie des matériaux qu'on doit livrer directement vers les chantiers, et les trois restants m2, m3 et m4 peuvent être envoyés directement ou à partir des CCC. Afin de mettre en œuvre le modèle proposé, et fournir une estimation réaliste, un projet de construction d'un bâtiment résidentiel a été choisi, les données collectées sont partiellement basées sur l'étude de Saïd et El-Rayes (2013). On considère que ce projet de construction renferme six activités fictives. Les opérations logistiques impliquent l'achat, le stockage et la livraison de (4) matériaux : 1) barre d'armature m1, 2) blocs de maçonnerie m2, 3) panneaux de cloison sèche m3 et 4) tuiles de céramique m4. Ces matériaux sont généralement requis pour les activités au premier étage, les activités typiques au rez-de-chaussée et les activités en extérieur pour les immeubles de grande hauteur dans les zones urbaines. Pour cela nous utilisons des données approximatives et réelles recueillies dans la littérature et les sources de données pertinentes.

4.1 Logiciel utilisé

Pour le développement du modèle mathématique, nous utilisons le solveur LINGO17. Notre ordinateur utilisé pour les expérimentations possède les caractéristiques suivantes :

Windows 7 Enterprise 2009 (64 bits) comme système d'exploitation, Intel Core, i7-2600, 3.4 GHz comme processeur et 16 Go de RAM. Le modèle génère 1926 variables entières et 2131 contraintes. Le modèle est résolu en utilisant l'algorithme Branch and Bound en 24 secondes en moyenne.

4.2 RÉSULTAT

La centralisation de la chaîne d'approvisionnement et l'intégration des activités sur site avec les activités logistiques de l'ensemble de la chaîne d'approvisionnement ont permis de réduire le coût total de 4,18 %. En fait, le coût de l'approvisionnement est passé de 2 987 371 \$ à 2 712 412 \$, ce qui représente une baisse de près de 274 959 \$; comme le montre le tableau 1. Nous pouvons également constater que l'utilisation du centre de consolidation peut générer des économies de transport. Cela peut s'expliquer par le fait que nous utilisons désormais une stratégie de coordination lors du transport des matériaux de construction des centres de consolidation vers les chantiers.

Comme prévu, la coordination des activités d'approvisionnement entraîne une augmentation des coûts de stockage, avec une hausse de plus de 30 % par rapport au scénario de base. En outre, on peut constater que la coordination des activités logistiques mène à des retards dans les projets de construction, ce qui se traduit par l'apparition d'un coût de retard dans la structure des coûts logistiques dans le scénario d'intégration de la chaîne d'approvisionnement.

Afin de comparer l'impact de l'intégration de la chaîne d'approvisionnement sur le flux de matériaux, nous avons choisi le chantier 3 où il y a des délais dans les travaux de planification. Nous avons également pris comme référence les matériaux m2 qui sont en grande partie livrés par le CCC. On notera que le flux entrant de m2 à chaque site de construction est la somme du flux direct et des livraisons du centre de consolidation de la construction.

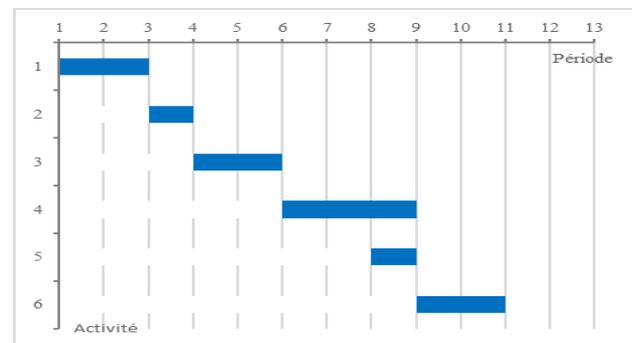


Figure 2. Diagramme de Gantt– site de construction 3 scénario de base

Dans le chantier de construction 3, la consolidation du flux de matériaux génère un retard de deux périodes. Il est certain que la coordination du flux de matériaux peut entraîner des retards des certains projets.

Dans le premier scénario : scénario de base, représenté dans la figure 2 les planifications des différents chantiers non coordonnées. La première activité de construction commence à la période 1, ce qui peut entraîner des opérations de livraison peu efficaces. Cependant, dans le deuxième scénario : scénario de coordination des projets, représentée dans la figure 3 on peut voir l'impact de la stratégie consolidation, on observe que le fait que la première activité ne soit pas lancée à la période 1 provoque un effet domino pour les autres activités sur le chantier.

Tableau 1. Comparaisons entre les deux scénarios

	Scénario de base	Scénario d'intégration de CCC	Économies de coûts	Économies de coûts en pourcentage
Coût d'approvisionnement	2 987 371,09 \$	2 712 412,89 \$	274 958,20 \$	9,20 %
Coût de transport	833 270,15 \$	798 555,72 \$	34 714,43 \$	4,17 %
Coût d'entreposage	411 529,92 \$	540 689,77 \$	129 159,85 \$	-31,39 %
Coût de retard	0 \$	3 600,00 \$	3 600,00 \$	0,00 %
Coût total	4 232 171,17 \$	4 055 258,39 \$	176 912,78 \$	4,18 %

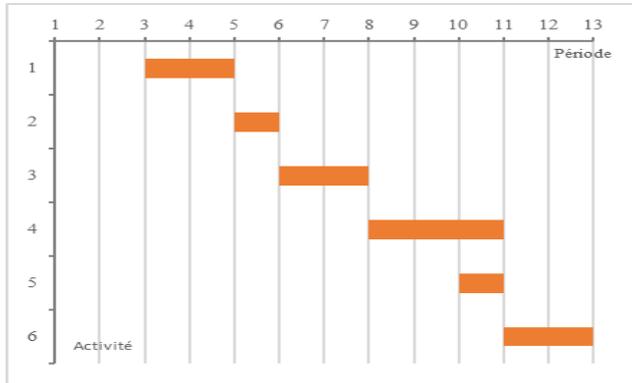


Figure 3. Diagramme de Gantt – site de construction 3 scénario avec collaboration

4.3 Analyse de sensibilité

Nous modifions le rabais offert sur les achats de grandes quantités de lots. Ainsi, nous plaçons ce paramètre dans une plage de 0 à 24 % de rabais appliqués sur les prix offerts par les fournisseurs dans le scénario de base. La figure 4 présente le nombre total d'unités du type de M¹ livrées directement et les quantités livrées par le centre de consolidation aux quatre sites de construction, pour l'ensemble de la période de planification par apport aux demandes totales. Nous pouvons nettement voir que l'escompte offert par les fournisseurs influence la décision de méthode de livraison des matériaux de construction. En fait, nous pouvons observer un tournant à partir de laquelle, les livraisons de matériaux de construction sont désormais effectuées par le centre de consolidation de la construction, cette valeur correspond à 8 % dans notre étude. Comme le révèle également notre analyse précédente, une fois que le centre de consolidation est en activité, il est plus judicieux de l'utiliser pour répondre à la majorité des demandes des chantiers de construction et pour effectuer la majorité des opérations logistiques.

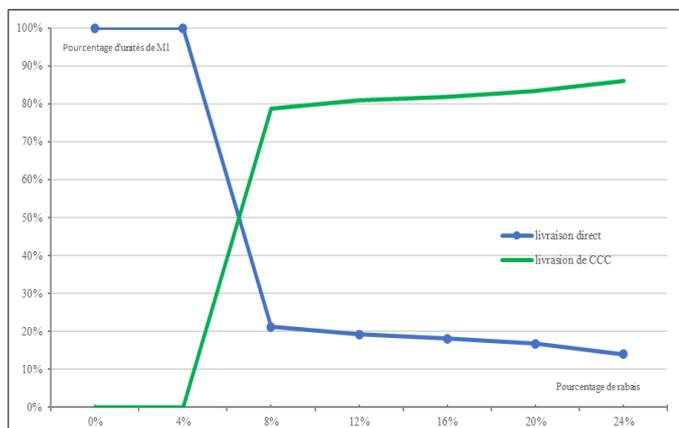


Figure 4. Impacts du rabais offert par les fournisseurs sur la méthode d'approvisionnement

5 CONCLUSION

Nous avons présenté le résultat du développement d'un modèle d'optimisation MILP pour la logistique de construction congestionnée. À cet égard, nous générons la décision optimale de sélection des fournisseurs, l'allocation optimale des CCC, les niveaux de stock optimaux dans les CCC ainsi que sur le site de construction, et la planification optimale des activités de construction. Nous avons généré les plans logistiques optimaux qui minimisent le coût logistique total tout en minimisant les impacts négatifs de la congestion dans les zones urbaines en réduisant le flux direct de matériaux et en détournant plutôt le flux vers le centre de consolidation de la construction pour des opérations logistiques mieux coordonnées. Cela garantit un flux plus faible vers la zone urbaine.

Le lecteur doit être conscient que nous optimisons la logistique de la construction du point de vue du destinataire, mais que nous parvenons à inclure les parties prenantes impliquées dans nos modèles. Nous avons pu étudier l'impact des parties impliquées sur la solution logistique proposée. Nous avons proposé deux scénarios différents et nous avons sélectionné des paramètres liés au coût de transport, au facteur d'actualisation. Les résultats suggèrent que le succès du centre de consolidation de la construction est fortement corrélé avec l'implication des parties prenantes et la validité de la proposition de valeur dans le modèle commercial proposé.

6 RÉFÉRENCES

- United Nations-Department of Economic and social affairs. (2018, May 16) 2018-revision-of-world-urbanization_prospects.dRécupéré sur <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>
- Government du Québec, ministère des Transports. (2019). transports.gouv.qc. Récupéré sur [road-freight-intervention-framework.pdf](https://www.transports.gouv.qc.ca/en/Documents/road-freight-intervention-framework.pdf)
- Nazarnia, N., Schwick, C., & A.G. Jaeger, J. (2016). Accelerated urban sprawl in Montreal, Quebec City, and Zurich: Investigating the differences using time series 1951–2011. *Ecological Indicators*, 1229–1251.
- Ville.Montreal. (2018, Juillet). Montreal.ca. Récupéré sur http://ville.montreal.qc.ca/pls/portal/docs/PAGE/MTL_stats_fr/media/documents/chantiers_juillet2018.pdf
- Statistics Canada. (2018, January 17). statcan. Récupéré sur <https://www150.statcan.gc.ca/n1/pub/11-402-x/2011000/chap/construction/construction-eng.htm>

- Boudoin, D., Morel, C., & Gardat, M. (2013). Supply Chains and Urban Logistics Platforms. dans J. Gonzalez-Feliu, Sustainable Urban Logistics: Concepts, Methods and Information Systems (pp. 1-20). Springer-Verlag Berlin Heidelberg.
- Cardenas, I., Borbon-Galvez, Y., Verlinden, T., Voorde, E. V., Vanelander, T., & Dewulf, W. (2017). City logistics, urban goods distribution and last mile delivery and collection. sage, 22-43.
- CIVIC project. (2018). Smart construction logistics.
- Guerlain, C., Renault, S., & Ferrero, F. (2019). Understanding Construction Logistics in Urban Areas and Lowering Its Environmental Impact: A Focus on Construction Consolidation Centres. sustainability.
- Radziszewska-Zielina, E., & Kania, E. (2017). Problems in Carrying Out Construction Projects in Large Urban Agglomerations on the Example of the Construction of the Axis and HighFive Office Buildings in Krakow. Theoretical Foundation of Civil Engineering.
- Scott Wilson Ltd. (2010). Freight Consolidation Centre Study Final Report. Glasgow.
- Meng, X. (2012). The effect of relationship management on project performance in construction. *International Journal of Project Management*, 188-198.
- Akintoye, A., McIntosh, G., & Fitzgerald, E. (2000). A survey of supply chain collaboration and management in the UK construction industry. *European Journal of Purchasing & Supply Management*, 159-168.
- Stevens, G. (1989). Integrating the supply chain. *International Journal of Physical Distribution*, 3-8.
- Tserng, H., Yin, S., & Li, S. (2006). Developing a resource supply chain planning system for construction projects. *Journal of Construction Engineering and Management*.
- Fang, Y., & Ng, S. T. (2011). Applying activity-based costing approach for construction logistics cost analysis. *Construction Innovation*, 259-281.
- Said, H., & El-Rayes, K. (2011). Optimizing Material Procurement and Storage on Construction Sites. *Journal of Construction Engineering and Management*, 421-431.
- Said, H., & El-Rayes, K. (2013). Optimal utilization of interior building spaces for material procurement and storage in congested construction sites. *Automation in Construction*, 292-306.
- Liu, Q., Xu, J., & Zhang, Z. (2015). Construction supply chain-based dynamic optimisation for the purchasing and inventory in a large scale construction project. *European J. of Industrial Engineering (EJIE)*.
- Ma, Y., Yan, F., Kang, K., & Wei, X. (2016). A novel integrated production-distribution planning model with conflict and coordination in a supply chain network. *Knowledge-Based Systems*, 119-133.
- Liu, Q., Xu, J., & Qin, F. (2017). Optimization for the integrated operations in an uncertain construction supply chain. *IEEE Transactions on Engineering Management*, 400-414.
- Sanjay, C. C., & Tindwani, A. (2017). Logistics optimisation in road construction project. 158-179.
- Jaśkowski, P., Sobotka, A., & Czarnigowska, A. (2018). Decision model for planning material supply channels in construction. *Automation in Construction*, 235-242.
- Sanjay, C. C., & Tindwani, A. (2017). Logistics optimisation in road construction project. 158-179.
- Dell'Amico, M., & Novellani, S. (2017). A two-echelon facility location problem with stochastic demands for Urban Construction Logistics An application within the SUCCESS project. *IEEE international conference on service operations and logistics, and informatics*, (pp. 90-95). Bari, Italy.
- Lu, H., Wanga, H., Xie, Y., & Wang, X. (2018). Study on construction material allocation policies: A simulation optimization method. *Automation in Construction*, 201-212.
- Hsua, P.-Y., Panagiotis, A., & Marco, A. (2018). Optimal logistics planning for modular construction using two-stage stochastic programming. *Automation in Construction*, 47-61.
- Deng, Y., Gan, V. J., Das, M., Cheng, J. C., & Anumba, C. (2019). Integrating 4D BIM and GIS for Construction Supply Chain Management. *Journal of Construction Engineering and Management*.
- Fang, Y., & Ng, S. T. (2019). Genetic algorithm for determining the construction logistics of precast components. *Engineering, Construction and Architectural Management*, 2290-2306.
- Golpıra, H. (2020). Optimal integration of the facility location problem into the multi-project multi-supplier multi-resource Construction Supply Chain network design under the vendor managed inventory strategy. *Expert Systems with Applications*.
- Tan, K. C. (2001). A framework of supply chain management literature. *European Journal of Purchasing & Supply Management*, 7(1), 39-48.
- Terzi, S., & Cavaliere, S. (2004). Simulation in the supply chain context: a survey. *Computers in industry*, 53(1), 3-16.
- Chan, F. T., Chung, S. H., & Wadhwa, S. (2004). A heuristic methodology for order distribution in a demand driven collaborative supply chain. *International Journal of Production Research*, 42(1), 1-19.
- Stadtler, H., & Kilger, C. (2002). *Supply chain management and advanced planning* (Vol. 4). Springer-Verlag.
- Stock, G. N., Greis, N. P., & Kasarda, J. D. (1998). Logistics, strategy and structure. *International Journal of Operations & Production Management*.

Conception d'un modèle de simulation pour l'évaluation de la pertinence de l'Internet Physique pour les chaînes d'approvisionnement humanitaires

MANON GREST¹, MATTHIEU LAURAS¹, BENOIT MONTREUIL²

¹ IMT MINES ALBI

CAMPUS JARLARD 81000 ALBI, FRANCE

MANON.GREST@MINES-ALBI.FR & MATTHIEU.LAURAS@MINES-ALBI.FR

² GEORGIA INSTITUTE OF TECHNOLOGY

NORTH AVE NW, ATLANTA, GA 30332, ETATS-UNIS

BENOIT.MONTREUIL@ISYE.GATECH.EDU

Résumé – La livraison de produits de première nécessité aux victimes des catastrophes naturelles représente un véritable défi pour les organisations humanitaires. Critiquées pour un manque d'efficacité, elles doivent aujourd'hui chercher à se réinventer. L'approche Internet Physique, qui propose une réorganisation de la logistique pour une meilleure gestion des ressources et coordination des acteurs tout en répondant aux exigences croissantes des parties prenantes et environnementales, pourrait s'avérer pertinente pour le secteur humanitaire. Cependant, cette hypothèse demande à être vérifiée au travers d'une démarche scientifique rigoureuse. Dans cette perspective, l'article propose la structuration d'une expérience basée sur un modèle de simulation de flux afin d'évaluer la pertinence de l'application de concepts de l'Internet Physique à un contexte humanitaire. Les étapes définies, inspirées des concepts de l'ingénierie des systèmes, permettent de garantir à la simulation son adéquation avec l'objectif de l'expérience et sa fiabilité.

Abstract – In the aftermath of a disaster, supplying relief items to affected people is a real challenge for humanitarian organizations. Showing a lack of effectiveness in the response, their practices are put into question. The Physical Internet approach, aiming at reorganizing logistics in a way for improved resource management and better coordination of actors while meeting the stakeholders and environment growing expectations, seems relevant for the humanitarian sector reinvention. However, this assumption needs to be proved through a rigorous scientific process. In this perspective, the paper offers to frame the design of an experiment based on a flow simulation model to evaluate the relevance of the Physical Internet concepts application in a humanitarian context. The steps defined, inspired by the system engineering field, ensure that the simulation meets the objective while ensuring the experiment reliability.

Mots clés - chaîne d'approvisionnement humanitaire, internet physique, simulation, ingénierie des systèmes

Keywords – humanitarian supply chain, physical internet, simulation, systems engineering

1 INTRODUCTION

Lorsqu'une catastrophe se produit, des opérations de secours humanitaires sont organisées afin de sauver des vies et de soulager au maximum les souffrances des personnes affectées. Cependant, la réponse apportée par les acteurs humanitaires souffre d'un manque d'efficacité (Chiappetta Jabbour et al., 2017; Haavisto and Goentzel, 2015). En effet, l'urgence de la situation, l'environnement perturbé et le manque de visibilité sont des difficultés additionnelles à gérer en plus de l'organisation des secours elle-même (Tomasini and Wassenhove, 2009). Face à ce constat, les pratiques des Organisations Humanitaires (OHs) sont remises en question. Avec une approche en rupture de la logistique actuelle, l'Internet Physique (IP), un récent concept destiné à répondre aux problématiques économiques, sociales et environnementales rencontrées dans un contexte de chaîne d'approvisionnement, semble présenter des atouts intéressants pour le secteur humanitaire (Montreuil et al., 2012). En effet, en redéfinissant la manière dont les objets sont déplacés, déployés, réalisés, fournis, conçus et utilisés, l'IP entend rendre la logistique plus efficace, durable, intelligente, agile, adaptable,

évolutive et résiliente (Ballot, 2014). L'idée d'une approche IP appliquée au secteur humanitaire a récemment été examinée, d'un point de vue conceptuel, par un petit nombre d'auteurs (Abdoulkadre et al., 2014; Grest et al., 2019). Aujourd'hui, tout l'enjeu réside dans la démonstration que la mise en œuvre des concepts issus de l'IP dans les chaînes d'approvisionnement humanitaires (CAHs) entraînerait des résultats positifs sur leur performance. Pour cela, une démarche d'expérimentation structurée s'impose et c'est là l'objet du présent article qui s'attache à décrire l'approche envisagée ainsi que la conception d'un modèle de simulation ad hoc permettant de mener à bien cette étude.

L'article est donc structuré comme suit : la section 2 est consacrée à une revue de littérature relative aux CAHs et aux grandes lignes du concept d'IP et propose un aperçu des caractéristiques d'une CAH hyperconnectée. La section 3 décrit quant à elle (i) la méthode proposée et (ii) le modèle de simulation suggéré. La partie 4 est dédiée à la conclusion et aux suites à mener.

2 REVUE DE LITTÉRATURE

2.1 Les chaînes d'approvisionnement humanitaire

Au fil des ans, le nombre et la complexité des catastrophes n'ont cessé d'augmenter et cette tendance devrait se poursuivre à l'avenir (Cozzolino, 2012). Les dommages causés privent ou rendent difficile l'accès aux populations à leurs biens et à leurs ressources et peuvent même mettre leurs vies en danger. Quand cela arrive, les populations peuvent compter sur l'aide des organisations humanitaires pour restaurer leurs dignités et les aider à retrouver des conditions de vie proches de l'avant incident (Kovacs and Spens, 2007). Pour répondre à cette demande, les OHs s'appuient sur leur logistique, qui pèse pour environ 80 % de l'ensemble des efforts à fournir (Van Wassenhove, 2006). Étant donné que des vies sont en jeu, les organisations doivent s'assurer de la complétion de leurs activités, et ce, de la manière la plus efficace possible. Bien qu'évident, cela est beaucoup plus complexe en pratique. En effet, le contexte particulier et les nombreux défis auxquels elles ont à faire face compromettent le bon déroulement des opérations. Pour en approfondir le sujet, les sections suivantes développent la manière dont les OHs gèrent aujourd'hui les catastrophes et les défis rencontrés lors de leurs activités.

2.1.1 La gestion des catastrophes

La gestion des catastrophes concentre un ensemble d'activités qui varient selon la phase de l'urgence (Kovacs and Spens, 2012). Tout d'abord, bien avant que la catastrophe ne se produise, on se trouve dans la phase de prévention où des actions peuvent être entreprises pour limiter les potentiels dommages humains et matériels. Toujours en amont de l'événement, mais avec une certaine certitude de son occurrence, c'est la phase de préparation. Elle consiste en la mise en place de mesures stratégiques pour faciliter par anticipation le secours. La phase de réponse débute lorsque la catastrophe survient. Elle concentre, en premier lieu, des activités d'assistance aux victimes comme l'accès à des soins, des aides psychologiques ou encore la remise en service des réseaux de communication qui requièrent l'organisation de la venue d'experts et de volontaires sur place. Dans un deuxième temps et en parallèle, l'assistance se fait au travers de l'apport de produits suivant les besoins comme des médicaments, de la nourriture, du matériel ou des produits de première nécessité de type couverture ou kit d'hygiène par exemple (Blecken, 2010). Enfin, la phase de réparation, dont le démarrage dépend de la fin de la situation d'urgence, comprend des actions visant à aider la population à se remettre de l'événement et à retrouver un cours proche de la normale (Kusumastuti et al., 2010). La section suivante détaille le processus d'approvisionnement humanitaire tel qu'il est généralement déployé lors de la phase de réponse.

2.1.2 Défis rencontrés & gestion des chaînes d'approvisionnement humanitaires

Lorsqu'une catastrophe se produit, le gouvernement ainsi que des OHs présentes sur le territoire prennent des mesures rapides pour assister les populations touchées. Lorsque l'ampleur des dégâts est trop importante, les instances locales peuvent faire appel à un soutien international (Weiss, 2016). En conséquence, le nombre d'acteurs peut rapidement être important et poser des problèmes de coordination (Van Wassenhove, 2006). De plus, chaque catastrophe requiert la formation d'un nouveau réseau d'acteurs (Salaun, 2017) où différents types de flux s'entrecroisent : matériel, information, finance. Tout d'abord, pour déplacer les articles de première nécessité aux zones touchées, le flux physique peut être amené à transiter par des

entrepôts (Pishvae et al., 2009) par le biais de divers moyens de transport. De nombreux échanges d'information ont également lieu dans pareille situation pour s'informer et prendre des décisions adéquates. Cependant, la grande incertitude quant à la survenue des catastrophes et quant à l'estimation de leurs impacts limite les OHs dans leurs possibilités d'anticiper la demande et de se préparer comme ils le souhaiteraient (Kovacs and Spens, 2007). Enfin, en ce qui concerne le flux financier, il provient généralement de donateurs souhaitant soutenir les organisations dans leur mission. Les dons peuvent être restreints dans leur utilisation par les contributeurs (Abdoulkadre et al., 2014), ce qui peut limiter le nombre de ressources disponibles. Également, la présence des médias a un véritable impact sur les conditions des opérations de secours. En effet, l'exposition médiatique tend à multiplier les dons qui s'avèrent souvent inappropriés entravant la livraison des marchandises véritablement utiles (Charles et al., 2010).

2.2 Vers une chaîne d'approvisionnement humanitaire hyperconnectée

Compte tenu des nombreux défis auxquels sont confrontées les OHs dans la poursuite leur mission et les difficultés rencontrées lors des opérations de secours après certains événements majeurs (Thomas and Kopczak, 2005) certains questionnements sur un plan organisationnel s'avèrent légitimes. À ce stade, a priori, deux pistes sont envisageables : soit optimiser l'existant et tenter de se rapprocher des niveaux de maturité des chaînes d'approvisionnement classique en avance de phase par rapport aux humanitaires ou bien accepter de remettre en cause certains fondamentaux.

L'idée de s'inspirer des chaînes d'approvisionnement plus traditionnelles ou commerciales a été explorée ces dernières années par le monde académique (Behl and Dutta, 2018). Il est reconnu que les deux types de chaînes ont des éléments en commun notamment au niveau des activités logistiques (gestion des achats, transport...), de la structure incluant les types de flux transitant (financier, information & physique), les installations et infrastructures (entrepôts, fournisseurs, moyens de transport...) ou encore les besoins (agilité, adaptabilité, flexibilité...). Cependant, elles se distinguent aussi sur de nombreux points (Beamon and Balcik, 2008). Notons d'abord les nuances existantes en matière d'objectifs. Là où les chaînes commerciales recherchent le profit et la satisfaction du client, les humanitaires se concentrent sur des services aux bénéficiaires permettant de sauver des vies, quel qu'en soit le prix. Ensuite, la demande est plus volatile et difficile à anticiper, car il est compliqué d'estimer où et quand une catastrophe va survenir et quels dégâts elle est supposée engendrer (Kovacs and Spens, 2007). La notion de temps est également différente puisque des vies en dépendent, il faut livrer rapidement dans un contexte humanitaire alors que les clients d'une chaîne commerciale peuvent se permettre un certain temps d'attente. Enfin et c'est probablement le plus important, la différence en matière de variété et du nombre d'acteurs à prendre part dans une chaîne. Dans un contexte commercial, les acteurs sont relativement connus et stables. Dans un contexte humanitaire, les acteurs sont variés (OHs, donateurs, associations, instances gouvernementales, etc.) et ont des objectifs différents, ce qui peut créer des conflits d'intérêts et représente un véritable défi (Beamon and Balcik, 2008). De plus, chaque catastrophe étant unique en raison du lieu, du moment et des dégâts générés ce qui induit la mise en place nouvelle chaîne d'acteurs à chaque fois (Tomasini and Wassenhove, 2009).

Récemment, il a été démontré que les recherches effectuées sur différentes thématiques concernant l'amélioration des CAHs

n'avaient pas produit d'impacts significatifs et n'avaient conduit qu'à peu de mises en pratique par les praticiens faute de réelle prise en compte des réalités terrains (Kunz et al., 2017). Sur cette base, nous avons choisi d'explorer une nouvelle piste dans cette optique d'amélioration qui se trouve être en rupture des schémas classiques et qui semble pouvoir répondre aux enjeux actuels.

2.2.1 Les intérêts de l'Internet Physique pour le secteur humanitaire

L'IP est un concept récent et innovant de réseaux logistiques interconnectés capitalisant sur la possibilité de partager les ressources et l'information (Montreuil et al., 2012). La définition de l'IP est la suivante : « L'Internet Physique est un système logistique global construit à partir de l'interconnexion des réseaux logistiques par un ensemble standardisé de protocoles de collaboration, de conteneurs modulaires et d'interfaces intelligentes pour une efficacité et une durabilité accrue » (Ballot, 2014). Comme mis en avant dans cette définition, l'IP propose une refonte des bases de la logistique. Le terme « interconnexion » fait ainsi référence à l'étroite et intensive connexion entre les acteurs et les composants du réseau. Le deuxième aspect clé de l'IP réside dans la volonté d'ouvrir les réseaux logistiques et de partager les actifs. Aujourd'hui, les entreprises forment des réseaux privés et relativement stables qui possèdent leurs propres entrepôts et véhicules en général. L'IP est en rupture avec cette logique et suppose que les actifs devraient être partagés entre tous les utilisateurs de ce nouveau réseau et utilisés en fonction des besoins. De plus, l'IP en tant que « réseau de réseaux » cherche à augmenter la visibilité et la possibilité de nouvelles connexions entre les acteurs.

Ce concept d'IP est né afin de satisfaire les exigences croissantes en matière d'environnement et de performance des services. En effet, le système logistique actuel présente des dysfonctionnements néfastes pour l'environnement et qui tendent à compromettre les objectifs des accords de Paris (Ballot, 2014). Avant qu'il ne soit trop tard, l'IP a été conçu afin d'offrir une chance aux services logistiques de se montrer plus résistants, efficaces, durables et plus adaptables pour ses utilisateurs en changeant la façon dont les objets physiques transitent à travers le réseau (Montreuil et al., 2012).

Il est à noter qu'un tel changement de paradigme de la logistique, comme suggéré par l'IP, requiert des transformations importantes, et ce, à divers niveaux. Tout d'abord, il sera nécessaire de développer des systèmes d'information adaptés et utilisant des technologies avancées pour permettre l'hyperconnexion des acteurs, un partage accru et standardisé de l'information et le stockage massif de données. Les impacts n'auront pas seulement lieu au niveau du fonctionnement logistique, mais également sur la manière dont les objets vont être conçus, produits et livrés (Montreuil et al., 2010). Ensuite, il conviendra de faire accepter aux acteurs d'évoluer vers ce nouveau mode de fonctionnement et d'engager des efforts dans ce sens. Les investissements attendus seront principalement d'ordre financier et de temps (Grest et al., 2019) et demanderont également une bonne organisation à différentes échelles. À ce sujet, le groupement ALICE (Alliance for Logistics Innovation through Collaboration in Europe) a d'ailleurs, récemment publié une feuille de route formulant les étapes importantes et les prérequis associés pour la mise en œuvre de l'Internet Physique d'ici à 2030 (ALICE-ETP, 2020). L'IP se met doucement en place, mais il demande à être plus largement connu et surtout convaincant pour être plus largement adopté, notamment au niveau humanitaire où il n'a pas encore

été réellement étudié. C'est donc tout l'enjeu de cet article que de montrer ou non les intérêts de l'IP pour le secteur humanitaire.

2.2.2 Définition d'une chaîne d'approvisionnement hyperconnectée

Cette section vise à présenter dans les grandes lignes de ce que présage une chaîne d'approvisionnement hyperconnectée. En d'autres termes ce à quoi ressemblerait une CAH où des concepts d'IP ont été appliqués dans le but de parvenir à un système plus efficace, efficient et durable. Compte tenu des défis à relever par les acteurs humanitaires, il y a en effet plusieurs domaines où l'IP pourrait être bénéfique. Premièrement, l'opportunité de partager les moyens permettrait d'optimiser les capacités de stockage et de transport utilisées et permettre certaines économies d'échelle. De plus, la mutualisation des stocks, c'est-à-dire permettre le stockage de produits appartenant à différentes organisations dans un même entrepôt, permettrait d'augmenter la couverture de chaque acteur et donc, sa capacité à porter assistance aux victimes et ainsi réduire les délais de livraison (Grest et al., 2019). La consolidation des transports (c'est-à-dire le remplissage des camions avec des colis provenant de différents expéditeurs) engendrerait une augmentation des taux d'utilisation tout en diminuant le nombre de camions mobilisés et donc, réduire principalement les coûts et les émissions de CO₂ (Grest et al., 2019). Un autre aspect clé où l'IP peut faire une différence se situe au niveau de la connexion des acteurs. En effet, en augmentant les possibilités de liaisons entre les différents acteurs humanitaires, la communication serait plus aisée ce qui favoriserait la collaboration. De plus, le développement d'un système centralisé, comme le suggèrent Abdoukkadre et al., permettrait de résoudre les problèmes d'inefficacité et de redondance des opérations. En effet, en centralisant et en partageant en temps réel les rapports de situation et les actions en cours à tous les acteurs, on peut s'attendre à voir s'améliorer la qualité des décisions prises (Abdoukkadre et al., 2014). La performance serait également suivie non plus à l'échelle d'une organisation, mais à celui de l'opération de manière plus globale. Par conséquent, la coordination et la collaboration entre les acteurs se verraient également être améliorées. Une telle plate-forme pourrait aussi aider à la gestion des dons ce qui augmenterait la disponibilité et l'adéquation des ressources. Constatez ici un début d'aperçu de formalisations des caractéristiques d'une CAH hyperconnectée. En effet, ce sujet est récent et il reste encore de nombreux aspects à examiner pour proprement définir un tel concept dont les possibilités sont larges.

2.3 L'ingénierie des systèmes

Sur un autre sujet, aujourd'hui, le nombre de systèmes complexes est en nette augmentation, et ce, au travers de différents domaines en lien avec le vivant (écosystèmes, systèmes vivants, climat...) jusqu'à des domaines plus « artificiels » comme les systèmes sociaux ou économiques, les villes ou encore Internet par exemple. La complexité s'explique par une demande croissante des utilisateurs pour des systèmes plus fiables et stables sur le plan de la performance et qui se doivent également d'être capables d'interagir avec une riche variété d'autres systèmes, eux aussi complexes. Par conséquent, ces systèmes sont plus compliqués à développer puisqu'ils nécessitent la coordination d'un grand nombre de participants issus de divers univers disciplinaires (Ramos et al., 2012). De là, force est de constater qu'il faille être attentif à la définition des besoins et à la conception de tels systèmes en amont. En effet, ces dernières sont des étapes cruciales, ayant un impact

sur les activités suivantes, et qui bien menées peuvent être source de certaines économies d'argent et de temps (Boehm, 1976). Dans cette perspective, le domaine de l'Ingénierie des Systèmes (IS) a vu le jour aux alentours de 1970 et est de plus en plus apprécié des chercheurs (Ramos et al., 2012) comme des ingénieurs (Cloutier, 2020). Destinée à accompagner le développement de systèmes complexes et interdisciplinaires pour assurer leur succès (Ramos et al., 2012), l'IS est basée sur une approche holistique (c'est-à-dire qu'elle considère le système dans son ensemble et ses interactions avec son environnement). Elle permet également de décrire le système de manière complète par le biais de différents angles de vue de modélisation (Ramos et al., 2012). En outre, elle présente les avantages d'améliorer la communication entre les parties prenantes en fournissant une base de langage commune et propose une méthodologie basée sur les processus (Cloutier, 2020). En effet, chaque étape du cycle de vie du système (c'est-à-dire de la définition des besoins jusqu'à la remise en main du système à son utilisateur) est décrite par un processus selon la norme ISO/IEC/IEEE15288:2015 (ISO et al., 2015). Enfin, l'IS intègre de nombreuses étapes de validation et de vérification visibles à chaque étape du cycle de vie du produit. Les étapes de validation d'un côté garantissent que le système serait en mesure répondre aux attentes et besoins des parties prenantes, et ce, dans le l'environnement opérationnel prévu. De l'autre, les étapes de vérification permettent de comparer le système par rapport au comportement attendu (ISO et al., 2015). Ces la combinaison de ces deux types d'interventions qui garantissent la conformité du système avec son cahier des charges.

3 PROPOSITIONS

3.1 Méthode expérimentale

Dans la perspective d'évaluer l'impact de l'IP appliqué au domaine des CAHs et tirer des conclusions fiables, et ce, avant même d'entreprendre toute expérience, nous devons nous assurer de la validité, crédibilité et la pertinence de l'expérience et de la simulation associée pour répondre à l'objectif.

En ce qui concerne la crédibilité, selon Law, elle vient de l'approbation du client quant à la conformation du comportement du système face à ses attentes. Cependant, le fait d'être crédible ne garantit pas que le modèle est valide tandis que la validité, elle, est nécessaire à établir de la crédibilité (Law, 2014). En ce sens, l'IS est un bon outil du fait de son approche holistique et des nombreuses étapes de validation qui assurent la conformité du système vis-à-vis de son commanditaire (ISO et al., 2015).

D'un autre côté, le sujet à observer, ici les CAHs, est reconnu comme un système complexe si l'on se base sur les critères suivants, développés par (Sheard and Mostashari, 2011). En premier lieu, les CAHs sont complexes par leur structure. En effet, elles concentrent une multitude d'acteurs différents, d'actifs associés (comme les entrepôts, véhicules...) et échanges pour subvenir aux personnes dans le besoin. Deuxièmement, la complexité provient de la dynamique du système. De nombreuses activités logistiques sont menées et doivent régulièrement s'adapter aux contraintes de l'environnement (par exemple, une route détruite entraînant un changement d'itinéraire...). Enfin, la présence des médias, des donateurs et des gouvernements génère de la complexité sociopolitique additionnelle au système (Tomasini and Wassenhove, 2009).

Par conséquent, sur la base de (i) la complexité des CAHs (ii) des attendus de recherche concernant la fiabilité de l'expérience à mener et (iii) les propriétés de l'IS permettant de traiter avec un système complexe et pluridisciplinaire tout en assurant la

vérification et la validation du système tout au long de son cycle de vie, l'IS semble une approche pertinente pour définir une méthode expérimentale fiable et développer la simulation associée pour réaliser notre étude.

À partir de ce constat, nous avons formalisé la méthode expérimentale reprenant une partie des étapes du processus de définition d'un système en IS (voir figure 1). Nous avons également tenu à conserver une des particularités de l'IS à savoir, la présence d'actions de validité à chaque étape. Cette méthode comprend donc trois étapes : la mission du système, l'architecture du système et la conception du système. La première étape vise à définir le problème traité et son contexte et ainsi, justifier de l'intérêt du système. L'étape d'architecture décrit les composants du système et leurs interactions entre eux. La dernière étape donne des indications quant aux prérequis et tâches à réaliser pour mener les expériences via la simulation.



Figure 1 Méthode expérimentale en trois étapes

3.2 Simulation

3.2.1 Mission de la simulation

Pour repositionner l'expérience dans son contexte ; il a été reproché aux organisations humanitaires, par le passé, d'un manque d'impact positif sur les sociétés recevant de l'aide ainsi qu'un manque d'efficacité (Haavisto and Goentzel, 2015). À cela, s'ajoutent des difficultés à se coordonner et cela a été principalement mis en avant lors de la gestion de la catastrophe qui a touché Haïti en 2010 (Noyer, 2010). La question qui se pose aujourd'hui est : comment améliorer la performance et les impacts lors des opérations de secours ? Jusqu'à peu, la recherche se concentrait principalement sur l'optimisation des ressources déjà présentes et cherchait à transférer certaines pratiques des chaînes d'approvisionnement classiques dans une optique d'amélioration (Kunz et al., 2017). Cet article propose une alternative de réorganisation plus en profondeur tirant inspiration de l'IP. L'hypothèse de l'étude est : une transformation des CAHs vers une organisation plus hyperconnectée impacterait positivement les résultats de performance des opérations menées par les OHs. L'objectif réside désormais dans la capacité de l'étude à prouver ou non que cette hypothèse est valide. Pour cela, nous faisons le choix de conduire une expérimentation quantitative afin d'observer la relation cause et effet de la réorganisation au travers d'indicateurs de performance. Les données utilisées seront issues de la littérature et du terrain afin de permettre de représenter au mieux la réalité et de fournir des résultats tangibles. Ces choix introduisent le besoin de concevoir une simulation permettant la mise en œuvre pratique des expériences.

3.2.2 Architecture de la simulation

Pour répondre à l'objectif énoncé précédemment, nous envisageons une expérimentation capable de comparer, au moyen d'indicateurs de performance, le comportement d'une

chaîne d'approvisionnement de référence (inspirée des pratiques actuelles) avec une chaîne de test, à savoir jumelle dans ses fondements, mais intégrant des concepts de l'IP. Afin d'étudier le comportement dans des contextes variés, comme c'est le cas à chaque fois qu'une catastrophe se produit, nous souhaitons soumettre les modèles à différents scénarios et ainsi observer leur réponse.

L'architecture de la simulation est donc composée de trois systèmes clés : le scénario, le système à l'étude (suivant le paramétrage pourra représenter la chaîne test ou bien de référence) et le tableau de suivi des indicateurs de performance (voir Figure 2).



Figure 2 Éléments composants l'architecture de la simulation

Le scénario

Le scénario fournit au système à étudier des indications quant au contexte de l'expérience. En effet, il définit le périmètre géographique (une région, un pays...) ainsi que la période considérée (semaine, mois, année...). Ces indications servent à l'initialisation de l'environnement du système. Le paramétrage intervient également au niveau du système lui-même concernant certaines hypothèses fixes (ex. : la vitesse des camions), mais aussi des choix d'options de variables à tester. Par exemple, on peut chercher à observer les conséquences d'un transport classique versus un transport consolidé. Cela demandera dans les deux cas de configurer le même environnement et hypothèses, mais de sélectionner une option de variable (par exemple, « classique »), lancer la simulation, récupérer les résultats puis avec ce même contexte, sélectionner l'autre option (par exemple, « consolidé ») afin de récupérer les données associées pour réaliser la comparaison.

Le système à l'étude

Selon l'Organisation Internationale de Normalisation (ISO), un système est une combinaison d'éléments en interaction organisés pour atteindre un ou plusieurs objectifs définis (ISO et al., 2015). Un élément est un composant d'un système voire d'un sous-système. Le terme « interaction » désigne la manière dont les éléments s'influencent mutuellement. Enfin, l'élément « objectif » de la définition réfère au fait de répondre à un besoin.

Pour commencer à définir le système à l'étude, nous pouvons nous appuyer sur l'une des définitions les plus citées relatives aux CAHs. En effet, comme la logistique représente 80% des activités lors des opérations de secours (Van Wassenhove, 2006), nous allons limiter le périmètre du modèle à ce niveau. Ensuite, selon Kovacs et Spens 2012, la logistique humanitaire englobe la planification et la gestion de toutes les activités en lien avec les flux physiques, d'informations et financiers. Il est important de noter que cela comprend également la coordination et la collaboration des membres de la chaîne, les prestataires de services et les organisations humanitaires (Kovacs and Spens, 2012). Sur cette base, on distingue que : (i) les éléments du système à l'étude se composent des activités logistiques et des différents acteurs (ii) les interactions sont représentées au travers des trois types de flux issus lors des efforts de coordination et de collaboration et tandis que (iii) l'objectif est, et c'est ce qu'ont en commun les différentes organisations, sauver des vies et soulager la souffrance (Haavisto and Goentzel, 2015).

Ensuite, comme le suggère l'approche holistique, la description de l'environnement et de son influence doivent être pris en compte. À cette fin, selon l'article de Kovacs et Spens (2009) relatif à l'identification des défis liés à la logistique humanitaire, il est mis en avant que les véritables défis rencontrés lors des catastrophes dépendent de leur type et de la région dans laquelle la catastrophe se produit (Kovacs and Spens, 2009). Les auteurs ont également souligné que les différentes phases de gestion de la crise ont un impact sur les activités logistiques. Enfin, les deux derniers défis identifiés précédemment (c'est-à-dire la variété des organisations et la coordination des acteurs) se doivent d'être pris en compte dans l'environnement du système.

À partir de cette enquête autour des éléments clés des CAHs, nous avons développé une illustration résumant les éléments composant le système à l'étude pour l'expérimentation (voir Figure 3). On y retrouve donc un sous-système dénommé « réponse humanitaire » et en interaction avec ce dernier, le sous-système de demande. À eux deux, ils forment un système d'intérêt qui va être central pour les observations. Ensuite, autour de ce dernier, se distinguent les sous-systèmes dits « territoire » et « catastrophe » qui sont considérés comme des systèmes externes faisant partie de l'environnement et qui influencent le comportement du système d'intérêt (Cloutier, 2020).

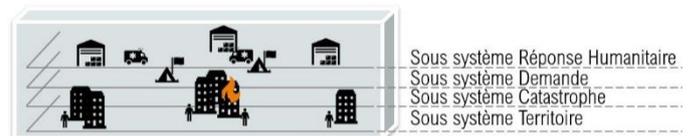


Figure 3 Système à l'étude et ses sous-systèmes clés dans un contexte de chaînes d'approvisionnement humanitaires

Le tableau de suivi de performance

Le tableau de suivi de performance concentre l'ensemble des indicateurs à consulter et est un élément clé pour statuer sur l'hypothèse définie précédemment et répondre à l'objectif fixé. En effet, chercher à améliorer la performance sous-entend de savoir au préalable ce qu'est la performance et qui plus est, dans un contexte humanitaire. La performance se définit au travers d'une mesure de résultats et d'objectifs fixés. Si les résultats sont en deçà de l'objectif, on peut parler de « contre-performance » alors qu'il y aura performance si les résultats dépassent les attentes. Maintenant, dans un contexte humanitaire qu'est-ce que cela signifie de manière plus précise ? Si l'on prend la définition de (Ramalingam et al., 2009), on retient que la performance dans un cadre humanitaire s'articule autour du fait de sauver des vies, soulager des souffrances et maintenir la dignité humaine, et ce, sans différenciation possible. Cet objectif s'applique à tout type d'organisation humanitaire, et ce, quelle que soit la phase. Il est également mis en avant que l'aide apportée doit être efficace et bien gérer l'utilisation des ressources (Ramalingam et al., 2009). De plus, il est généralement admis que les organisations qui suivent leur performance surpassent celles qui ne le font pas puisque cela permet entre autres de s'améliorer (Abidi et al., 2014). D'autres aspects sont également mis en avant comme le fait que la réponse doit être « agile » (Tomasini and Wassenhove, 2009), « flexible » (Kabra and Ramesh, 2016), ou encore « durable » (Salvadó and Lauras, 2017). De telles contraintes doivent être mesurées afin de s'assurer que le comportement est celui attendu.

Au niveau de la dynamique autour des éléments de l'architecture, les trois systèmes (scénario, système d'étude et

performance) s'influencent en séquence. Le scénario contraint le système en fournissant le contexte, les hypothèses et le paramétrage de certaines variables du système à l'étude. Le comportement du système observé fournit, durant la période définie par le scénario, des résultats de performance qui pourront ensuite être analysés à la fin de l'expérimentation.

3.2.3 Conception et prérequis à la simulation

Pour mettre en œuvre l'expérience imaginée, un certain nombre d'étapes ont été planifiées. Tout d'abord, et afin d'éviter tout biais dans l'expérience, il est nécessaire de définir en premier les indicateurs de performance pour comparer les résultats de comportements du système à l'étude. Il s'agira de les sélectionner de manière pertinente par rapport à l'objectif fixé et de définir la manière de les mesurer. Dans notre cas, les premières expérimentations seront évaluées au travers d'indicateurs tels que l'OTIF (On Time Delivery In Full) pour l'évaluation en termes d'efficacité, les coûts de transport et la valeur du stock pour la partie efficacité et les émissions de CO2 générées par les transports pour la partie durabilité.

Ensuite, il conviendra de décrire les scénarios que l'on souhaite soumettre au système. Dans un premier temps, il faut définir le contexte. Notre premier cas d'étude se trouve être Aceh, une province d'Indonésie et se concentre sur la période de juillet 2013 où un tremblement de terre de magnitude 6.1 a frappé. Ensuite, il est nécessaire de définir les hypothèses et les options possibles de certains éléments. À titre d'illustration, pour démarrer le développement de la simulation, nous considérons un seul type de camion de type cinq tonnes avec une vitesse constante de 50km/h. Puis, il conviendra de définir le plan d'expérience, c'est-à-dire les leviers clés d'une CAH sur lesquels nous souhaitons tester les effets de l'approche IP. Par exemple, au niveau des entrepôts, on peut envisager de tester la différence entre la version classique (où les bâtiments sont privés) versus la version hyperconnectée (où cette fois les infrastructures sont ouvertes et partagées). Il est donc attendu à en parallèle de cette étape, de lister les composants caractéristiques de chaque « type » de système. À titre d'illustration, sur le cas classique, on considère les infrastructures classiques (entrepôts, dépôt ...) alors que la version IP contiendra en plus des plates-formes de transbordement. Il sera également nécessaire de décrire les processus notamment ceux de la réponse humanitaire afin de décrire le séquençement d'activités qui rythment la gestion de la catastrophe ainsi que les critères pour la prise de décision.

Alors que les étapes précédentes s'intéressaient au développement d'un modèle pour la simulation, il reste désormais à coder l'expérience. Pour rappel, le choix de la simulation est justifié par le fait que c'est une des méthodes privilégiées dans le cas d'expérimentations sur un système complexe (Law, 2014). Pour cette dernière étape, on peut conseiller la rédaction de l'algorithme, au préalable de l'étape de codage, et la mise en place d'un planning de développement par niveau de complexité. Une étape transverse à toutes celles évoquées précédemment sera de s'attacher à ce que l'ensemble des choix et hypothèses faits soient validés par des praticiens dans le but d'assurer la crédibilité de l'expérience et la fiabilité des conclusions qui pourront être tirées.

structurer les étapes nécessaires à la réponse de l'objectif fixé à savoir l'amélioration de la performance lors des opérations humanitaires. Ces phases ont permis de déterminer un ordre d'actions à réaliser au préalable afin de réaliser la simulation nécessaire à l'expérience tout en garantissant la validité des résultats. Celle-ci sera réalisée à l'aide du logiciel AnyLogic qui permet la modélisation, entre autres de chaînes logistiques, selon différents degrés d'abstraction (événement discret, agent, dynamique des systèmes, et même de réaliser des hybrides). Dans un premier temps, la modélisation et programmation se concentreront sur la mise en forme du système de référence et de sa validité et s'en suivra, dans un second temps, la version hyperconnectée. Deux premiers aperçus de la version de référence sous AnyLogic sont visibles ci-dessous (voir Figure 4 & 5).

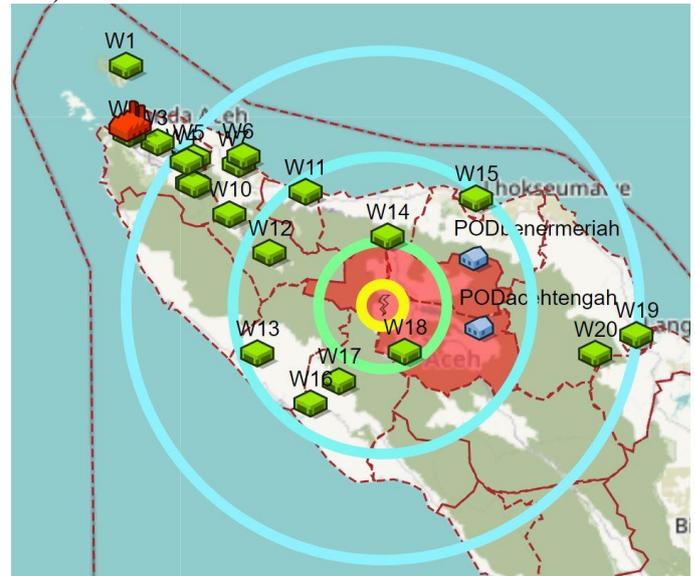


Figure 4 Aperçu des entrepôts identifiés sur la région d'Aceh en Indonésie et reconstitution d'un tremblement de terre historique affectant les populations des districts de Bener Meriah et Aceh Tengah

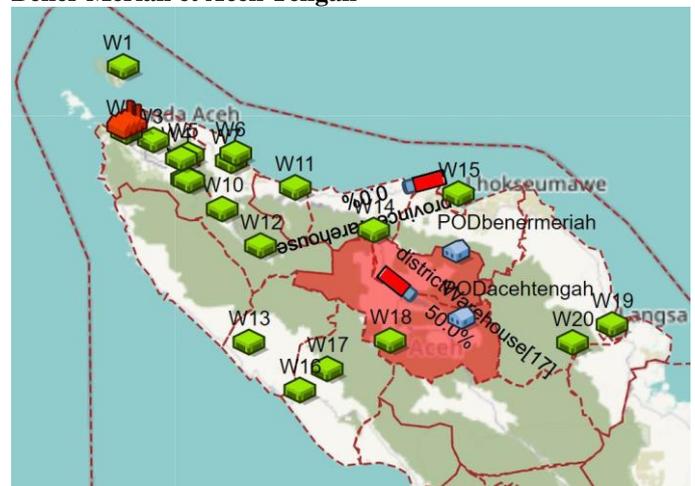


Figure 5 Aperçu de la livraison de produits de première nécessité à destination des points de distribution (POD) à la suite d'un tremblement de terre historique

5 REFERENCES

Abdoulkadre, A., Intissar, B.O., Marian, M., and Benoit, M. (2014). Towards Physical Internet Enabled Interconnected Humanitarian Logistics. 1st International Physical Internet Conference 30.

4 CONCLUSION & AGENDA DE RECHERCHE

L'ingénierie des systèmes développée, pour les systèmes complexes, multidisciplinaires dans le but de garantir le bon déroulement de conception et l'adéquation du système, est apparue comme une approche adaptée pour aider à la définition de l'expérience. En s'en inspirant, nous avons pu identifier et

- Abidi, H., de Leeuw, S., and Klumpp, M. (2014). Humanitarian supply chain performance management: a systematic literature review. *Supply Chain Management: An International Journal* 19, 592–608.
- ALICE-ETP (2020). ALICE Roadmap to Physical Internet. (Shenzhen, China), p.
- Ballot, E. (2014). L'Internet physique: le reseau des reseaux des prestations logistiques (Paris: la Documentation française).
- Beamon, B.M., and Balcik, B. (2008). Performance measurement in humanitarian relief chains. *International Journal of Public Sector Management* 21, 4–25.
- Behl, A., and Dutta, P. (2018). Humanitarian supply chain management: a thematic literature review and future directions of research. *Ann Oper Res*.
- Blecken, A. (2010). Supply chain process modelling for humanitarian organizations. *International Journal of Physical Distribution & Logistics Management* 40, 675–692.
- Boehm, B. (1976). Software Engineering. *IEEE Transactions on Computers* C-25, 1226–1241.
- Charles, A., Lauras, M., and Van Wassenhove, L. (2010). A Model to Define and Assess the Agility of Supply Chains: Building on Humanitarian Experience. *International Journal of Physical Distribution & Logistics Management* 40, 722–741.
- Chiappetta Jabbour, C.J., Sobreiro, V.A., Lopes de Sousa Jabbour, A.B., de Souza Campos, L.M., Mariano, E.B., and Renwick, D.W.S. (2017). An analysis of the literature on humanitarian logistics and supply chain management: paving the way for future studies. *Ann Oper Res* 283, 289–307.
- Cloutier, R.J. (2020). The Guide to the Systems Engineering Body of Knowledge (SEBoK) (Hoboken, NJ: SEBoK Editorial Board).
- Cozzolino, A. (2012). Humanitarian Logistics: Cross-Sector Cooperation in Disaster Relief Management (Berlin Heidelberg: Springer-Verlag).
- Grest, M., Lauras, M., and Montreuil, B. (2019). Toward Humanitarian Supply Chains Enhancement by using Physical Internet Principles. In 8th International Conference on Industrial Engineering and Systems Management, p.
- Haavisto, I., and Goentzel, J. (2015). Measuring humanitarian supply chain performance in a multi-goal context. *Jrnl Hum Log and Sup Chn Mnage* 5, 300–324.
- ISO, IEC, and IEEE (2015). ISO/IEC/IEEE 15288:2015 Ingénierie des systèmes et du logiciel - Processus du cycle de vie du système.
- Kabra, G., and Ramesh, A. (2016). Information Technology, Mutual Trust, Flexibility, Agility, Adaptability: Understanding Their Linkages and Impact on Humanitarian Supply Chain Management Performance. *Risk, Hazards & Crisis in Public Policy* 7, 79–103.
- Kovacs, G., and Spens, K. (2009). Identifying challenges in humanitarian logistics. *International Journal of Physical Distribution & Logistics Management* 39, 506–528.
- Kovacs, G., and Spens, K.M. (2007). Humanitarian logistics in disaster relief operations. *International Journal of Physical Distribution & Logistics Management* 37, 99–114.
- Kovacs, G., and Spens, K.M. (2012). Relief supply chain management for disasters: humanitarian aid and emergency logistics (Hershey, PA: Information Science Reference).
- Kunz, N., Wassenhove, L., Besiou, M., Hambye, C., and Kovács, G. (2017). Relevance of humanitarian logistics research: best practices and way forward. *International Journal of Operations & Production Management* 37, 1585–1599.
- Kusumastuti, R.D., Insanita, R., and Wibowo, S.S. (2010). Relief Logistics Practices in Indonesia: A Survey (Rochester, NY: Social Science Research Network).
- Law, A.M. (2014). Simulation Modeling and Analysis (Dubuque: McGraw-Hill Higher Education).
- Montreuil, B., Meller, R.D., and Ballot, E. (2010). Towards a Physical Internet: the Impact on Logistics Facilities and Material Handling Systems Design and Innovation. (Milwaukee, Wisconsin, USA), p.
- Montreuil, B., Meller, R.D., and Ballot, E. (2012). Physical Internet Foundations. *IFAC Proceedings Volumes* 45, 26–30.
- Noyer, O. (2010). L'urgence d'une logistique humanitaire internationale. *Les Echos*.
- Pishvaei, M.S., Jolai, F., and Razmi, J. (2009). A stochastic optimization model for integrated forward/reverse logistics network design. *Journal of Manufacturing Systems* 28, 107–114.
- Ramalingam, B., Mitchell, J., Borton, J., and Smart, K. (2009). Counting what counts: performance and effectiveness in the humanitarian sector | ALNAP (ALNAP).
- Ramos, A.L., Ferreira, J.V., and Barcelo, J. (2012). Model-Based Systems Engineering: An Emerging Approach for Modern Systems. *IEEE Trans. Syst., Man, Cybern. C* 42, 101–111.
- Salaun, V. (2017). Les logistiques temporaires : faire vite en prenant son temps. *Supply Chain Magazine* 80–81.
- Salvadó, L.L., and Lauras, M. (2017). Sustainable Performance Measurement for Humanitarian Supply Chain Operations. 9.
- Sheard, S.A., and Mostashari, A. (2011). Complexity Types: From Science to Systems Engineering. *INCOSE International Symposium* 21, 673–682.
- Thomas, A., and Kopczak, L.R. (2005). From logistics to supply chain management: the path forward in the humanitarian sector.
- Tomasini, R., and Wassenhove, L.V. (2009). Humanitarian Logistics (Springer).

Van Wassenhove, L.N. (2006). Humanitarian aid logistics: supply chain management in high gear. *Journal of the Operational Research Society* 57, 475–489.

Weiss, T.G. (2016). *Humanitarian Intervention* (John Wiley & Sons).

CIGI 2021

Logistics Network Design in a Product-Service System case

AGATHE MÉTAIREAU¹, VAN-DAT CUNG¹

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP*, G-SCOP

* Institute of Engineering Univ. Grenoble Alpes
38000, Grenoble, France

agathe.metaireau@grenoble-inp.org, van-dat.cung@grenoble-inp.fr

Résumé – De nos jours, les entreprises proposent de plus en plus d’offres intégrant des produits et leurs services associés plutôt que de vendre uniquement des produits. Les producteurs proposent alors des services comme la reprise, la gestion des consommables et des déchets ou encore la maintenance. Dans ce cas, ils doivent développer des réseaux logistiques pour supporter les activités associées à ces services. L’objectif de ce travail est de mener une étude de conception de réseaux logistiques dans le cadre de l’économie d’usage pour une entreprise en croissance. Pour ce faire, nous avons bâti un modèle de données permettant de représenter le cas étudié, et utilisé un programme d’optimisation afin de concevoir différents scénarios de réseaux logistiques permettant de répondre aux besoins de l’entreprise partenaire. Ce travail a permis de fournir aux entreprises une représentation du réseau logistique en fonction des activités de l’entreprise concernée, ainsi qu’une estimation des coûts globaux du système.

Abstract – Nowadays, companies offer more and more not just to sell products, but rather integrated products and services bundles. Producers offer then services as take-back, waste and refill management or even maintenance. In this case, they have to develop logistic networks to support the activities related to these services. The aim of this work is to carry out a study to design logistic networks as part of the Product-Service System for a growing company. To do so, we built a data model to represent the studied case and used an optimization program in order to design a logistic network compliant to the needs of the partner company. This work allowed us to provide a representation of the logistic network according to the activities of the company, as well as an estimation of the global costs of the system.

Mots clés – Economie d’usage, Conception de réseaux de service, Optimisation, Réseaux de logistique et transport, Entreprise en croissance.

Keywords – Product-Service System, Service networks design, Optimization, Logistics and transport networks, Growing companies.

1 INTRODUCTION

The aim of this work is to carry out a study, with the help of modelling and optimizing tools, to build a logistics and transport network that will support the activities of a company which has adopted the Product-Service System (PSS) business model. In addition, the network to be built is aimed at an on-growing company, so it will have to evolve according to the expansion of the company’s activities. We use a two-year-old start-up company located in Grenoble metropolitan area as a case study for the problem we explored, and they provided a certain amount of data used to experiment the method developed. At the time of the study, the company has one production site in the Grenoble metropolitan area. They have products located on more or less ten towns in France and wish to expand its activities on the French metropolitan territory in a five-years-horizon.

1.1 Context and Challenges

The business model studied is based on the PSS concept. This implies that the company does not sell the product, but its functions and services. This type of business model has the potential of improving product and service management for the company and product and service quality for its customers, but also sustainability. Companies are indeed responsible for a larger part of the life cycle of the product, which implies a better resource and waste management. In our case, the product offered by the company is similar to a vending machine. It

requires little care, except for maintenance and consumable refurbishment. The company offers a two-year rental agreement, during which the company is responsible for the wellness of the product, and in particular its installation and maintenance. This means that every time a failure occurs on an installed product, the company has to repair it within a certain time span. This business model is, according to the literature [Reim et al., 2015], a use-oriented PSS.

All the installation and maintenance activities are currently operated from the production site at the company. This strategy seems to be viable for a small amount of products, but it can quickly be expensive in terms of transport costs as and when the company’s activities expand. The company has indeed planned to install a given number of new products every year for a time horizon of five years. To sustain this expansion, the company wants to know how to design its logistics network in order to fulfil the distribution and maintenance functions. Therefore, a method is developed to design a logistics network to be used in such a context. The network will have to support the installation and maintenance activities at each phase of the company’s growth while optimizing the overall costs of the system. To do so, the main decisions to take are where to place the facilities serving the maintenance and distribution functions, and how much workforce to hire at every facility. The components of the problem must be clearly modelled in order to meet the requirements of a PSS business model. For instance, the company is required to repair every failure with a full service

level since with the PSS business model the company is responsible for the wellness of the product.

This work is to help a company with such business model take strategic decisions on its logistics network when designing a PSS offer.

1.2 Methodology

In order to help the company make good quality decision, we set up the following methodology. The first step is to study the academic work published on this topic. This will allow to position our work and have an insight on how related problems are handled. This study is presented in Section 2. Then, an important step is to formalize the problem. First, the problem has to be clearly defined. Then, we can express the requirements mathematically by building a data model in Section 3 and an Integer Linear Program (ILP) in Section 4 to represent and solve the problem we address. As the core problem we want to address is locating facilities and allocate customers to these facilities, we wrote the ILP based on an Uncapacitated Facility Location which target is precisely to answer such questions. Next, the relevance of the model has been tested on the data provided by the company. The experimental results on the case study are presented in Section 5. Thus, a sensitivity analysis has been conducted evaluate the robustness of the solutions. This sensitivity analysis also give insight on which parameters can influence the results. This is presented in Section 6. Finally, we explored the possibility of outsourcing some activities in order to suggest recommendations to the company. This is detailed in the Section 7.

2 RELATED WORK IN THE LITERATURE

The PSS business model studied is based on the fact that the producer keeps the ownership of the product and is responsible for its proper functioning. It is very similar to the business model of vending machines, as in this case the machine is made available in a public place and its owner has to refurbish it and repair it in case of failure. It is also close to the activities implemented in the context of reverse logistics, because in this case a product which has already been sold is reinjected in the supply chain to be resold are reused. After-sales services are also close to use-oriented PSS, because it implies replacing or repairing a product returned by a customer. This is why we chose to explore the literature related to these topics, in addition to articles related to the PSS.

2.1 Product-Service System

Several aspects of the PSS have been studied in the literature, and we focused on the ones related to our problem. On the logistics network design, [Sun et al., 2019] determine the configuration of a multi-level logistics network using a Location-Inventory problem modelled with a Mixed Integer Linear Program (MILP). Concerning costs estimation, [Shen et al., 2017] suggested a methodology to identify, classify and calculate costs inherent to a PSS business model.

Activities related to maintenance have been studied from different angles. [Derrouiche and Dubruc, 2014] designed a method to set up maintenance processes as part of a PSS offer. [Kuo, 2011] studied the parameters influencing the feasibility of such a business model. [Kuo et Wang, 2012] designed a model based on Multi-Attribute Utility Theory to optimize the service level of a network with several types of maintenance. [Rezapour et al., 2016] used a stochastic approach considering demand and production uncertainties to determine the optimal quantities of spare parts, semi-finished and finished goods to produce in order to satisfy the demand. [Xia et al., 2017] developed a model

allowing to optimize the preventive maintenance planning for leasing contracts. [van Ooijen et al., 2019] identified the parameters of the transport between customers and repair centres in order to determine the best strategy to adopt with a transport service provider.

2.2 Vending Machines

The literature about logistics aspects of the vending machine business model is mainly focused on tactical aspects. [Rusdiansyah and Tsao, 2005] built a non-linear mixed integer program called Integrated Inventory and Periodic Vehicle Routing Problem with Time-Window to answer the problem of periodic refurbishment of the vending machines. [Kim and Yoo, 2012] addressed the manpower question and built a non-linear program to calculate how many permanent and temporary technicians to hire in order to fulfil the tasks related to preventive and curative maintenance.

2.3 After-Sales Services

Literature on after-sales services is mainly focused on operations aspects. On another hand there are also a few studies pertaining to the information system required by such a business model. For example, [Sagama et al., 2016] studied several scenarios to determine the optimal preventive-curative maintenance strategy for a wind farm. [Chen et al., 2009] focused on the set up of an information system supporting after-sales maintenance.

2.4 Reverse Logistics

The literature related to reverse logistics is mainly about take-back activities. Most of the references are focused on the design of logistics networks required for such activities, and the most common approach is to use a model of Facility Location Problem type. Researchers use several declinations of this type of problem, like Uncapacitated Facility Location [Srivastava, 2008]), Capacitated Facility Location with a deterministic [Melo et al., 2009] or stochastic [Listes and Dekker, 2001] approach to optimize the positioning of the facilities.

Other approaches have also been used. [Cruz-Rivera and Ertel, 2009] used a Fixed Charge Facility Location model to place waste collecting facilities. This problem is also addressed with different strategies like discrete event modelling [Kara et al., 2007] or continuous optimization ([Fleischmann, 2001]).

About maintenance activities, [Amini et al., 2005] use a What-if analysis to decide where to stock spare parts, and a Facility Location based Integer Linear Program (ILP) to place technicians dedicated to maintenance.

2.5 Conclusion and Positioning

To the best of our knowledge, articles studying maintenance and PSS either concentrate on preventive maintenance, or search to optimize the parameters of an existing system, rather to build a new system supporting maintenance activities for a PSS business model. Articles pertaining to the other topics we identified as relevant are more focused on tactical or operational problematics than on strategical ones.

The problem addressed by our study is designing a logistics network for maintenance and distribution as part of a PSS. These three aspects (*i.e.* maintenance, logistics network design and PSS) have all been studied, but not altogether to the author's knowledge.

Furthermore, we need to consider the growing aspect of the network for a start-up company. To the best of our knowledge, the literature about this is rather scarce and does not apply to the case we study.

Facility location problems are sometimes extended to consider other logistics problematics. We can cite Location-Inventory, where facility location questions are integrated with inventory management decisions. We can also cite Location-Routing, where the routing between facilities is jointly addressed with their location.

In our case, these aspects are not taken into account, mostly because of the lack of data provided by the case study company. In fact, as the company is still relatively young, the current available data they have provide small amount of maintenance, installation and refurbishing demand.

3 DATA MODEL

To address the problem identified, we choose to adapt an Uncapacitated Facility Location (UFL) based Integer Linear Program (ILP) to our study. The UFL-ILP model lends well to the case study. As we stated before, the company currently operates all of the distribution and maintenance activities from its production site, in the Grenoble metropolitan area. But with the expansion of its activities, this strategy will be very expensive, mostly in terms of transport costs. As a result, the problem we address is how to locate maintenance and distribution sites on the covered territory so that the overall costs of the system will be optimized.

This requires building a data model which will represent the system and be the input of the UFL-ILP set up. The aim of this section is to explain this model.

3.1 Geographical repartition

To decide where to place facilities, there is a need to know where the customers – respectively the products – will be located. The start-up company does not have this information, at least not in long term. We propose to build an expansion scenario based on their commercial strategy over the French metropolitan territory. With the business plan on five years the company provided, this time horizon has been selected as well.

The initial state for the expansion was the roster of cities in which the company had already placed products in 2019. All these cities have been classified by region, and then the roster has been completed with the most populated cities of each French region, so that it contains three cities per region (Figure 1). The repartition of cities obtained is rather consistent with the French population density in 2019 (Figure 2). To be as close as possible of this repartition, several big regional cities such as Clermont-Ferrand, Limoges, Perpignan and Chambéry have been added to the roster and obtain the final geographical scenario (Appendix 1).

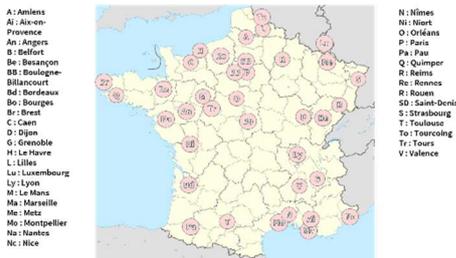


Figure 1. Map of the selected cities

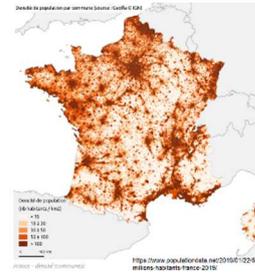


Figure 2. French population density in 2019

To decide in which year the products will be placed in the cities for the first time (i.e. activation year; in other words, year in which there is new clients in these cities), we apply a random draw to these cities and obtain a classification of the cities per activation year (Appendix 2).

3.2 Products repartition

To design an efficient logistics network, we also have to determine how much products will be located in the cities each year. To do so, the repartition of the products is modelled based on the initial state, communicated by the company. We estimate the number of products to be placed in every city each year using two parameters, that are the occupation rate (i.e. number of products in the city on total number of commercialised products the year before) and the population rate (i.e. population of the city regarding to the total population in the activated cities of the year in question). For each city j , every year t , let γ_j^t and δ_j^t be respectively those parameters, which are positive if the city is activated and zero otherwise. Those parameters are weighted by the proportions α_1 for the population rate and α_2 for the occupation rate. For every year t , let BP^t be the total number of products in all the cities (given by the company). So, for the city j , the year t , the number of products will be calculated as follows:

$$n_j^t = BP^t (\alpha_1 \times \delta_j^t + \alpha_2 \times \gamma_j^t).$$

Let us take a simple example to illustrate this formula. For a year t , let's assume we will have $BP^t = 150$ new products on the entire territory. Let's assume there was 50 products installed the year $t - 1$, and that the total population of the territory is 10000. Let's select a city j in which there is already $n_j^{t-1} = 10$ products. This city has 1000 inhabitants.

As a result, its occupation rate and population rate will be the following:

$$\gamma_j^t = \frac{10}{50} = 0.2,$$

$$\delta_j^t = \frac{1000}{10000} = 0.1.$$

Let us weight those parameters equally, in other words $\alpha_1 = \alpha_2 = 0.5$. Finally, the number of products to be placed in the city j at year t will be:

$$n_j^t = 150 (0.5 \times 0.1 + 0.5 \times 0.2) = 22.5.$$

We chose to round down non-integer numbers to the nearest integer, so here:

$$n_j^t = 22.$$

3.3 Product ageing and breakdown modelling

We choose to set up a model in which the products will have an evolutive failure rate according to their ageing. Three age categories are defined for the products: young products (from 0 to 1 year old), medium-aged products (from 1 to 2 years old) and old products (more than 2 years old). This choice of categories is also defined regarding the 2 years rental duration.

The failure rate will be shared in two types of failures: the short ones and the long ones. Each failure type has its own component cost, repair duration and proportion of the failure rate. This proportion will also evolve with the age of the product (i.e. the older the product will go, the more long or serious failures it will have). Let t_{pm} be the type p failure rate of a product of age m , and β_{pm} the proportion of failure of type p for a product of type m . The number of failures of type p for a product of type m will be as follows:

$$nbp_{pm} = t_{pm} \times \beta_{pm}.$$

3.4 Cost structure

To define the cost estimation purpose of our work, we had to set up a cost modelling of the system. To do so, all the costs that the system could entail are cataloged and classified into categories. This allows to identify which activity would be the most expensive.

Table 1. Costs details

Cost	Notation	Detail
Structural facility cost	coe_i for each facility i	Opening cost of a facility (one-year rent and agency fee on the first year), calculated by the UFL-ILP
Structural installation cost	cti	Transport costs related to installation (one direct round trip per new product installed), calculated by the UFL-ILP
Structural maintenance cost	ctm	Transport costs related to maintenance (one direct round trip per failure), calculated by the UFL-ILP
Functional facility cost	cfi_i for each facility i	Depreciation costs of the technician's trucks, rental expenses, electricity fees
Functional installation cost	$cins$	Technician's wage during the installation
Functional maintenance cost	$cmal$	Component cost, technician's wage during the maintenance

Two types of costs were distinguished: structural costs, which are related to the design of the network and are calculated by the UFL-ILP, and functional costs, which are calculated from the network obtained by the solving of the UFL-ILP. Each one of these cost types are subdivided into three categories: facility cost, installation cost and maintenance cost. These costs are detailed in Table 1. Figure 3 summarizes schematically the inputs and outputs of the model.

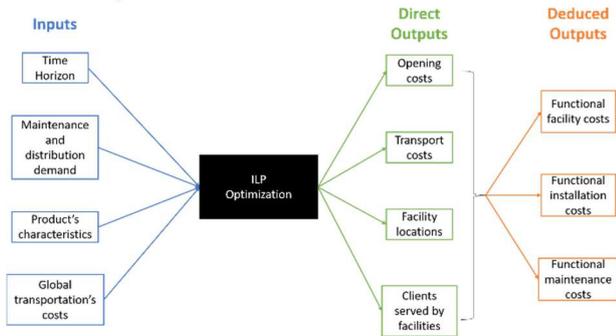


Figure 3. Diagram of the inputs and outputs of the model

4 INTEGER LINEAR PROGRAM

To model our logistics network design problem, we choose to formulate it as an Uncapacitated Facility Location based ILP (UFL-ILP) as follows.

4.1 Problem Data

T – Time horizon.

$T = \llbracket 1, T \rrbracket$.

$T_1 = \llbracket 2, T \rrbracket$.

I – Set of potential facilities.

J – Set of customer cities.

M – Set of product types.

P – Set of failure types.

t_{ins} – Installation time.

\mathcal{R} – Action range (i.e. maximal duration between a failure and its repair).

d – Annual availability of a technician in hours.

$\forall i \in I, co_i$ – Opening cost of the facility i ,

fa_i – Agency fee of the facility i .

$\forall i \in I, \forall j \in J, ct_{ij}$ – Transport cost from i to j ,

dt_{ij} – Transport duration from i to j .

du – Mean time to do a direct round trip from a facility to a customer (this is the sum of all transport durations on the number of possible trips).

$\forall t \in T, \forall j \in J, n_j^t$ – Number of products in the city j the year t .

$\forall t \in T, \forall j \in J, \forall m \in M, n_{jm}^t$ – Number of products of type m in the city j the year t , especially,

$\forall t \in T, \forall j \in J, \forall m \in M, n_{j0}^t$ – Number of new products in the city j the year t .

$\forall p \in P, ti_p$ – Repair time for a failure of type p .

$\forall p \in P, \forall m \in M, nbp_{pm}$ – Number of failures of type p for a product of type m .

4.2 Decision Variables

$\forall t \in T, \forall i \in I, ntech_i^t$ – Number of technicians to be hired the year t at the facility i .

$\forall t \in T, \forall i \in I,$

$$y_i^t = \begin{cases} 1 & \text{if the facility } i \text{ is opened the year } t, \\ 0 & \text{otherwise.} \end{cases}$$

$\forall t \in T, \forall i \in I, \forall j \in J,$

$$x_{ij}^t = \begin{cases} 1 & \text{if the facility } i \text{ serves the customer } j \text{ the year } t, \\ 0 & \text{otherwise.} \end{cases}$$

4.3 Objective function

$$z = \min \sum_{t \in T} coe^t + cti^t + ctm^t + \sum_{t \in T} \sum_{i \in I} w ntech_i^t$$

4.4 Constraints

$$coe^0 = \sum_{i \in I} (co_i + fa_i) y_i^0; \quad (1)$$

$$\forall t \in T_1, coe^t = \sum_{i \in I} co_i y_i^t + fa_i (y_i^t - y_i^{t-1}); \quad (2)$$

$$\forall t \in T, cti^t = \sum_{i \in I} \sum_{j \in J} 2n_{j0}^t ct_{ij} x_{ij}^t; \quad (3)$$

$$\forall t \in T, ctm^t = \sum_{i \in I} \sum_{j \in J} 2ct_{ij} x_{ij}^t \sum_{m \in M} n_{jm}^t \sum_{p \in P} nbp_{pm}; \quad (4)$$

$$\forall t \in T, \forall i \in I, n_{tech_i^t} \geq \frac{1}{d} \left[(t_{ins} + du) \sum_{j \in J} n_{j0}^t x_{ij}^t + \sum_{p \in P} (t_{ip} + du) \sum_{j \in J} \sum_{m \in M} n_{jm}^t n_{pjm} x_{ij}^t \right]; \quad (5)$$

$$\forall t \in T, \forall i \in I, \forall j \in J, x_{ij}^t \leq y_i^t; \quad (6)$$

$$\forall t \in T, \forall j \in J, \sum_{i \in I} x_{ij}^t = 1; \quad (7)$$

$$\forall t \in T, \forall i \in I, \forall j \in J, 2dt_{ij} x_{ij}^t \leq \mathcal{R}; \quad (8)$$

$$\forall t \in T_1, \forall i \in I, y_i^t \geq y_i^{t-1}. \quad (9)$$

Constraints (1), (2), (3) and (4) define the different costs (facility, installation, maintenance) of the problem while constraint (5) set the number of technicians necessary. Constraint (6) ensures that the customer cities are served only by opened facilities. The customer cities are served by one and only facility is ensured by constraint (7). Constraints (6) and (7) are the constraints coming from the classic UFL model. Constraint (8) allows to respect the action range, and finally constraint (9) ensures that an opened facility stays open the next year.

5 EXPERIMENTAL RESULTS

The problem has been modelled and solved using the data characterizing the company. The UFL-ILP have been run on CPLEX (12.6.2 Version), with an Intel Core i7-8550U CPU and 8Gbytes of usable RAM. The runs lasted less than one second for each data instance.

5.1 Data Generation

A time horizon of five years, an action range of 48 hours (quality of service), an installation time of half an hour are used. These parameters are deduced from the data provided by the partner company. The annual technician's availability of 1470 hours is used, calculated from the mean annual days of work of a technician worker. The charged wage per hour of the technicians is set to 19.45€, based on the mean wage of a technician. Three age categories have been defined for the products, and two failure types, again based on the data of the company (Tables 2 and 3).

Table 2. Failures' Characteristics

Failure type	Short	Long
Repair time	1h	2h
Component cost	32€	120€

Table 3. Products' Characteristics

Age category	Young	Medium	Old
Failure rate	0.33	0.38	0.45
Long failure proportion	0.8	0.75	0.7
Short failure proportion	0.2	0.25	0.3

For the data pertaining to the facilities, we choose to locate facilities only on the cities where products are located (cf. section 3.1), except for Luxembourg for which we could not find sufficient information. The number of client cities is 41, and the number of candidate cities for facilities is 40. For each candidate city to locate a facility, we search for an available suitable facility and use its annual rent as opening cost. If the

facility is opened for the first time, the real estate agency costs are also considered in the opening cost.

For the functioning cost of the facilities, the annual electricity consumption cost are calculated and added to the annual charges paid to the real estate agency.

5.2 Centralized Solution

The first step is to solve the problem using the strategy currently in process at the company. It consists in operating all the activities from the production site (centralized solution). In this part, there is only one possible facility, and the UFL-ILP will not decide where to locate the facilities. It will just calculate the costs and the number of technicians to hire.

Figure 4 shows a representation of the results per cost category and per year.

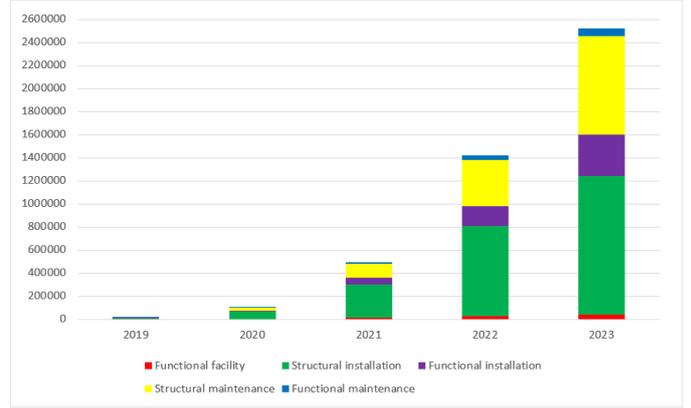


Figure 4. Centralized solution's cost repartition

In this solution, there is no structural facility cost, because no facility is opened, so there is no rent fee. We observe that the most expensive costs are the structural installation cost and the structural maintenance cost, which represent the transport costs of the system.

5.3 Decentralized solution

In this part, the program is run with the possibility of opening several facilities, to answer the problem first defined (*i.e.* design the logistics network).

Figure 5 shows a representation of the results per cost category and per year. Maps showing the network obtained each year are available in Appendix 3.

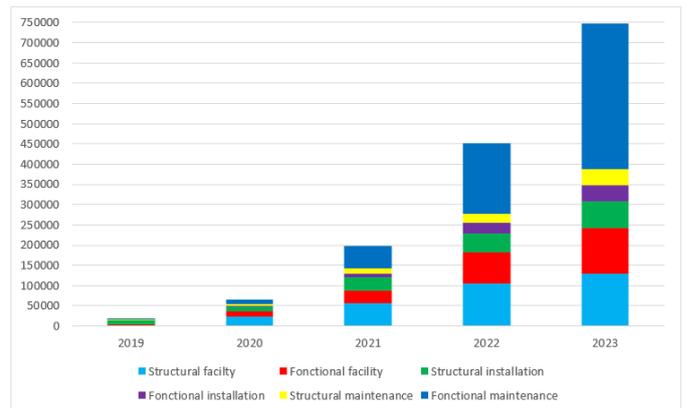


Figure 5. Decentralized solution's cost repartition

In this solution, the most expensive costs are the functional maintenance cost and the structural facility cost.



Figure 6. Comparison of the two solutions

5.4 Comparison of the Solutions: Centralized vs. Decentralized

After we ran the program for the two solutions and calculated all the costs, the results have been compared to identify which solution would be the most interesting for the company. Figure 6 shows the cost comparisons for each activity and for the total costs over 5 years of the business plan.

The centralized solution is more interesting in terms of facility cost. Indeed, only the Grenoble site is opened every year, but its opening cost is set to zero because the site already exists. Regarding this, the facility cost is reduced in the centralized solution. But this facility cost savings is low compared to the savings made on installation and maintenance costs in the decentralized solution. As the facilities are closer to the customers thanks to the decentralized logistics network, the transport costs are indeed much lower than in the centralized solution. Finally, the total costs (bottom right chart) of the decentralized system are lower (and even three times lower in 2023) when compared to those of the centralized one.

Table 4. Variations of the failure rate

Age category	Young	Medium	Old
Variation 1	0.25	0.30	0.37
Variation 2	0.30	0.35	0.42
Variation 3	0.35	0.40	0.47
Variation 4	0.40	0.45	0.52

The failure types proportions, β_1 and β_2 , are related by the formula $\beta_1 + \beta_2 = 1$. Table 5 details the values of β_1 used in the sensitivity analysis depending on the products' age category.

Table 5. Variations of the failure types' proportions

Age category	Young	Medium	Old
Variation 1	0.60	0.55	0.50
Variation 2	0.70	0.65	0.60
Variation 3	0.80	0.75	0.70
Variation 4	0.90	0.85	0.80

6 SENSITIVITY ANALYSIS

A sensitivity analysis has been conducted to identify which parameters could influence the results. We have chosen to vary four parameters: the action range, the products repartition in the cities, the failure rate, and the failure types proportion regarding the failure rate.

6.1 Parameters

The action range was initially 48 hours, which is a specification of the company's rental contracts. We choose to run the program with an action range of 24 hours, 20 hours and 10 hours.

The product repartition depends on two parameters, the population rate and the occupation rate. These parameters are respectively weighted by the values α_1 and α_2 . These values are initially set to $\alpha_1 = 0.1$ and $\alpha_2 = 1 - \alpha_1 = 0.9$. The program is run using $\alpha_1 = 0.2$ to $\alpha_1 = 0.9$ incremented of 0.1 each time and with the corresponding values for α_2 .

Table 4 details the values used for the failure rate depending on the products age category.

6.2 Results

The results of the sensitivity analysis are detailed in Table 6. A '0' means that none of the variations influences the result. A '+' and a '++' mean that respectively one and two variations influence the result. A '+++' means that at least three variations influence the result. Looking at the results, it is to note that the most influencing parameters are the products repartition and the failure rate. Each variation of these parameters has indeed a significant impact on most of the cost types.

About the action range, this parameter impacts the result only when it is set to 10 hours (otherworldly divided by nearly 5 from 48h). Thus, we can conclude that the solution is not very sensitive to this parameter, except if this constraint becomes very strong, that is to say in the hypothesis of a high quality service.

Finally, the failure types proportions impact only the functional maintenance cost. This can be explained by the fact that these parameters do not appear in the UFL-ILP, and hence have no influence on the structural costs.

Table 5. Variations of the failure types proportions

	Structural costs			Functional costs		
	Facilities	Installation	Maintenance	Facilities	Installation	Maintenance
Action range	+	+	+	+	0	+
Products repartition	+++	+++	+++	+++	++	++
Failure rate	+++	+++	+++	+++	0	+++
Failure types proportions	0	0	0	0	0	+++

7 OUTSOURCING AND RECOMMENDATIONS

For the start-up company, different configurations (centralized vs. decentralized, with outsourcing or not) of the logistics network are designed. We try to find which one is the most interesting for the company in terms of costs. The first step is the identification of the network configuration, the different service activities and the outsourcing options available for each one.

7.1 Logistics Network Structure

As stated before, the logistics network can be centralized or decentralized. To compare both options, we totalize of the structural and functional facility costs.

7.2 Installation Service

This concerns the installation of the products at the clients' premises. This includes the transport of the product from the company to the installation site, and the installation intervention. This can be outsourced or not. To compare both situations, an offer made by a logistics service provider is used. We classify the costs of the provider's offer into categories and calculate the invoice for the company. Then, we calculate what it would cost to the company if it were internalized, using the same cost classification.

7.3 Maintenance Service

This is about the products maintenance at the clients' premises (i.e. failure repair). This includes the intervention cost, the component cost, and the technician's transport cost. This activity is necessarily internalized because it is a core component of the company's service and product offer.

7.4 Transport of Products

The transport of products is subdivided in two components.

- The upstream transport includes the way from the company to the facility inventory. This subcomponent is considered only when we choose a decentralized network.
- The downstream transport includes the way from the facility inventory to the customer. In case of a centralized network, this is the way from the production site to the customer's premises.

If the installation service is outsourced, the service provider manages this part, and so we do not have to make a decision on this. Otherwise, we can choose to outsource it or not. To estimate the cost invoiced by a transport provider, the transport costs we calculated for the UFL-ILP are used, and are increased by 15% to consider the functioning fees and the margin of the carrier.

7.5 Recommendations

Six different scenarios are built using all the possibilities we enumerated and calculated the overall costs of each one. Based on this comparison, the least expensive configuration was the

one with a decentralized network, internalized installation service and outsourced transport. In this configuration, the company is in charge of the facility, installation and maintenance activities. The service provider will manage the upstream and downstream transports. The cost of this configuration is close to the one of the configurations in which we have a decentralized network, with outsourced installation.

As the outsourced transport costs are an estimation, the study could be more precise if we considered a real carrier's offer, so that the choice between those two scenarios could be more reliable.

8 CONCLUSION

The aim of our work was to design a logistics network dedicated to a growing company which uses a PSS business model.

To reach this goal, a complete methodology have been built based on a data model representing the system, the product life cycle and the expansion strategy of the company over its 5 years business plan. This data model has been used as an input for our adaption of an Uncapacitated Facility Location based ILP which allowed to design optimized logistics networks and estimate the overall system costs.

We completed the study by a report on the different strategies that the company could adopt. The comparison of these strategies leads to recommendations for the company regarding outsourcing possibilities.

The goals of this study have been globally answered, but the work we carried out can be extended.

First, we built a linear life cycle model for the products. As the product of the company are still young, there is a lack of data about the way they will work with the time. A more complete life cycle analysis on the products could help to understand better how they evolve, and then offer a more precise model in terms of maintenance activities.

Then, with the upcoming expansion of the company, more data will be available. This can allow to consider other aspects of the problem which are currently overlooked, such as vehicle routing for simultaneous installations or maintenances at the clients' premises or spare parts and consumable refurbishment management at the facilities. We could use more sophisticated models as Location-Inventory-Routing Problems to address these aspects altogether.

Furthermore, as stated before, the outsourced transport costs are estimated, this can be further refined with real offers to better specify our recommendations.

9 REFERENCES

Reim, W., Parida, V., & Örtqvist, D. (2015). Product-Service Systems (PSS) business models and tactics - A systematic literature review. *Journal of Cleaner Production*, 97, 61–75.

Sun, J., Qu, T., Nie, D., & Li, P. (2019). Research on “location-inventory” problem of spare parts supply chain based on product service system. *Procedia CIRP*, 83, 819–825.

Shen, J., Erkoyuncu, J. A., Roy, R., & Wu, B. (2017). A framework for cost evaluation in product service system configuration. *International Journal of Production Research*, 55(20), 6120–6144.

Derrouiche, R., & Dubruc, N. (2014). Approche d’analyse et de simulation des différents scénarii système produit-service (PSS) : cas d’Envie Loire. *10ème Conférence Francophone de Modélisation, Optimisation et Simulation-MOSIM’14*, <https://hal.inria.fr/emse-01085281/>.

Kuo, T. C. (2011). Simulation of purchase or rental decision-making based on product service system. *The International Journal of Advanced Manufacturing Technology*, 52, 1239–1249.

Kuo, T. C., & Wang, M. L. (2012). The optimisation of maintenance service levels to support the product service system. *International Journal of Production Research*, 50(23), 6691–6708.

Rezapour, S., Allen, J. K., & Mistree, F. (2016). Reliable product-service supply chains for repairable products. *Transportation Research Part E: Logistics and Transportation Review*, 95, 299–321.

Xia, T., Xi, L., Pan, E., Fang, X., & Gebraeel, N. (2017). Lease-Oriented Opportunistic Maintenance for Multi-Unit Leased Systems under Product-Service Paradigm. *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, 139(7), 1–10.

van Ooijen, H., Bertrand, J. W. M., & Buyukkaramikli, N. C. (2019). Coordinating failed goods collecting and repair capacity policies in the maintenance of commoditized capital goods. *International Journal of Production Economics*, 208, 29–42.

Rusdiansyah, A., & Tsao, D. B. (2005). An integrated model of the periodic delivery problems for vending-machine supply chains. *Journal of Food Engineering*, 70(3), 421–434.

Kim, S., & Yoo, S. (2012). The combined manpower planning and preventive maintenance strategies by aggregate planning for the service division of vending machines. *Journal of the Operational Research Society*, 63(8), 1034–1050.

Sagarna, I., Uribetxebarria, J., Castellano, E., & Erguido, A. (2016). After-sales maintenance service strategies optimization. An offshore wind farm case study. *IFAC-PapersOnLine*, 49(28), 156–161.

Chen, X., Sheng, T., & Yi, S. (2009). Probe on network-based collaborative maintenance model for after-sales equipment. *Chinese Journal of Mechanical Engineering (English Edition)*, 22(5), 702–709.

Srivastava, S. K. (2008). Network design for reverse logistics. *Omega*, 36(4), 535–548.

Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2009). Facility location and supply chain management - A review. *European Journal of Operational Research*, 196(2), 401–412.

Listes, O., & Dekker, R. (2001). Stochastic approaches for

product recovery network design : a case study. *Econometric Institute Report EI 2001-08*, 1–16.

Cruz-Rivera, R., & Ertel, J. (2009). Reverse logistics network design for the collection of End-of-Life Vehicles in Mexico. *European Journal of Operational Research*, 196(3), 930–939.

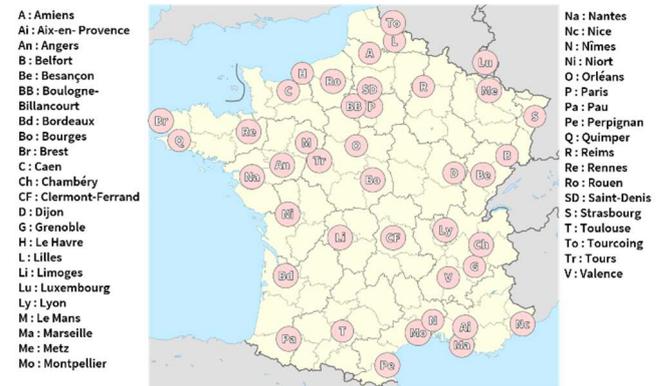
Kara, S., Rugrungruang, F., & Kaebnick, H. (2007). Simulation modelling of reverse logistics networks. *International Journal of Production Economics*, 106(1), 61–69.

Fleischmann, M. (2001). Reverse Logistics Network Structures and Design. *Business Perspectives on Closed-Loop Supply Chains*, 1–21.

Amini, M. M., Retzlaff-Roberts, D., & Bienstock, C. C. (2005). Designing a reverse logistics operation for short cycle time repair services. *International Journal of Production Economics*, 96(3), 367–380.

10 APPENDICES

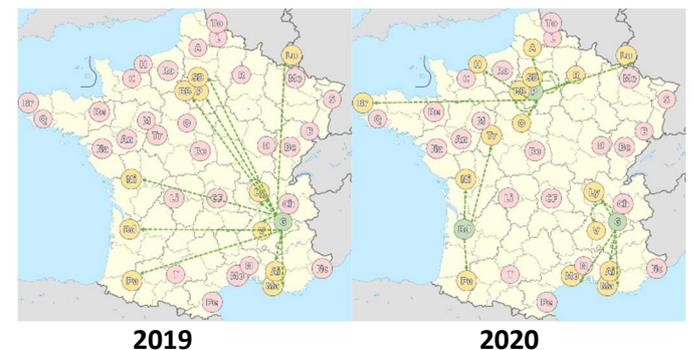
10.1 Appendix 1 – Final Geographical Scenario

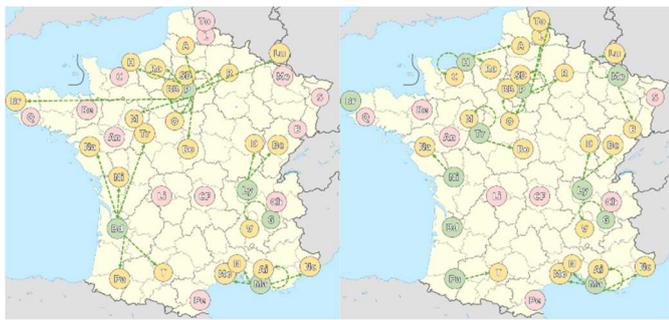


10.2 Appendix 2 – Roster of Cities per Activation Year

	Cities
2019	Lyon, Grenoble, Valence, Paris, Boulogne-Billancourt, Saint-Denis, Bordeaux, Pau, Niort, Marseille, Aix-en-Provence, Luxembourg
2020	Brest, Tours, Orléans, Reims, Amiens, Le Havre, Montpellier
2021	Dijon, Besançon, Bourges, Rouen, Toulouse, Nîmes, Nantes, Le Mans, Nice
2022	Belfort, Metz, Caen, Lille, Tourcoing
2023	Rennes, Quimper, Strasbourg, Angers, Limoges, Clermont-Ferrand, Chambéry, Perpignan

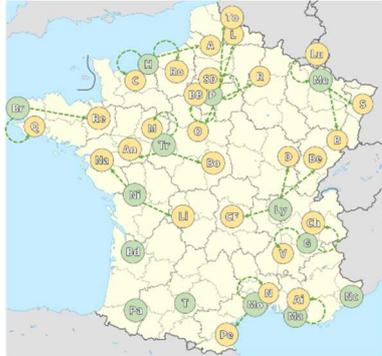
10.3 Appendix 3 – Maps of the Logistics Networks over Five Years





2021

2022



2023

Légende des icônes :

- A – Amiens
- Ai – Aix-en-Provence
- An – Angers
- B – Belfort
- Be – Besançon
- BB – Boulogne-Billancourt
- Bd – Bordeaux
- Bo – Bourges
- Br – Brest
- C – Caen
- Ch – Chambéry
- CF – Clermont-Ferrand
- D – Dijon
- G – Grenoble
- H – Le Havre
- L – Lille
- Li – Limoges
- Lu – Luxembourg
- Ly – Lyon
- M – Le Mans
- Ma – Marseille
- Me – Metz
- Mo – Montpellier
- Na – Nantes
- Nc – Nice
- N – Nîmes
- Ni – Niort
- O – Orléans
- P – Paris
- Pa – Pau
- Pe – Perpignan
- Q – Quimper
- R – Reims
- Re – Rennes
- Ro – Rouen
- SD – Saint-Denis
- S – Strasbourg
- T – Toulouse
- To – Tourcoing
- Tr – Tours

Étude sur le potentiel de mutualisation logistique dans la grande région de Montréal

SARA TARIQ¹, MAHA BEN ALI^{1,2}, MARTIN TRÉPANIÉ^{1,2}

¹ Département de mathématiques et de génie industriel, Polytechnique Montréal
2500 Chemin de Polytechnique, Montréal, Canada

² Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport (CIRRELT)
sara.tariq@polymtl.ca, maha.ben-ali@polymtl.ca, martin.trepanier@polymtl.ca

Résumé – La collaboration est considérée comme une solution pour assurer la compétitivité des différentes entreprises. Cet article s'intéresse principalement à la mutualisation logistique, une pratique efficace de la logistique qui a attiré une attention accrue ces dernières années. La sélection des partenaires fait partie des clés de réussite d'un réseau mutualisé et se base sur différents critères, tels que la répartition spatiale et le secteur d'activités. Cette étude vise à évaluer le potentiel de réalisation d'une mutualisation logistique dans la région de Montréal. D'abord, une revue de littérature est présentée, suivie d'une présentation de la base de données utilisée et de la méthodologie d'analyse. Ensuite, une analyse préliminaire sur les coalitions possibles dans la grande région de Montréal est exposée, ce qui pourrait servir de guide pour toute entreprise désirant débiter un projet de mutualisation des sites logistiques dans cette région.

Abstract – Collaboration is considered as a solution to ensure the competitiveness of different companies. This article focuses on logistics pooling, an effective logistics practice that has gained increased attention during the last few years. Partner selection is one of the keys to the success of a pooled network and is based on various criteria, such as spatial distribution and business sector. This study aims to assess the potential for achieving a logistic pooling in Montreal. First, a literature review is presented, followed by the presentation of the database used and the analysis methodology. Then, a preliminary analysis on the possible coalitions in Montreal is exposed, that would serve as a guide for any company wishing to initiate a logistic pooling in this region.

Mots clés – Mutualisation logistique, Sélection des partenaires, Répartition spatiale, Système de classification des industries de l'Amérique du Nord.

Keywords – Logistics Pooling, Partner selection, Spatial distribution, North American Industry Classification System.

1 INTRODUCTION

De nos jours, les entreprises doivent satisfaire les besoins de clients de plus en plus exigeants en termes de réactivité, de taux de service et de flexibilité. Or, assurer un niveau de service acceptable implique une augmentation importante des coûts logistiques : des livraisons fréquentes à faibles quantités et une dispersion géographique des sites logistiques et des points de livraison (Ouhader et al. 2016). Afin de demeurer compétitives, les entreprises ont besoin plus que jamais d'optimiser leurs coûts et de repenser leur organisation à travers l'amélioration de la chaîne logistique (Moutaoukil et al. 2013).

La mutualisation logistique figure parmi les solutions qui permettent de réduire les coûts et d'assurer une meilleure gestion de la chaîne logistique (Lozano et al. 2013). Il s'agit d'une forme de collaboration horizontale qui se fait entre des acteurs du même niveau (par exemple, entre un ensemble de fournisseurs ou un ensemble de producteurs). L'idée est de mettre en commun des moyens physiques (comme les entrepôts et les moyens de transport), des données nécessaires à la gestion, de risques et de bénéfices dans le but d'améliorer la performance du réseau logistique et du niveau de service (Makaci et al. 2014). Le transport et l'entreposage sont les deux formes de mutualisation logistique les plus répandues. La mutualisation du transport permet d'optimiser et d'améliorer le taux de remplissage du matériel roulant. La mutualisation des entrepôts consiste à considérer un entrepôt partagé entre plusieurs acteurs permettant de partager les espaces physiques,

les ressources et les informations logistiques (Makaci et al. 2017). De plus, elle permet d'assurer une réduction des émissions de gaz à effet de serre (Senkel et al. 2013) et de ressentir un impact social par la réduction du taux de risque, du nombre de voyages et des distances parcourues (Zouari 2019).

L'objectif de cet article est d'identifier des coalitions de mutualisation logistique les plus prometteuses dans la grande région de Montréal (c-à-d, l'île de Montréal, Laval, les Laurentides, Lanaudière et Montérégie). Un état de l'art est d'abord présenté. Ensuite, la méthodologie suivie et les résultats de l'analyse sont exposés.

2 MUTUALISATION LOGISTIQUE : ETAT DE L'ART

Dans cette section, une revue de littérature sur la mutualisation logistique est menée afin de clarifier cette notion et de présenter les clés de réussite d'un réseau mutualisé.

2.1 Sélection des partenaires

La première étape pour l'implémentation d'une mutualisation logistique est de sélectionner les bons partenaires (Amer et al. 2014). La Figure 1 résume les différents critères de sélection des partenaires cités dans plusieurs études (Glaister et al. 1997, Asawasakulsorn 2009, Bahinipati et al. 2009, Mat et al. 2009, Naesens et al. 2009, Anand et al. 2012, Audy et al. 2012, Winkelhaus et al. 2014, Pomponi et al. 2015, Martin et al. 2018, Sitadewi et al. 2018).

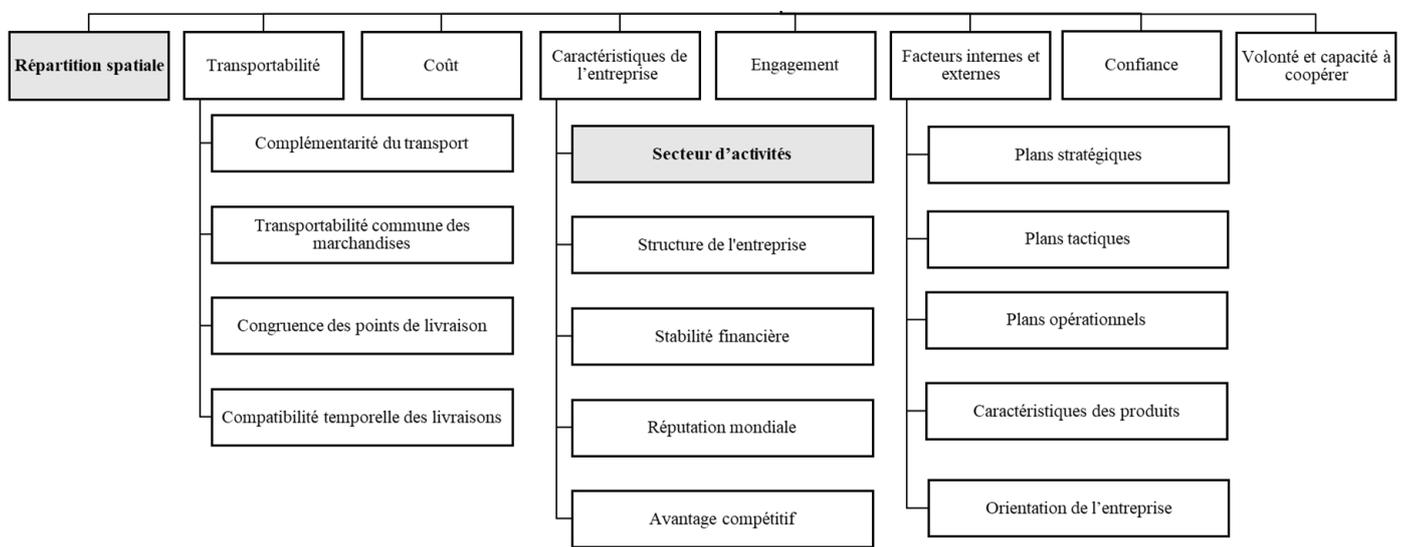


Figure 1 Critères de sélection des partenaires cités dans la littérature

Il existe des méthodes bien structurées pour évaluer ces différents critères. Winkelhaus et al. (2014) ont développé un modèle de choix des partenaires qui réalise une sélection pondérée des acteurs potentiels tout en évitant l'analyse détaillée des partenaires incompatibles. Naesens et al. (2009) ont utilisé la méthode AHP (*Analytic Hierarchy Process*) pour vérifier la correspondance stratégique des partenaires pour la mutualisation des entrepôts. L'étude de Bahinipati et al. (2009) a considéré l'aspect d'incertitude dans les jugements apportés par les décideurs. Les différents critères ont été évalués en utilisant la méthode AHP-FLM (*Analytic hierarchy process-fuzzy logic model*). Anand et al. (2012) ont utilisé une approche basée sur la théorie des graphes (*Graph-Theoretic Approach*), qui détermine la pondération du jugement subjectif apporté par les décideurs.

Des études plus récentes ont utilisé la répartition spatiale des sites logistiques des entreprises pour sélectionner les partenaires. L'idée est d'identifier des *clusters* qui peuvent réaliser une mutualisation logistique. Donahue et al. (2018) définissent les *clusters* par des groupes d'entreprises qui acquièrent un avantage concurrentiel grâce à leur proximité offrant, ainsi un cadre convaincant aux dirigeants pour analyser et soutenir leurs économies. Sakai et al. (2020) ont utilisé les données d'emplacement pour le développement des sites logistiques. Les entreprises ayant des sites dans la même zone géographique peuvent donc former un *cluster* et ont tendance à avoir des transactions entre elles (Kadokawa 2013). Les entreprises du même *cluster* peuvent facilement se réunir pour collaborer et développer des propositions de valeurs communes. Ces valeurs conjointes sont basées par exemple sur une réduction des coûts de transport, résultant de la proximité. La localisation commune permet également de faciliter le développement de la confiance mutuelle entre les entreprises (Sheffi et al. 2019). Par conséquent, la localisation géographique aide à créer des *clusters* qui ont un potentiel de mettre en place la mutualisation logistique. Cette technique sera utilisée par la suite pour analyser le potentiel de mise en place de la mutualisation dans la grande région de Montréal.

2.2 Partage des coûts et des bénéfices

La mutualisation logistique permet de générer des bénéfices supplémentaires grâce à la réduction des coûts. Audy et al. (2008) proposent de partager les coûts entre l'ensemble des membres de la coalition au lieu de répartir les économies

réalisées. Le défi est d'assurer à l'ensemble des partenaires que leur intégration à la coalition générera des coûts inférieurs à ceux générés en travaillant d'une façon autonome. L'approche la plus simple pour la répartition des coûts est basée sur le volume ou le poids global des produits transportés, tandis que les approches les plus avancées consistent à utiliser la théorie des jeux coopératifs (Guajardo et al. 2016). Selon cette théorie, une méthode d'allocation des coûts doit satisfaire certaines propriétés : l'efficacité, la rationalité individuelle, la symétrie, l'additivité, la propriété *dummy* et la stabilité (Frisk et al. 2010). Il est possible aussi d'utiliser la méthode d'allocation proportionnelle qui consiste à répartir équitablement entre les partenaires les profits de la grande alliance, en la pondérant par le coût autonome de chaque acteur (Frisk et al. 2010, Liu et al. 2010). Bien que cette méthode soit facile à comprendre et à calculer, elle n'est pas stable car elle peut fournir des allocations qui ne sont pas équitables.

La méthode de la valeur de Shapley, développée en 1953, est l'une des méthodes de répartition des coûts les plus fréquemment utilisées (Guajardo et al. 2016). Cela est dû principalement à l'unicité de la solution retenue, sa facilité de mise en œuvre et son équité (Krajewska et al. 2008) : chaque membre devrait se voir attribuer un coût égal à sa contribution moyenne dans la coalition (Wong et al. 2007). Au fur et à mesure que chaque participant entre dans la coalition, on lui attribue le coût marginal, correspondant à l'augmentation du coût total de la coalition due à son entrée (Frisk et al. 2010). Quant à la méthode de Nucleolus datant de 1969, elle permet une répartition des coûts qui minimise l'excédent maximal, c'est-à-dire la différence entre la somme des profits alloués aux partenaires et le profit total de la coalition (Liu et al. 2010). Cette méthode permet une répartition des coûts qui minimise la pire iniquité, de sorte que la rationalité individuelle soit satisfaite (Frisk et al. 2010). Les méthodes de la valeur de Shapley et de Nucleolus partagent une même propriété qui est l'unicité. La méthode de la valeur de Shapley est facile à mettre en œuvre, alors que la méthode de Nucleolus nécessite de résoudre une séquence de programmes linéaires (Krajewska et al. 2008).

Il existe aussi des méthodes d'allocation répartissant le coût total en un coût séparable et un coût non séparable (Frisk et al. 2010), parmi lesquelles on cite la méthode ECM (*Equal Charge Method*), la méthode ACAM (*Alternative Cost Avoided Method*) et la méthode CGM (*Cost Gap Method*), la méthode

EPM (*Equal Profit Method*). La méthode EPM, qui s'assure que les économies relatives soient aussi similaires que possible pour tous les participants, a été utilisée par exemple dans le contexte d'une étude de cas dans le domaine forestier (Guajardo et al. 2016).

L'étude d'Audy et al. (2011) a proposé une version modifiée de l'allocation des coûts séparables et non séparables qui tient compte des modifications de la planification collaborative (par exemple, pour répondre à des exigences spécifiques). Les mêmes auteurs ont considéré dans une étude plus récente (Audy et al. 2012) que l'aspect comportemental des entreprises peut également influencer la répartition des coûts et des bénéfices entre les partenaires de la coalition. Dans ce sens, ils répartissent les entreprises qui forment la coalition en deux catégories : les entreprises leaders qui initient et conduisent la planification de la collaboration et les entreprises non-leaders qui supportent le processus. Les leaders détenant une attitude opportuniste visent à obtenir un avantage supplémentaire aux dépens des entreprises non-leaders, ce qui influence la répartition des coûts et des bénéfices. En présence d'un comportement opportuniste, l'augmentation marginale réalisée à l'entrée de nouveaux acteurs n'est partagée qu'entre les partenaires qui viennent d'intégrer la coalition et les sites logistiques faisant partie des entreprises leaders. Dans le cas contraire, l'augmentation marginale est partagée entre tous les acteurs de la coalition.

2.3 Conception d'un réseau mutualisé

Il s'agit d'un problème de décision stratégique qui consiste à localiser un certain nombre d'installations logistiques (Ouhader et al. 2016). Différents modèles de programmation linéaire ont été proposés dans la littérature, permettant d'assurer une réduction des coûts logistiques et/ou des émissions de gaz à effet de serre. Des études telles que Daoud et al. (2015), Ouhader et al. (2016), Verdonck et al. (2016) et Ouhader et al. (2017), ont utilisé des modèles de programmation linéaire en nombres entiers mixtes (MILP), tandis que d'autres comme Pan (2010) et Pan et al. (2011), ont eu recours à des modèles de programmation linéaire en nombres entiers (PLNE).

2.4 Évaluation de l'impact de la mutualisation

L'implémentation de la mutualisation logistique a des impacts économiques, environnementaux et sociaux. Différents coûts peuvent être pris en compte pour évaluer l'impact économique, tels que les coûts en relation avec le transport, le stockage, les retards de livraison, l'ouverture des entrepôts mutualisés, la préparation des commandes (Mrabti et al. 2018), ainsi que les coûts d'immobilisation du véhicule lors d'un arrêt de chargement ou de déchargement dans une des installations du réseau logistique (Pan et al. 2014). L'impact environnemental de la mutualisation est mesuré principalement par les émissions de CO₂ (Pan et al. 2013, Mrabti et al. 2019). Pour l'impact social, Mrabti et al. (2019) estiment qu'il est important d'évaluer l'impact de la mutualisation sur les habitants et les citoyens, et particulièrement le niveau de bruit des véhicules et le niveau de sécurité des personnes sur les routes. Neubert et al. (2014) ont considéré en plus les conséquences physiques des émissions polluantes sur la santé publique. Morana et al. (2015) proposent de tenir compte également des effets du système de livraison sur les aspects organisationnel (habitudes, changements sur les modes d'organisation, réticences), réglementaire et ergonomique. Les deux indicateurs les plus utilisés dans la littérature pour évaluer l'impact social de la mutualisation logistique sont le niveau de bruit des véhicules (Mrabti et al. 2020) et le taux d'accident reflétant le niveau de sécurité des

personnes sur les routes (Moutaoukil et al. 2013, Mrabti et al. 2019).

2.5 Partage des informations

Le partage des informations est défini par Badraoui et al. (2020) comme l'acte de partager en temps opportun des informations pertinentes, exactes et complètes. En particulier, il a été démontré que le partage d'informations entre les partenaires logistiques permet d'améliorer l'ensemble des activités (Audy et al. 2012), en plus de créer un environnement de confiance (Badraoui et al. 2020).

Selon Gonzalez-Feliu et al. (2011), la mutualisation nécessite un système d'information efficace pour assurer un haut niveau de performance. Cruijssen et al. (2007) subdivisent les informations en deux catégories : exclusives et partagées. Tel que présenté par la Figure 2, un système d'information d'un réseau logistique mutualisé se base principalement sur 3 éléments (Gonzalez-Feliu et al. 2011, Morana et al. 2014) : le module de gestion logistique, les technologies de l'information et de la communication (réparties en des outils internes tels que les logiciels professionnels et des outils externes tels que les codes à barres ou l'identification par radiofréquence), ainsi que la liste des acteurs. Ces derniers sont impliqués dans l'identification des objectifs de la mutualisation et des risques potentiels, et dans le suivi de la performance du réseau mutualisé grâce à des indicateurs en relation avec les trois aspects économiques, environnemental et social.

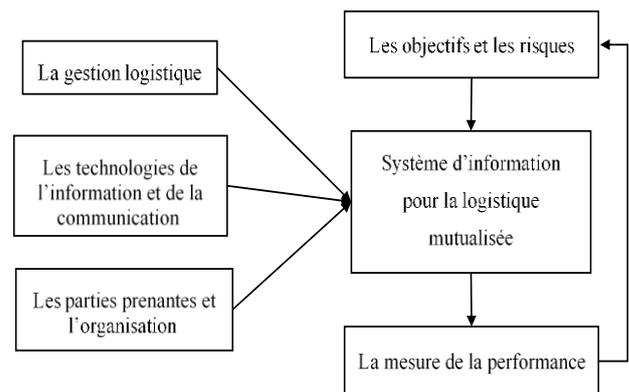


Figure 2 Éléments de base d'un système d'information pour la logistique mutualisée. Inspiré de Morana et al. (2014)

Le « Guide pratique de la mutualisation logistique » de l'Institut du commerce (2013) a proposé d'avoir recours aux solutions techniques suivantes : les portails collaboratifs (soit les plateformes interactives de stockage et de consultation des données), les systèmes de transport intelligents ITS (*Intelligent Transport System*), les logiciels de gestion de transport TMS (*Transport Management System*) et les logiciels de gestion mutualisée des approvisionnements. En plus de ces outils, l'Internet des objets (IoT) permet d'assurer la disponibilité des informations en temps réel, ce qui représente un des éléments essentiels pour la réussite de la mutualisation logistique (Reaidy et al. 2015). En effet, l'IoT soutient l'efficacité opérationnelle en fournissant des informations sur l'utilisation des actifs et en surveillant les équipements à distance comme les camions de transport, leurs emplacements et leurs vitesses. Reaidy et al. (2015) ont proposé une infrastructure IoT basée sur la technologie RFID (*Radio Frequency Identification*), pour assurer la gestion des entrepôts collaboratifs. La RFID permet, grâce à des étiquettes collées sur les produits, de mémoriser et de récupérer les données à distance.

Différentes études avaient proposé le prestataire logistique de 4^e génération (4PL) comme solution pour un partage efficace des informations dans le cadre de la mutualisation logistique. Le 4PL, un prestataire qui coordonne entre les différents acteurs en utilisant ses propres ressources, capacités et technologies, permet d'assurer la neutralité et l'efficacité des opérations logistiques (Rodrigues et al. 2015) et de réduire les coûts d'investissement dans l'implantation de nouvelles technologies d'information (Dircksen et al. 2017).

2.6 Transport collaboratif

Le transport collaboratif est la principale forme de la mutualisation logistique qui se retrouve dans la littérature. La hausse des prix du carburant et la concurrence féroce sur le marché poussent les entreprises à améliorer leur planification des transports et à assurer une utilisation optimale des capacités de transport disponibles (Danloup et al. 2013). Le transport collaboratif permet non seulement d'améliorer l'efficacité, mais également de réduire les émissions des gaz à effet de serre, la congestion routière et la pollution sonore (Gansterer et al. 2018).

Différentes études ont présenté des modèles permettant une affectation optimale des demandes de transport aux transporteurs et une planification réussie dans un contexte de mutualisation (Berger et al. 2010, Wang et al. 2014, Gansterer et al. 2018). Des solutions techniques sont présentées dans la littérature, telles qu'un système de planification collaborative du transport basé sur un système d'information géographique basé sur le web (Lu 2004), et un simulateur à événements discrets (Makhloufi et al. 2015).

Cet état de l'art a mis l'accent sur les principaux thèmes en relation avec la mutualisation logistique abordés dans la littérature et présente les différents critères de sélection des partenaires (voir Figure 1). Parmi ces critères, cet article propose d'analyser les critères de répartition spatiale et de secteurs d'activités pour l'évaluation du potentiel de mise en place de la mutualisation entre les sites logistiques. Dans ce qui suit, nous présenterons la méthodologie suivie pour une étude de cas dans la grande région de Montréal.

3 METHODOLOGIE

Comme mentionné dans la section 2.1, les sites logistiques se trouvant dans une même zone géographique peuvent communiquer et instaurer un climat de confiance plus facilement (Sheffi et al. 2019). Ceci est important pour réussir une mutualisation logistique, d'où l'intérêt de former des *clusters* sur la base de la répartition spatiale des sites. Nous proposons d'analyser les données des sites logistiques qui se trouvent dans la grande région de Montréal et de former les *clusters* en se basant sur leurs localisations géographiques. Nous analysons une base de données qui inclut les coordonnées (x,y) en projection MTM des sites se trouvant dans la grande région de Montréal et les classifie par secteur d'activité. Chaque secteur est identifié par un code NAICS (*North American Industry Classification System*).

La Figure 3 représente l'ensemble des étapes suivies pour identifier les coalitions les plus prometteuses pour réaliser une mutualisation logistique dans la grande région de Montréal. D'abord, nous sélectionnons les secteurs (ou sous-secteurs) d'activités les plus pertinents à considérer dans cette étude. Les données proviennent de CanMap Content Suite, publiées par la compagnie DMTI (Desktop Mapping Technologies Inc) Spatial (DMTI Spatial 2020). Ensuite, nous réalisons une analyse descriptive des données grâce aux tableaux croisés dynamiques d'Excel. Finalement, nous menons une analyse géographique

des données avec le logiciel QGIS (*Quantum Geographical Information System*).

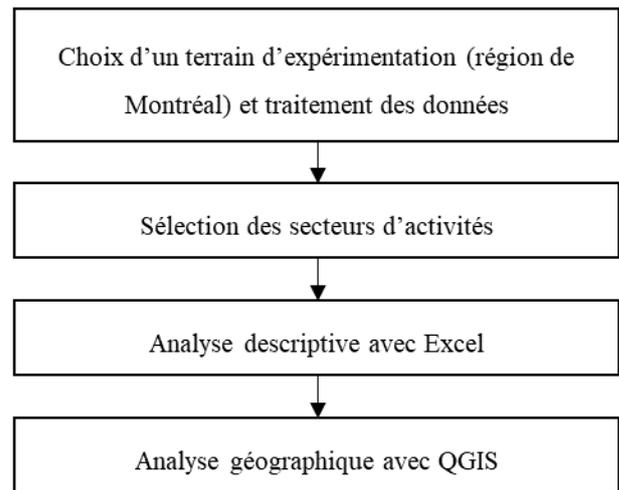


Figure 3 Étapes suivies pour l'analyse des données des sites logistiques présents dans la grande région de Montréal

Le logiciel QGIS considère les localisations géographiques des différents sites qui se trouvent dans la grande région de Montréal et permet d'identifier ceux qui peuvent faire partie du même réseau mutualisé (soit ceux qui sont à proximité). En premier lieu, nous avons calculé le centre moyen et l'écart-type au centre par secteur d'activités. Les coordonnées du centre moyen sont respectivement la moyenne des x et la moyenne des y des coordonnées des sites logistiques faisant partie du même secteur. L'écart type, quant à lui, reflète la dispersion géographique des sites du secteur en question sur la grande région de Montréal autour du centre moyen. En second lieu, nous avons utilisé un outil pour déterminer les *hot spots* afin d'identifier les coalitions possibles. Selon CrimeState (2013), un *hot spot* représente une concentration d'activité dans une zone géographique limitée. Dans notre cas, la notion de *hot spot* représente un *cluster* formé par une concentration de sites qui exercent des activités similaires. Parmi les différentes techniques qui existent pour réaliser les *hot spots*, nous avons utilisé une combinaison entre : i) le mode, reflétant la localisation avec le plus grand nombre de sites logistiques, et ii) le mode *Fuzzy* permettant de définir un petit rayon de recherche autour de chaque emplacement pour inclure les événements qui se produisent autour ou à proximité. En d'autres termes, il s'agit de regrouper les sites logistiques qui sont localisés dans un rayon bien défini. Dans cette étude, les *clusters* sont formés avec un minimum de 10 sites se trouvant dans un rayon de 2 kilomètres entre elles.

Dans ce qui suit, les résultats obtenus sont analysés pour identifier le potentiel de mise en place de réseaux mutualisés dans la grande région de Montréal.

4 RESULTATS ET DISCUSSION

Dans cette étude, nous nous sommes concentrés sur des secteurs liés à la fabrication et avons évalué via trois étapes le potentiel de mise en place de la mutualisation logistique dans la grande région de Montréal.

Dans la première étape, l'analyse descriptive des données grâce aux tableaux croisés dynamiques d'Excel nous a permis d'identifier les cinq sous-secteurs les plus importants en termes de nombre de sites logistiques et de somme d'emplois (voir le Tableau 1).

Tableau 1 : Les cinq sous-secteurs de fabrication les plus importants en termes de nombre de sites logistiques et de somme d'emplois à la grande région de Montréal

Sous-secteur (code NAICS)	Nombre de sites	Somme d'emplois
Fabrication du matériel informatique (334)	1209	16306
Fabrication des machines (333)	1434	17522
Fabrication des produits métalliques (332)	977	13061
Fabrication des aliments (311)	653	6642
Fabrication des vêtements (315)	323	3471

Dans la deuxième étape, grâce au logiciel QGIS, nous avons visualisé la répartition spatiale des sites logistiques de chaque sous-secteur sur la grande région de Montréal avec les différents *clusters* possibles qui peuvent mettre en place une mutualisation logistique. La Figure 4 illustre, comme exemple, les *hot spots* du sous-secteur de fabrication des aliments (311). Le numéro affiché pour chaque groupe sur la figure représente le nombre de sites associés au *cluster*. Un réseau mutualisé pourrait être mis en place au niveau du *cluster* qui se trouve sur l'île de Montréal et qui regroupe 63 sites logistiques. De même, à Laval, la mutualisation des ressources pourrait être intéressante au niveau du *cluster* formé de 16 sites. La même analyse a été reproduite pour les quatre autres sous-secteurs d'activité, soit la fabrication du matériel informatique (334), la fabrication des machines (333), la fabrication des produits métalliques (332), et la fabrication des vêtements (315).

Comme troisième étape, nous avons mené une analyse comparative des cinq sous-secteurs sélectionnés, en utilisant le centre moyen et l'écart type.

4.1 Centre moyen

La Figure 5 montre la répartition géographique des sites logistiques et des centres moyens associés à chaque sous-secteur. Les coordonnées (\bar{x}, \bar{y}) du centre moyen d'un sous-secteur correspondent aux moyennes des coordonnées (x_i, y_i) des N sites logistiques associés au sous-secteur, soit :

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (1) \quad \text{et} \quad \bar{y} = \frac{\sum_{i=1}^N y_i}{N} \quad (2)$$

Tel qu'illustré par la figure 5, les centres moyens des cinq sous-secteurs étudiés se trouvent sur l'île de Montréal. Notons que les centres moyens les plus rapprochés sont ceux du sous-secteur de fabrication des vêtements (315), de fabrication des produits informatiques (334) et de fabrication des aliments (311).

4.2 Zone rectangulaire de l'écart type

La Figure 6 illustre l'écart type pour chaque sous-secteur, ce qui reflète la concentration des sites logistiques autour du centre moyen. Elle démontre que le sous-secteur de fabrication des machines (333) est le sous-secteur le plus étendu. Une concentration moyenne est observée avec le sous-secteur fabrication des produits métalliques (332), le sous-secteur de fabrication des aliments (311) et le sous-secteur de fabrication des produits informatiques (334). En dernière position, le plus petit écart type est celui du sous-secteur fabrication des vêtements (315), dont les sites logistiques sont plus concentrés autour du centre moyen.

4.3 Discussion

En raison de leurs proximités relatives, les sites logistiques appartenant aux différents *clusters* pourraient débiter une mutualisation logistique. Cependant, il faut s'assurer d'abord d'avoir les éléments de base nécessaires pour la réussite de ce projet, tels que la confiance et l'engagement. Par exemple, au niveau du cluster formé de 63 sites dans le sous-secteur de fabrication des aliments (Figure 4), chaque entreprise doit réaliser une analyse individuelle pour s'assurer de collaborer avec les bons partenaires. Après la formation de la coalition et avant de commencer la mise en œuvre pratique de la mutualisation, il faut répartir les coûts entre les différents partenaires, en utilisant la méthode de la valeur de Shapley par exemple (la plus citée dans la littérature). Ensuite, les différents acteurs doivent décider des systèmes d'information à utiliser pour assurer un partage efficace des informations. Le recours à l'IoT pourrait être une bonne solution pour assurer la disponibilité des informations en temps réel et une prise de décision plus efficace. L'identification des meilleures localisations des sites mutualisés (soit celles générant le minimum de coûts et des émissions de CO₂) pourrait se faire par un modèle de programmation linéaire.

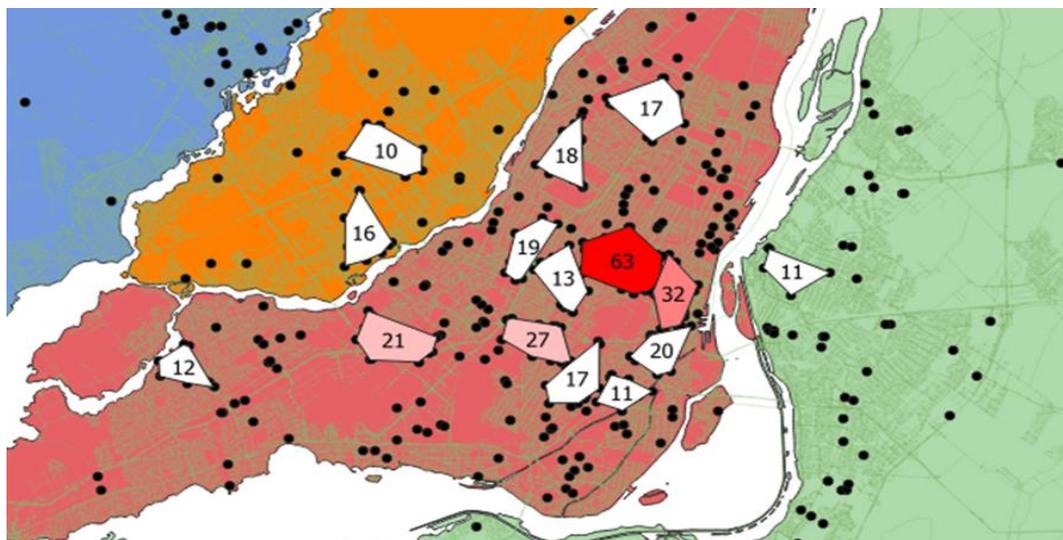


Figure 4 Hot spots du sous-secteur de fabrication d'aliments (311)

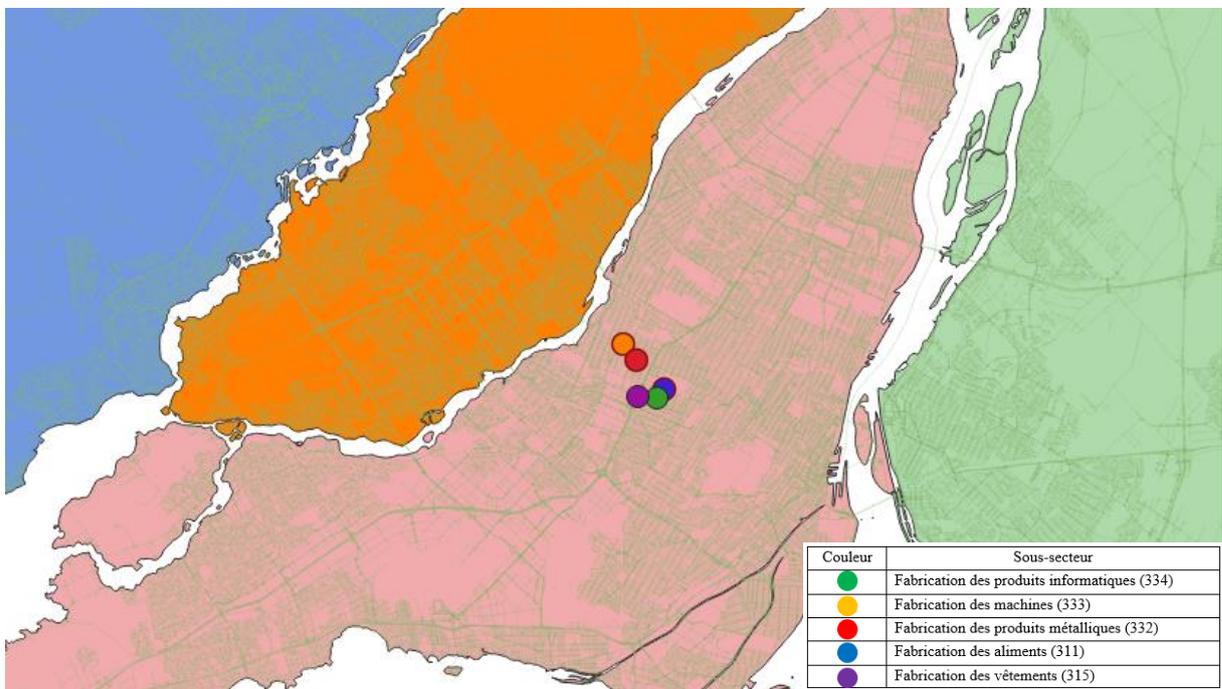


Figure 5 Répartition des centres moyens associés aux sous-secteurs de fabrication

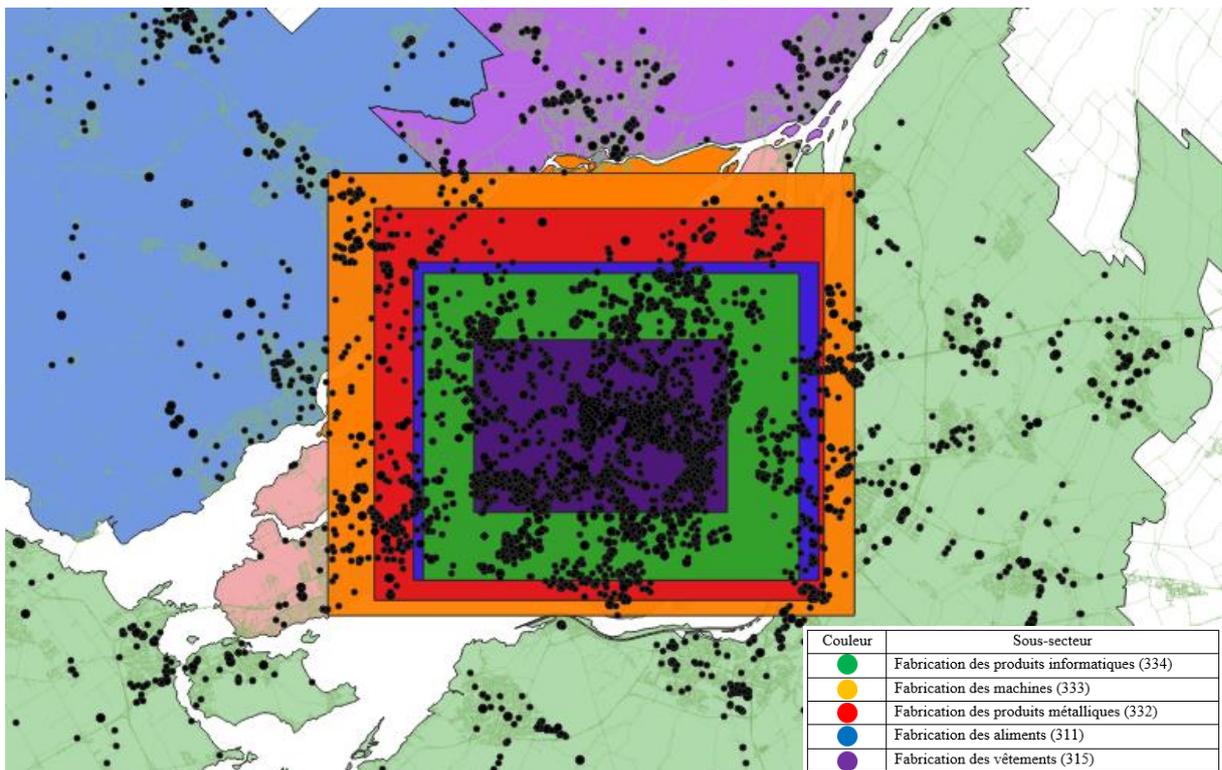


Figure 6 Répartition des sites logistiques associés aux sous-secteurs de fabrication avec leur écart type

5 CONCLUSION

L'objectif principal de cette étude est de mettre la lumière sur l'importance de la mise en œuvre de la mutualisation logistique pour atteindre des bénéfices importants au niveau de la chaîne d'approvisionnement et d'évaluer le potentiel de sa mise en place dans la grande région de Montréal.

Pour cette raison, une revue de littérature est réalisée pour tracer un portrait complet sur la mutualisation logistique. Les thèmes étudiés permettent d'accompagner un projet de mutualisation depuis les premières étapes jusqu'à la réalisation pratique pour assurer une mise en œuvre efficace. La revue de littérature a débuté par l'identification des critères et des méthodes pour assurer une sélection efficace des bons partenaires, ensuite

analyser les techniques de répartition des coûts et des bénéfices entre les entreprises pour réaliser une allocation efficace entre tous les membres. La conception d'un réseau mutualisé pour décider les meilleures localisations des plateformes a également été traitée. L'évaluation de l'impact de la mutualisation logistique sur les trois axes du développement durable, à savoir : économique, environnemental et social, est également analysé. La disponibilité des informations en temps réel est un élément essentiel pour la réussite de la mutualisation logistique et pour établir la confiance entre ses membres d'où l'importance de définir des solutions et des systèmes convenables dans cette optique. La revue de littérature a également présenté des solutions pour assurer une planification efficace du transport.

Par la suite, une analyse de la possibilité de mettre en place une mutualisation logistique entre les sites logistiques présents dans la grande région de Montréal a été traitée en utilisant à la fois Excel par sa fonctionnalité des tableaux croisés dynamiques et le logiciel de gestion des informations géographiques QGIS. Cet article a analysé en particulier les sous-secteurs du secteur de fabrication et a présenté les hot spots trouvés au niveau du sous-secteur de fabrication des aliments. Finalement, une analyse comparative des sous-secteurs du secteur de fabrication est réalisée en utilisant le centre moyen et l'écart type.

Cet article peut constituer une référence pour les entreprises souhaitant débiter un projet de mutualisation logistique dans la grande région de Montréal. Comme perspectives, il serait pertinent d'étendre les recherches futures sur l'analyse de la confiance et de l'engagement qui sont des aspects importants pour la réussite de la mutualisation et qui doivent être présents pour chaque membre du réseau mutualisé. Une attention particulière doit être portée à la sécurité des sites et des ressources mutualisés. De plus, combiner la mutualisation logistique qui se fait entre des acteurs du même niveau du réseau logistique) et la collaboration entre des acteurs de différents niveaux du réseau logistique (par exemple entre les fournisseurs et les fabricants ou entre les fabricants et les distributeurs) permettrait de mettre les ressources mutualisées à la disposition d'un plus grand nombre d'acteurs.

6 REFERENCES

Amer, L., et al. (2014). Collaborative sustainable supply chain network design: state of the art and solution framework.

Anand, G., et al. (2012). "Measuring horizontal collaboration intensity in supply chain: a graph-theoretic approach." Production Planning & Control **23**(10-11): 801-816.

Asawasakulsorn, A. (2009). "Transportation Collaboration: Partner Selection Criteria and Interorganizational System Design Issues for Supporting Trust." International Journal of Business and Information **4**(2).

Audy, et al. (2008). Impact of benefit sharing among companies in the implantation of a collaborative transportation system-An application in the furniture industry, Springer.

Audy, et al. (2012). "An empirical study on coalition formation and cost/savings allocation." International Journal of Production Economics **136**(1): 13-27.

Audy, et al. (2011). "Cost allocation in the establishment of a collaborative transportation agreement—an application in the furniture industry." Journal of the Operational Research Society **62**(6): 960-970.

Audy, et al. (2012). "A framework for an efficient implementation of logistics collaborations." International transactions in operational research **19**(5): 633-657.

Badraoui, I., et al. (2020). "Horizontal logistics collaboration: an exploratory study in Morocco's agri-food supply chains."

International Journal of Logistics research and applications **23**(1): 85-102.

Bahinipati, B. K., et al. (2009). "Horizontal collaboration in semiconductor manufacturing industry supply chain: An evaluation of collaboration intensity index." Computers & Industrial Engineering **57**(3): 880-895.

Berger, S., et al. (2010). "Solutions to the request reassignment problem in collaborative carrier networks." Transportation Research Part E: Logistics and Transportation Review **46**(5): 627-638.

CrimeState (2013). A spatial statistics program for the analysis of crime incident locations.

Cruijssen, F., et al. (2007). "Horizontal cooperation in transport and logistics: a literature review." Transportation journal: 22-39.

Danloup, N., et al. (2013). Literature review on or tools and methods for collaboration in supply chain, IEEE.

Daoud, I., et al. (2015). Network design and planning with resource pooling: The context of merging two logistics entities, IEEE.

Dirksen, M., et al. (2017). "Evaluation of synergy potentials in transportation networks managed by a fourth party logistics provider." Transp. Res. Procedia **25**: 824-841.

DMTI Spatial. (2020). "CanMap Content Suite." from <https://www.dmtispacial.com/canmap>.

Donahue, R., et al. (2018). "Rethinking cluster initiatives."

Frisk, M., et al. (2010). "Cost allocation in collaborative forest transportation." European Journal of Operational Research **205**(2): 448-458.

Gansterer, M., et al. (2018). "Collaborative vehicle routing: a survey." European Journal of Operational Research **268**(1): 1-12.

Glaister, K. W., et al. (1997). "Task-related and partner-related selection criteria in uk international joint ventures." British Journal of Management **8**(3): 199-222.

Gonzalez-Feliu, J., et al. (2011). Collaborative transportation sharing: from theory to practice via a case study from France. Technologies for supporting reasoning communities and collaborative decision making: Cooperative approaches, IGI Global: 252-271.

Guajardo, M., et al. (2016). "A review on cost allocation methods in collaborative transportation." International transactions in operational research **23**(3): 371-392.

Institut du commerce (2013). "Guide pratique de la mutualisation logistique."

Kadokawa, K. (2013). "A search for an industrial cluster in Japanese manufacturing sector: evidence from a location survey." GeoJournal **78**(1): 85-101.

Krajewska, M. A., et al. (2008). "Horizontal cooperation among freight carriers: request allocation and profit sharing." Journal of the Operational Research Society **59**(11): 1483-1491.

Liu, P., et al. (2010). "Allocating collaborative profit in less-than-truckload carrier alliance." Journal of Service science and Management **3**(01): 143.

Lozano, S., et al. (2013). "Cooperative game theory approach to allocating benefits of horizontal cooperation." European Journal of Operational Research **229**(2): 444-452.

Lu, X. (2004). A new approach for Web-GIS based collaborative transportation planning system design, IEEE.

Makaci, M., et al. (2017). "Pooled warehouse management: An empirical study." Computers & Industrial Engineering **112**: 526-536.

Makaci, M., et al. (2014). "La mutualisation des entrepôts dans la chaîne logistique."

- Makhloufi, R., et al. (2015). "Simulation of mutualized urban logistics systems with real-time management." Transportation Research Procedia **6**: 365-376.
- Martin, N., et al. (2018). "Horizontal collaboration in logistics: decision framework and typology." Operations Management Research **11**(1-2): 32-50.
- Mat, N. A. C., et al. (2009). Partner selection: Criteria for successful collaborative network.
- Morana, J., et al. (2015). "Le tableau de bord durable d'un système mutualisé des livraisons urbaines à l'aune des préoccupations du XXIème siècle." Revue Française de Génie Industrielle **34**(1): 45-65.
- Morana, J., et al. (2014). Urban consolidation and logistics pooling. Sustainable urban logistics: Concepts, methods and information systems, Springer: 187-210.
- Moutaoukil, A., et al. (2013). Modélisation d'une stratégie de mutualisation logistique en intégrant les objectifs de Développement Durable pour des PME agroalimentaires.
- Mrabti, N., et al. (2018). Mutualisation logistique, approche par simulation.
- Mrabti, N., et al. (2019). "Sustainability indicators for a pooled logistics chain."
- Mrabti, N., et al. (2020). "The pooling of sustainable freight transport." Journal of the Operational Research Society: 1-16.
- Naesens, K., et al. (2009). "A swift response framework for measuring the strategic fit for a horizontal collaborative initiative." International Journal of Production Economics **121**(2): 550-561.
- Neubert, G., et al. (2014). "Pooling logistics as a new mean for sustainable urban freight distribution".
- Ouhader, H., et al. (2017). "Combining facility location and routing decisions in sustainable urban freight distribution under horizontal collaboration: how can shippers be benefited?" Mathematical Problems in Engineering **2017**.
- Ouhader, H., et al. (2016). A two-echelon location-routing model for designing a pooled distribution supply chain, IEEE.
- Pan, S., et al. (2013). "The reduction of greenhouse gas emissions from freight transport by pooling supply chains." International journal of production economics **143**(1): 86-94.
- Pan, S., et al. (2014). "Environmental and economic issues arising from the pooling of SMEs' supply chains: case study of the food industry in western France." Flexible Services and Manufacturing Journal **26**(1-2): 92-118.
- Pomponi, F., et al. (2015). "Trust development and horizontal collaboration in logistics: a theory based evolutionary framework." Supply Chain Management: An International Journal.
- Reaidy, P. J., et al. (2015). "Bottom-up approach based on Internet of Things for order fulfillment in a collaborative warehousing environment." International Journal of Production Economics **159**: 29-40.
- Rodrigues, V. S., et al. (2015). "Horizontal logistics collaboration for enhanced supply chain performance: an international retail perspective." Supply Chain Management: An International Journal.
- Sakai, T., et al. (2020). "Location factors for logistics facilities: Location choice modeling considering activity categories." Journal of Transport Geography **85**: 102710.
- Senkel, M.-P., et al. (2013). "La mutualisation logistique: entre théories et pratiques." Logistique & Management **21**(1): 19-30.
- Sheffi, Y., et al. (2019). "New forms of partnership: the role of logistics clusters in facilitating horizontal collaboration mechanisms." European Planning Studies **27**(5): 905-931.
- Sitadewi, D., et al. (2018). "The Conceptual Framework of Horizontal Collaborative Transportation Management in Indonesian Trucking Industry." International Journal of Modern Trends in Business Research **3**(4): 42-58.
- Verdonck, L., et al. (2016). "Analysis of collaborative savings and cost allocation techniques for the cooperative carrier facility location problem." Journal of the Operational Research Society **67**(6): 853-871.
- Wang, X., et al. (2014). "Collaborative transportation planning of less-than-truckload freight." OR spectrum **36**(2): 357-380.
- Winkelhaus, M., et al. (2014). "Partner Selection for Horizontal Logistic Cooperation."
- Wong, H., et al. (2007). "Cost allocation in spare parts inventory pooling." Transportation Research Part E: Logistics and Transportation Review **43**(4): 370-386.
- Zouari, A. (2019). "Transport Pooling: Moving Toward Green Distribution." Solving Transport Problems: Towards Green Logistics: 63-95.

Truck platooning transportation planning: A review with emphasis on Operations Research methods

SABA GAZRAN^{1,2,3}, TASSEDA BOUKHERROUB^{1,2,3}, MIKAEL RÖNNQVIST^{2,3,4}, MARC PAQUET^{1,2}

saba.gazran.1@ens.etsmtl.ca, tassedaboukherroub@etsmtl.ca,
mikael.ronnqvist@gmc.ulaval.ca, marc.paquet@etsmtl.ca

¹ Department of Systems Engineering & Numerix Laboratory, École de technologie supérieure
1100 Notre-Dame West, Montréal, QC, H3C 1K3, Canada

² Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)
Université de Montréal, Pavillon André-Aisenstadt, CIRRELT, Canada

³ FORAC Consortium,
1065, av. de la Médecine, Université Laval, Québec, QC, G1V 0A6, Canada

⁴ Department of mechanical engineering
1065, av. de la Médecine, local 1504 Université Laval, Québec, QC, G1V 0A6, Canada

Résumé – Le concept de camions en peloton est le fait de relier deux ou plusieurs camions entre eux pour former un convoi, en utilisant des technologies de connectivité et de conduite autonome. Ce système de camions ainsi connectés permet de maintenir une courte distance entre les véhicules sur une partie du trajet. Ce système présente plusieurs avantages; les deux plus importants sont la réduction de la consommation du carburant et la réduction du nombre de conducteurs requis. Ce système pourrait donc permettre de réaliser des livraisons plus économiques et de générer moins d'émissions de GES. Dans les dernières années, le système de camions en peloton est devenu un domaine de recherche important pour l'industrie du transport. Par exemple, une planification efficace et la coordination entre les camions formant le peloton sont nécessaires. Cet article présente une revue de la littérature sur le problème de planification du transport utilisant les systèmes de camions en peloton dans le domaine de la Recherche Opérationnelle.

Abstract – Truck platooning links two or more trucks in a convoy, by using connectivity and automated driving technologies. These connected vehicles automatically maintain close distance between each other for certain parts of a journey. Truck platooning has many benefits; the two most prominent are lower fuel consumption and convoy transportation and less human drivers. Therefore, truck platooning might lead to more economical shipping and fewer greenhouse gas (GHG) emissions. In recent years, truck platooning has become an important research field for the transportation industry. For example, efficient planning and coordination among the vehicles are required. This paper presents a literature review on the problem of truck platooning transportation planning in the field of Operations Research.

Mots clés – Camions en peloton, planification du transport, Optimisation, Recherche Opérationnelle

Keywords – Truck platooning, transportation planning, Optimization, Operations Research.

1 INTRODUCTION

Freight transportation is a primary component of the supply chains, ensuring the movement and delivery of raw materials and finished products (Crainic, 2003). Based on product supply chain tendencies in recent decades, which nowadays are more complex and global, the importance of freight transportation has become higher. In a general transportation problem, specific amounts of one or multiple products are available in limited quantities at several origins in a typical transportation problem. Products need to be moved to destinations at a minimum total cost. The demand amounts at the destinations are satisfied, and the availability of supplies at the sources is respected. The transportation could be done by the company itself, carrier, or a Logistics Service Provider (LSP). Furthermore, the transporter could consider starting and delivering time constraints, time windows, the size and quantities of products, and legal and technical requirements to satisfy customers' demands (Wieberneit, 2008).

Companies are exploring solutions to improve their process efficiency and lower their cost. Consequently, as a significant section, logistics-related companies have been obliged to decrease their costs using improved planning, efficient use of resources, and collaboration with other companies (SteadieSeifi et al., 2014). As an example, in the forestry sector, transportation costs in countries such as Canada and Sweden account for a significant proportion (around one-third) of the industry's overall raw materials cost (Frisk et al., 2010). Based on an analysis of operational costs of trucking in the USA (Murray et al., 2019), the two highest costs are driver-related costs (driver wages represent 30% and driver benefits 10%) and fuel costs (24%). Moreover, the forestry industry in Canada faces driver shortages in forestry transportation operations, especially in raw material transportation from harvesting areas (Canadian Forest Industries magazine, 2018). Consequently, it is of great importance to improve fuel and driver economy in transportation industries.

Furthermore, governments and the public have brought new

challenges for logistics companies in terms of reducing their negative impacts on the environment. In Canada, transport activities (people and product transportations) are responsible for 30% of greenhouse gas (GHG) emissions in 2018 (Government of Canada, 2020); this percentage was 28% in the USA in 2018 (United States Environmental Protection Agency, 2020). As for 28 European countries (EU-28) in 2017, transportation accounted for 27% of total GHG emissions (The European Environment Agency, 2019). As it is clear, the transportation sector has produced more than a quarter of GHG emissions in North America and Europe, which is a significant portion. Therefore, logistics companies must manage their services to be more sustainable and contribute to less GHG emissions. Some initiatives have been implemented to tackle these challenges, especially using electric vehicles (Schiffer et al., 2018).

Some researchers have identified systematic information sharing, collaboration, using autonomous and connected vehicles in logistics processes as significant trends in supply chain management (Porter et al., 2015; Speranza, 2018). In addition to technological achievements, more effective transportation planning can contribute to improving economic and environmental performances. Concepts of the 4th generation of the industrial revolution (Industry 4.0) and the Internet of Things (IoT) in which machines, objects, persons, and vehicles of all types are interconnected, have affected different sectors of businesses such as production, logistics, and so on (Gilchrist, 2016). Autonomous and Connected Vehicles are one of the Industry 4.0 technologies as they exhibit an excellent connection between the digital and real-world (Pieroni et al., 2018).

Since the proper planning of truck platooning transportation planning plays an essential role in and the research in this area has gained more attention in recent years. There exist some recent literature reviews on truck platooning transportation. Bhoopalam et al. (2018) have provided a literature review paper that addressed truck platooning transportation planning from an Operations Research view up to 2017. Axelsson (2016) reviewed the safety aspects of vehicle platooning. In another review, L. Zhang et al. (2020) studied the literature about fuel savings aspects of truck platooning. According to an increasing number of studies about truck platooning transportation planning from 2017 to 2020, it is necessary to review the new studies and survey the uncovered research areas. Our literature review focuses on truck platooning transportation planning using Operations Research methods.

This paper's remainder is organized as follows: Section 2 describes the truck platooning transportation. In the third section, the review of the literature about truck platooning transportation planning will be presented. Section 4 will present our conclusions and future research directions.

2 TRUCK PLATOONING

This section introduces the concepts and benefits of intelligent and connected transportation systems and truck platooning as a part of it.

2.1 Intelligent and connected transportation systems

The progress of information and communication technology and intelligent transportation systems has led to significant transportation improvements in recent years. With current achievements and trends in communication-related technologies such as vehicle-to-vehicle (V2V), vehicle-to-

infrastructure, and vehicle-to-everything (V2X) communications, performance is expected to be further enhanced in transportation (Talebpoor et al., 2016; Van de Hoef, 2016). The mentioned types of communications provide a great infrastructure for linking transportation components together. The components could contain all the vehicles on the streets and roads (e.g., cars, trucks, motorcycles), traffic infrastructures (e.g., traffic lights, fuel, and charging stations), and even pedestrians' communication devices could communicate with each other (Amadeo et al., 2016). At the vehicle level, enormous progress has already been achieved in developing progressive driver assistance systems and autonomous vehicles (Van de Hoef, 2016). Technologies such as adaptive cruise control (ACC), cooperative adaptive cruise control (CACC), and lane-keeping assistance are basic enablers for designing truck platooning systems (Bhoopalam et al., 2018).

Modern technologies, such as connected and autonomous vehicles and electric vehicles, are progressing rapidly (Liu et al., 2018). It is likely to see a fleet of electric, autonomous vehicles connecting to the road system, the internet, and a network of transportation services in the near future (Mahmassani, 2016). Connected vehicles could lead to three primary purposes: safety improvements, mobility enhancements, and system efficiency (US Department of Transportation, 2020). The National Highway Transportation Safety Administration outlines six different levels of progressive automation. These levels start with a base level (Level 0) of no automation. All the other levels use an advanced driver assistance system (automated steering and braking/accelerating). Vehicles in Level 1 use an advanced driver assistance system (automated steering and braking/accelerating) that occasionally assist the driver, but the system's actions are not occurring at the same time. In Level 2 automated steering and braking/accelerating could be applied simultaneously under some conditions. In Level 3 the vehicle can make all the decisions of the driving under some conditions which the driver needs to be ready to take action if it is necessary. In Level 4 the vehicle is able to do all driving tasks and control the driving environment and there are some situations that the human driver does not need to watch out for the vehicle. Level 5, which shows full self-driving automation and the vehicle can do all the actions in all conditions. (National Highway Traffic Safety Administration, 2016).

The autonomous vehicles can drive with no external assistance using a system that can collect data from the environment and process them, then act based on the results through sensors, radars, lidars, and artificial intelligence. Sharing the vehicle's information (e.g., vehicle location, speed, origin, and destination), using vehicles' connectivity could lead to better performance of the vehicles and transportation system (Scherr et al., 2019).

2.2 Truck platooning transportation

A truck platoon is a set of virtually connected trucks (two trucks or more) that travel in a convoy with small intervehicle distances. The follower trucks automatically steer, brake, and accelerate with the platoon leader vehicle, using autonomous driving technology (L. Zhang et al., 2020). Vehicle platooning is an inspiring trend among developing technologies that could offer advantages such as reducing energy (fuel) consumption, decreasing road space usage, and safety improvement (Bhoopalam et al., 2018; Boysen et al., 2018; Larson et al., 2016). The trucks within a platoon move in close intervals, leading to air resistance reduction in the platoon and, as a result, lower rate fuel consumption of trucks and consequently GHG

emissions (Larson et al., 2013). Drag reductions from platooning depend on the distance between platoon vehicles, their order, and shapes (Wadud et al., 2016).

Studies reveal that platooning can decrease fuel consumption by 4-7% (Al Alam et al., 2010) and even up to 21% (Bonnet et al., 2000). The mentioned savings related to field experiments testing the fuel reductions depend on the environmental conditions, inter-vehicle distance, the driving speed, and the first weight. For example, in a platoon of size three, fuel savings could be up to 16% for the follower trucks, and the leader could save up to 7.5% (Tsugawa, Kato, and Aoki, 2011). In another research, Larson et al. (2013) showed by simulation which truck platooning could lead to fuel savings from 1-9% for the output of the whole transportation system.

From an automated driving perspective, there are three types of truck platooning; Bhoopalam et al. (2018) proposed the following classification to describe different levels of human involvement in platooning: Human-driven platooning with in-platoon resting, Hybrid platooning and Driverless platooning. The first one is when drivers are at the wheels of the vehicles (automation level 3). In the second case, only the leading truck has a driver (automation level 4). In the third model, there is no driver in any truck (automation level 5). The first type can only benefit from fuel consumption reduction, while the second and third types can tackle driver shortage issues and lead to driver costs reduction and fuel consumption decrease.

Hybrid platooning has more benefits in some aspects that need fewer drivers that lower transportation costs and lower need for human drivers. However, they have lower flexibility since the driverless trucks cannot move on to their next destinations without the lead truck. Also, some processing times may take longer for them, such as loading/unloading and refueling. One challenge in forming truck platoons is related to the need to synchronize the departure times, travel speeds, and truck routes (Bhoopalam et al., 2018). For example, a truck may change its route and perhaps even make a reasonable detour to join or form a platoon. Consequently, this needs high coordination and planning of platoons, as otherwise, trucks and drivers would wait idly for an unacceptable length of time and inefficiency (Boysen et al., 2018). The mentioned coordination comprises the matches between trucks and information sharing between companies that use truck platooning in their transportation. Coordination could enhance transport efficiency and create trust between companies and the willingness to participate in mutual plans, especially in the initial phases of implementing the platooning technology. Therefore, coordination and transportation planning play an essential role in increasing truck platooning transportation benefits.

3 LITERATURE REVIEW

This section reviews the literature about transportation planning for truck platoons and shows the methodology used for selecting relevant studies and classify the literature.

3.1 Search method

This paper considers only studies reporting about truck platooning, and excludes researches about other platooning studies, like passenger public transportation (Liu et al., 2018) and bus platooning (Sethuraman et al., 2019). Also, the review considers only truck platooning transportation planning using Operations Research methods. Therefore, other research methods such as data mining techniques (Meisen et al., 2008) and simulation (Liang et al., 2013) are excluded.

Studies were collected by using the following combinations of keywords: “truck platooning” AND “transportation” OR “logistics” AND “planning” OR “routing” OR “scheduling” OR “network design”. Online electronic databases used are: Web of Science, ScienceDirect, and Springer. Moreover, some articles were added by tracking cited references in Google Scholar. Next, the references were exported for selection and further analyses. A total number of 94 studies resulted from this process. Non relevant articles were removed by reviewing the title and abstract. After that, each article's full text was skimmed to identify whether it was related to the topic (Operations Research), and finally, 21 publications have been selected for further analysis. Consequently, each remaining article was reviewed and categorized by its methodology and content.

3.2 Literature classification

At the operational planning level, platoon coordination and vehicle routing problems in parallel is a combinatorial optimization problem (Sokolov et al., 2017a). Truck platooning routing and scheduling in some aspects is similar to other studied problems. These problems are the Swap-Body Vehicle Routing Problem (Lum et al., 2015), ridesharing (Furuhata et al., 2013), the dial-a-ride problem (Cordeau et al., 2007), automated guided vehicle scheduling and routing problems, Convoy Movement Problem (Gopalan, 2015; Mokhtar et al., 2020), which itself is comparable to routing of Automated Guided Vehicles in Flexible Manufacturing Systems environment (Fazlollahabbar et al., 2015), movement of luggage from different flights along with a standard automated transportation system to various pick-up points, strategic routing of hazardous materials and scheduling of trains on a single line track (Kumar et al., 2010). Furthermore, the Minimum Cost Sharing Pairwise Distance Preservers problem can be applied in vehicle platooning (Abdolmaleki et al., 2020).

Larsson et al. (2015) have shown that platooning's operational transportation planning problem is an NP-hard problem, even in its basic form, when all trucks start at the same time and level. Therefore, it is extremely hard to find exact solutions for medium and scale instances. Some studies, such as Duret et al. (2020), have tested the feasibility and demonstrate their proposed strategy's consistency and robustness, and tested the results through simulation experiments. Some researchers (Nourmohammadzadeh et al., 2016, 2019) have used metaheuristics algorithms to solve their problems.

It is crucial to put the advantages and disadvantages, and features of truck platooning into account for more suitable transportation planning. Transportation planning for truck platooning transportation can be studied at three stages: strategic, tactical, and operational. Strategic planning deals with the effects of truck platooning characteristics on long-term decisions in transportation, e.g., the structure of the transportation network, deciding about investments in infrastructures and vehicles. Tactical truck platooning, transportation planning is about mid-term decisions, e.g., material flow consideration between origin, terminal, transshipment, destination nodes, and service network design. In truck platooning operational transportation planning refers to short-term decisions, e.g., daily routing and scheduling.

We propose to classify the literature review following three main categories: truck platooning characteristics, criteria considered, and optimization methods. The following subsection analyzes the published studies regarding the mentioned categories. The main features of each study are

summarized in Table 1, which contains the information about each research such as modeling approach of the papers, decisions variables considered in them, type of objective function and constraints in mathematical models, solution methods to solving the numerical cases and also it includes the information about automation type of platooning, transportation planning level and countries/industrial sector if it is mentioned in each paper. Figure 1 shows the criteria considered in the objective function of the mathematical models proposed while Figure 2 shows the proportion of solution approaches developed.

3.2.1 Operational transportation planning

Larson et al. (2013) studied one of the first papers that appeared in this research area. This study initiated the scheduling of truck platoons with flexible routing decisions. Using information from local infrastructure in the transportation network, the authors optimized the objective function. They considered a piecewise concave function of fuel consumption for platoon members based on position, destination, and the speed of vehicles. Small instances of the problem were solved exactly, and the authors suggested a fast heuristic approach that could provide a solution for a large number of fleets in real-time. The authors also applied the model to simulated data of the German autobahn road network. Larson, Munson, and Sokolov (2016) used truck platooning problem features like the maximum length of the detour to decrease the problem's size. They proposed a bound on the maximum length of detour constraint under which platooning fuel savings outweigh the actual fuel costs for each vehicle and solved small and medium instances to optimality. Furthermore, the authors applied their method on large network and truck instances on the Chicago highway network with results within a one percent optimum gap.

Larsson et al. (2015) defined a truck platooning routing problem, which is used in a given line part of the road network in real-time. They proposed an integer linear programming model (ILP) for the problem and proposed two heuristics to solve the model. The problem's objective was to minimize fuel consumption, with no time windows consideration, making it an unlimited platooning problem (because there was no limitation on start and delivery time windows consideration in the problem). The authors were the first to introduce an ILP formulation for truck platooning routing. Generally, the more trucks moving on the road, the more fuel savings would be achieved in the truck platooning schedule and route planning.

Van De Hoef et al. (2015) studied the problem of coordinating the formation and breakup of platoons to optimize fuel consumption and consider various speeds in trucks' movement. The authors proposed a heuristics-solving approach, in which based on the shortest paths between nodes platoons are formed, and then a speed level is allocated to them. They also proposed an approximate solution, where each truck takes its shortest path, then the approach produces some convex optimization problems. In another research, Van de Hoef (2016) considered a problem in which truck routes and speeds are planned by a centralized coordinator given their destinations and delivery deadlines. The authors have formulated a MIP combinatorial optimization problem that combines these plans to obtain minimal fuel consumption. An algorithm to compute optimal solutions to this problem is developed in which routes are generated and all possible combinations of trucks that could form platoons along the routes are identified.

Nourmohammadzadeh et al. (2016) presented a MIP model for

the platoon formation problem. The inputs of this problem are a set of trucks with their origin, destination, ultimate times for pick-up and delivery, the road network structure, and the fuel-saving factor for the trucks in a platoon. The outputs are routes and transportation schedules for each truck. The authors tested their model on a case study for a various number of trucks on a network of cities in Germany. They implemented a Genetic algorithm and compared the outputs with optimal solutions. Later the same authors (Nourmohammadzadeh et al., 2019) studied a similar platoon formation and routing problem and implemented an Ant Colony Algorithm to solve large-scale networks.

Sokolov et al. (2017a) considered the problem of the central strategic coordination of platoon-able vehicles. The problem's objective is to minimize the vehicle's total fuel consumption by coordinating their departure time and routing, which is performed jointly. Moreover, a coordinated platooning optimization MIP model is presented, and the impact of optimal truck platoon routing is compared with the basic strategy. In another research, Sokolov et al. (2017b) have considered a similar problem which optimizes the fuel savings of platooning vehicles by routing and schedule their departure times. The resulting problem is combinatorial optimization and a MIP model, which considers platoon formation and routing of vehicles. The authors evaluate their methodology by comparing the benefits of a coordinated solution with the case without coordination.

W. Zhang et al. (2017) studied the planning of freight transport platoon and departure time scheduling problems while considering travel time uncertainty. The objective function contained the cost of trip time, the scheduling penalties, and the fuel cost in a MIP model. Platooning is shown to be useful only if the planned arrival times alter by less than a certain amount. They concluded that platooning on converging routes would lead to more significant schedule changes and driving costs than on diverging routes, where vehicles split after platooning. Johansson et al. (2018) considered the platoon formation problem in a group of trucks having the same origin but different destinations. The vehicles belong to carriers in the MIP model. Strategic cooperation has been modeled as a non-cooperative game in which the main decision is related to departure time coordination for the trucks. The authors modeled the platoon matching by game theory concept and proposed an algorithm for finding a Nash equilibrium of the game in which vehicles with the same departure time are matched to form a platoon.

Luo et al. (2018) offered an iterative route-then-schedule approach for centralized platoon transportation planning. The authors also suggest and evaluate a set of valid inequalities for the problem related to assigning vehicles to routes and providing the required times vehicles move in their routes. They formulated the problem as a tight MIP with a continuous-time variable with different transportation speeds to improve platooning formation through waiting times. They also proposed a heuristic approach called the cluster-first route-second decomposition method to solve the problem. In another research, Luo et al. (2020) proposed an iterative route-then-schedule approach for the centralized platooning formation and routing planning that converges to high-quality solutions in a reasonable time. In this research routing and scheduling, decisions are combined in two sub-problems in solution method which is quicker converges to the optimal solutions. They also recommended and evaluated a group of inequalities for the problems to reduce the computational time or optimality gap in

large instances.

Boysen et al. (2018) studied platooning and transportation-related factors such as the distribution of platooning technology, time windows, the maximum number of trucks in a platoon, and truck drivers' willingness to wait for joining a platoon. The authors focused on platoon formation from an operational level with no fixed routes to investigate platooning's advantages and its effect on increasing logistics performance. Finally, the authors concluded that, if schedules were tight, delay costs would exceed fuel savings' benefits. Sun et al. (2019) investigated platoon formation to optimize platooning's advantages regarding fuel-saving. The authors then used the results to fairly share and allocate the platooning benefits to form and promote stable platoon formation based on Game Theory concepts.

Abdolmaleki et al. (2019) focused on scheduling and routing for a given collection of trucks to optimize fuel savings. The authors formulated the problem as a minimum concave-cost on a time-expanded network. They applied an outer approximation algorithm, a dynamic programming-based heuristic, and a fast approximation algorithm to solve the large-scale instances. You et al. (2020) have focused on the truck platooning operation combined with the Local Container Drayage Problem in which only the leading truck is human-driven. In this study, two distinct drivers can serve one customer by coordination, and customers can share one container without returning to the depot. The authors introduced a heuristic method including a learning mechanism for solving the problem. They concluded that cost effectiveness of the model could be achieved by increasing the platoon's size and/or well-coordinating container sharing among customers.

Sebe, Kraus, et al. (2019) studied the problem of platoon formation with trucks from different Logistic Service Providers (LSPs) which are equipped with fleets of autonomously driving trucks in a same-day delivery context. This paper has studied platoon formation mechanisms and achieve the benefits of cross-provider platooning using simulation. The authors formulated the platoon formation as an ILP problem to find the longest sub-routes shared between vehicles. In a similar study, (Sebe & Müller, 2019) considered heterogeneous fleets and cost minimization objective. This paper's contribution is a formation and routing algorithm for heterogeneous vehicles.

Albiński et al. (2020) studied the centralized planning of platoon formation and routing. They considered the day-before planning problem for routing and scheduling of the trucks, and they suggested a new mathematical model that could be used in three automation stages of platooning under platoon size limits, hard time windows, and driving time regulations. Also, the authors developed a pre-processing procedure that facilitates solving large instances in a reasonable time. Furthermore, they tested the proposed models in different scenarios in city logistics. Larsen et al. (2019) presented a model for optimizing the problem of platoon formation at a hub while considering mandatory breaks and the possible case that driving time for the platoon followers can be counted as rest time and that drivers can skip breaks. This research aims to minimize the total travel cost of the whole system through fixed routes. The authors solved the problem using dynamic programming based local search heuristic and concluded that effectiveness depends on whether drivers are allowed to rest while driving in platoons and in other cases the savings are very limited.

3.2.2 Tactical transportation planning

Scherr et al. (2018) proposed a service network design solution for parcel delivery with truck platooning. The authors identified a two-tier city logistics scenario where autonomous trucks in different regions satisfy customer demands by the guide of trucks with a driver. In the mentioned problem, trucks form a platoon in which the leading truck has a human driver to move from the depot to outside the area or reverse. The authors formulated the problem as a MIP model. The results showed that this idea could lead to cost savings. Scherr et al. (2019) studied a similar problem. They considered heterogeneous vehicles in which autonomous vehicles can only drive in feasible zones but require control elsewhere by driver leaders in platoons. This model determines the size and mix of the fleet of vehicles and flow routing. The authors stated that cost savings depend on the strategies to coordinate the mix of the fleet, supply, and demand.

4 DISCUSSION

In this paper, we described recent studies reporting on research efforts conducted in truck platooning transportation planning. The review focused on OR modeling and solution approaches related to this problem. According to the review, most of the papers have studied the operational truck platooning transportation planning and only two papers have studied tactical platooning transportation planning. Furthermore, strategical transportation planning or network design problem considering truck platooning characteristics in transportation networks have not been studied in the literature.

At the operational decision level, studies mostly have considered human-driven truck platooning (platooning with drivers), but hybrid truck platooning (the case with only one driver in the leading truck); is not sufficiently investigated in the literature. Most of the research aimed to minimize fuel consumption to show the potential of truck platooning. Other relevant aspects such as driver costs and GHG emissions have not been adequately investigated in the mathematical models. Another potential research avenue is investigating truck platooning with electric trucks. As recharging these trucks would be considerably longer than refueling traditional trucks, transportation cost is likely to increase due to higher idle times. Truck platooning could reduce the number of recharges needed for electric battery vehicles since less energy is expended, and therefore, greater distances can be covered. Finally, in real-world cases, some parameters might be dynamic or stochastic. These aspects could affect the efficiency of truck platooning in the transportation networks. The number of research papers that consider this aspects in truck platooning transportation is limited.

Table 1 Truck platooning transportation planning studies using Operations Research methods

Author (Year)	Modeling approach	Decision variables	Objective function optimization	Constraints*	Solution method	Type of platooning	Planning level	Country/sector
(Luo et al., 2020)	MIP	Routing and platooning, time and platooning	Fuel consumption	PS, R, NB, T	Exact, Heuristics	Human-driven	Operational	Germany/-
(Albiński et al., 2020)	MIP	Platooning, time, break and rest	(drivers + fuel + platoon formation) costs	PS, NB, T	Exact	Human-driven, driverless	Operational	Germany, Canada, USA/-
(You et al., 2020)	MIP	Routing	(drivers + travel + trucks) costs	PS, R, D, F, NB, T, S&D	Exact, Metaheuristics	Hybrid	Operational	-/city logistics
(Sebe & Müller, 2019)	ILP	Routing, flow	Transportation distances	PS, R, F, NB, T	Exact, Simulation	Human-driven	Operational	Germany/city logistics
(Sebe, Kraus, et al., 2019)	ILP	Routing, flow	Transportation distances	PS, S, F, NB, T	Exact, Simulation	Human-driven	Operational	Germany/city logistics
(Larsen et al., 2019)	MIP	Platooning, time	Waiting times & transportation distances	PS, FP	Exact, Simulation	Human-driven	Operational	Germany/-
(Abdolmaleki et al., 2019)	MIP/ MINLP	Routing	Fuel consumption	R, S, F, T	Exact, Heuristic	Human-driven	Operational	Germany/-
(Nourmohammadzadeh et al., 2019)	MIP	Routing, fuel cost	Fuel consumption	D, NB	Exact, Metaheuristic	Human-driven	Operational	Sweden/-
(Scherr et al., 2019)	MIP	Routing, flow	(drivers + travel + trucks) costs	PS, F, NB, S&D, NC, AC, FN, PF	Exact, Heuristic	Hybrid	Tactical	-/city logistics
(Sun et al., 2019)	Game Theory (MIP)	Platoon formation, speed, Cost/benefits sharing	Fuel consumption & GHG emissions	PS, S, PF	Exact	Human-driven	Operational	-/-
(Boysen et al., 2018)	MIP	Platoon formation	(drivers + fuel) costs	PS, T, PF	Exact, Heuristic	Human-driven	Operational	Germany/-
(Luo et al., 2018)	MIP	Routing (speed-dependent), platoon formation (speed-dependent), time	Fuel consumption	PS, R, S, T, PF	Exact, Heuristic	Human-driven	Operational	USA/-
(Scherr et al., 2018)	MIP	Routing, flow	(drivers + travel + trucks) costs	PS, R, F, NB, S&D, AC, FN, PF	Exact, Heuristic	Hybrid	Tactical	-/city logistics
(W. Zhang et al., 2017)	MIP	Time and platooning	(fuel + time) costs	FP, T	Exact	Human-driven	Operational	-/-
(Sokolov et al., 2017b)	MIP	Routing, platooning, time	Fuel consumption	D, T, FP	Exact	Human-driven	Operational	USA/-
(Larson et al., 2016)	MIP	Routing, platooning, time	Fuel consumption	R, D, F, NB, T, PF	Exact	Human-driven	Operational	USA/-
(Van de Hoef, 2016)	MIP	Time and platooning	Fuel consumption	S, FP	Exact	Human-driven	Operational	Sweden /-
(Nourmohammadzadeh et al., 2016)	MIP	Routing, fuel cost	Fuel consumption	D, NB, T	Exact, Metaheuristic	Human-driven	Operational	Germany/-
(Van De Hoef et al., 2015)	MIP	Time and platooning	Fuel consumption	S, PF	Exact	Human-driven	Operational	Sweden /-
(Larsson et al., 2015)	ILP	Routing, time, fuel cost	Fuel consumption	R, NB,	Exact, Heuristic	Human-driven	Operational	Germany/-
(Larson et al., 2013)	-	Routing and platooning, time and platooning	Fuel consumption	PS, D, S, T	Exact, Heuristic	Human-driven	Operational	Germany/-

*PS = Platoon size, R=Routing, D=Detour, S=Speed, F=Flow conservation, NB=Node balancing, T=Time, S&D=Supply and demand, NC=Node capacity, AC=Arc capacity, FN=Fleet NO., PF=Platoon formation, DT=Driving time regulations

As it can be seen in Figure 1, fuel is the criterion mostly considered in the literature while GHG emissions is considered in only one study.

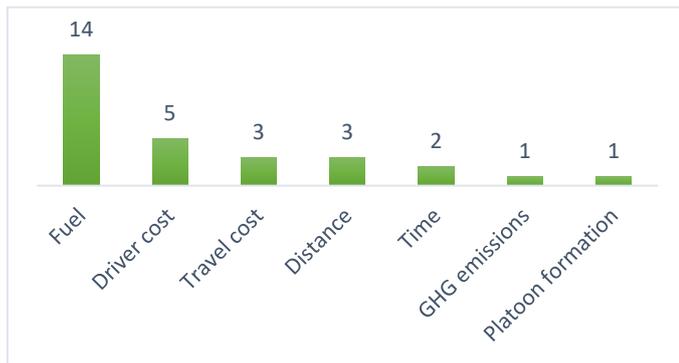


Figure 1 Criteria considered in the objective function

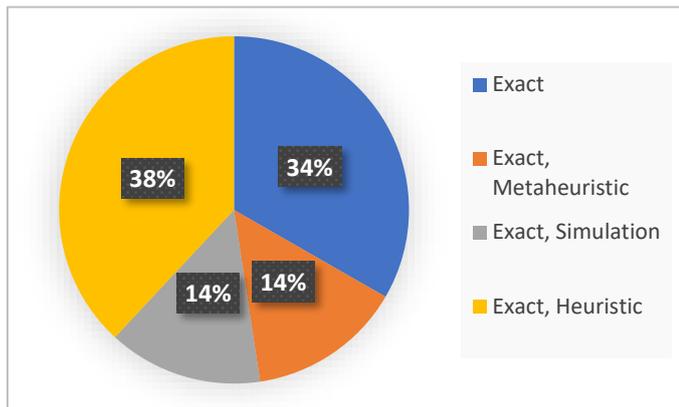


Figure 2 Proportion of solution methods proposed in the literature

The researches related to applications of truck platooning transportation in specific industries and supply chains are quite inadequate in the literature. Since material transportation in supply chains normally is performed in multi-tiers, it is more applicable to use truck platooning transportation in more complex and multi-echelon networks. Furthermore, analyzing platoon-able trucks in heterogeneous fleet transportation planning along with ordinary trucks has not been studied sufficiently in the literature. Also, Collaboration between logistics companies could lead to more truck platooning and transportation benefits for all the companies in the supply chains. In this area, fair benefits/cost-sharing between the companies would lead to more stable collaborations. These types of researches in the literature have not been studied vastly.

5 CONCLUSION AND RESEARCH PERSPECTIVES

Truck platooning could improve sustainability performance (financial, environmental, and social) in transportation operations. To achieve these potentials, efficient planning and optimization techniques are required to make the best use of platooning advantages. To the best of our knowledge, the following subjects have not been considered or have not been sufficiently investigated in the literature on truck platooning transportation planning:

- Strategic truck platooning transportation planning in the transportation networks.
- Platooning transportation planning in complex transportation and multi-echelon networks.
- Truck platooning transportation planning with stochastic parameters and uncertainty.

- Collaborative truck platooning transportation planning between multiple companies and fair costs/benefits sharing between them.

- Applying truck platooning transportation in specific supply chains (e.g., forestry and food supply chains) to investigate the advantages of this technology (compared to normal trucks).

6 ACKNOWLEDGMENT

This research project is funded by ÉTS establishment funds/NSERC Discovery grant of Pr. Tassedra Boukherroub and MITACS.

7 REFERENCES

- Abdolmaleki, M., Shahabi, M., Yin, Y., & Masoud, N. (2019). Itinerary planning for cooperative truck platooning. Available at SSRN 3481598.
- Abdolmaleki, M., Yin, Y., & Masoud, N. (2020). Minimum Weight Pairwise Distance Preservers. *arXiv preprint arXiv:2007.07554*.
- Al Alam, A., Gattami, A., & Johansson, K. H. (2010). *An experimental study on the fuel reduction potential of heavy duty vehicle platooning*. Paper presented at the 13th International IEEE Conference on Intelligent Transportation Systems.
- Albiński, S., Crainic, T. G., & Minner, S. (2020). *The Day-before Truck Platooning Planning Problem and the Value of Autonomous Driving*: CIRRELT.
- Amadeo, M., Campolo, C., & Molinaro, A. (2016). Information-centric networking for connected vehicles: a survey and future perspectives. *IEEE Communications Magazine*, 54(2), 98-104.
- Axelsson, J. (2016). Safety in vehicle platooning: A systematic literature review. *IEEE Transactions on Intelligent Transportation Systems*, 18(5), 1033-1045.
- Bhoopalam, A. K., Agatz, N., & Zuidwijk, R. (2018). Planning of truck platoons: A literature review and directions for future research. *Transportation Research Part B: Methodological*, 107, 212-228.
- Bonnet, C., & Fritz, H. (2000). *Fuel consumption reduction in a platoon: Experimental results with two electronically coupled trucks at close spacing* (0148-7191). Retrieved from
- Boysen, N., Briskorn, D., & Schwerdfeger, S. (2018). The identical-path truck platooning problem. *Transportation Research Part B: Methodological*, 109, 26-39.
- Canadian Forest Industries magazine. (2018). FPIInnovations and partners successfully test forestry truck platooning. Retrieved from <https://www.woodbusiness.ca/fpinnovations-and-partners-successfully-test-forestry-truck-platooning-5294/>
- Cordeau, J.-F., & Laporte, G. (2007). The dial-a-ride problem: models and algorithms. *Annals of operations research*, 153(1), 29-46.
- Crainic, T. G. (2003). Long-haul freight transportation. In *Handbook of transportation science* (pp. 451-516): Springer.
- Duret, A., Wang, M., & Ladino, A. (2020). A hierarchical approach for splitting truck platoons near network discontinuities. *Transportation Research Part B: Methodological*, 132, 285-302.
- Fazlollahab, H., & Saidi-Mehrabad, M. (2015). Methodologies to optimize automated guided vehicle scheduling and routing problems: a review study. *Journal of Intelligent & Robotic Systems*, 77(3-4), 525-545.
- Frisk, M., Göthe-Lundgren, M., Jörnsten, K., & Rönnqvist, M. (2010). Cost allocation in collaborative forest transportation. *European journal of operational research*, 205(2), 448-458.
- Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.-E., Wang, X., & Koenig, S. (2013). Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological*, 57, 28-46.
- Gilchrist, A. (2016). *Industry 4.0: the industrial internet of things*: Springer.
- Gopalan, R. (2015). Computational complexity of convoy movement planning problems. *Mathematical Methods of Operations Research*, 82(1), 31-60.
- Government of Canada. (2020). Greenhouse gas sources and sinks: executive summary 2020. Retrieved from <https://www.canada.ca/en/environment-climate-change/services/climate-change/greenhouse-gas-emissions/sources-sinks-executive-summary-2020.html>
- Johansson, A., Nekouei, E., Johansson, K. H., & Mårtensson, J. (2018). *Multi-fleet platoon matching: A game-theoretic approach*. Paper presented at the 2018 21st International Conference on Intelligent Transportation Systems (ITSC).
- Kumar, P. R., & Narendran, T. (2010). Convoy movement problem—an optimization perspective. In *Innovations in Defence Support Systems-1* (pp. 79-93): Springer.

- Larsen, R., Rich, J., & Rasmussen, T. K. (2019). Hub-based truck platooning: Potentials and profitability. *Transportation Research Part E: Logistics and Transportation Review*, 127, 249-264.
- Larson, J., Kammer, C., Liang, K.-Y., & Johansson, K. H. (2013). *Coordinated route optimization for heavy-duty vehicle platoons*. Paper presented at the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).
- Larson, J., Munson, T., & Sokolov, V. (2016). *Coordinated platoon routing in a metropolitan network*. Paper presented at the 2016 Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing.
- Larsson, E., Sennton, G., & Larson, J. (2015). The vehicle platooning problem: Computational complexity and heuristics. *Transportation Research Part C: Emerging Technologies*, 60, 258-277.
- Liang, K.-Y., Mårtensson, J., & Johansson, K. H. (2013). When is it fuel efficient for a heavy duty vehicle to catch up with a platoon? *IFAC Proceedings Volumes*, 46(21), 738-743.
- Liu, T., Ceder, A., & Rau, A. (2018). *Scheduled platoons of public transport autonomous modular vehicles*. Paper presented at the Proceedings of the 7th Symposium of the European Association for Research in Transportation (hEART2018), Athens, Greece.
- Lum, O., Chen, P., Wang, X., Golden, B., & Wasil, E. (2015). *A heuristic approach for the swap-body vehicle routing problem*. Paper presented at the 14th INFORMS computing society conference.
- Luo, F., & Larson, J. (2020). A Repeated Route-then-Schedule Approach to Coordinated Vehicle Platooning: Algorithms, Valid Inequalities and Computation. *arXiv preprint arXiv:2004.13758*.
- Luo, F., Larson, J., & Munson, T. (2018). Coordinated platooning with multiple speeds. *Transportation Research Part C: Emerging Technologies*, 90, 213-225.
- Mahmassani, H. S. (2016). 50th anniversary invited article—autonomous vehicles and connected vehicle systems: Flow and operations considerations. *Transportation Science*, 50(4), 1140-1162.
- Meisen, P., Seidl, T., & Henning, K. (2008). *A data-mining technique for the planning and organization of truck platoons*. Paper presented at the Proceedings of the International Conference on Heavy Vehicles.
- Mokhtar, H., Krishnamoorthy, M., Dayama, N. R., & Kumar, P. R. (2020). New approaches for solving the convoy movement problem. *Transportation Research Part E: Logistics and Transportation Review*, 133, 101802.
- Murray, D., & Glidewell, S. (2019). An Analysis of the Operational Costs of Trucking: 2019 Update. In: American Transportation Research Institute (ATRI), November.
- National Highway Traffic Safety Administration. (2016). Automated Vehicles for Safety. Retrieved from <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>
- Nourmohammadzadeh, A., & Hartmann, S. (2016). *The fuel-efficient platooning of heavy duty vehicles by mathematical programming and genetic algorithm*. Paper presented at the International Conference on Theory and Practice of Natural Computing.
- Nourmohammadzadeh, A., & Hartmann, S. (2019). Fuel-efficient truck platooning by a novel meta-heuristic inspired from ant colony optimisation. *Soft Computing*, 23(5), 1439-1452.
- Pironi, A., Scarpato, N., & Brillì, M. (2018). Industry 4.0 revolution in autonomous and connected vehicle a non-conventional approach to manage big data.
- Porter, B., Linse, M., & Barasz, Z. (2015). Six transportation trends that will change how we move. *Forbes*, January.
- Scherr, Y. O., Neumann-Saavedra, B. A., Hewitt, M., & Mattfeld, D. C. (2018). Service network design for same day delivery with mixed autonomous fleets. *Transportation research procedia*, 30, 23-32.
- Scherr, Y. O., Saavedra, B. A. N., Hewitt, M., & Mattfeld, D. C. (2019). Service network design with mixed autonomous fleets. *Transportation Research Part E: Logistics and Transportation Review*, 124, 40-55.
- Schiffer, M., Schneider, M., & Laporte, G. (2018). Designing sustainable mid-haul logistics networks with intra-route multi-resource facilities. *European journal of operational research*, 265(2), 517-532.
- Sebe, S.-M., Kraus, P., Müller, J. P., & Westphal, S. (2019). *Cross-provider Platoons for Same-day Delivery*. Paper presented at the VEHTS.
- Sebe, S.-M., & Müller, J. P. (2019). PFA: a Platoon Forming and Routing Algorithm for Same-Day Deliveries. *arXiv preprint arXiv:1912.08929*.
- Sethuraman, G., Liu, X., Bachmann, F. R., Xie, M., Ongel, A., & Busch, F. (2019). *Effects of bus platooning in an urban environment*. Paper presented at the 2019 IEEE Intelligent Transportation Systems Conference (ITSC).
- Sokolov, V., Larson, J., Munson, T., Auld, J., & Karbowski, D. (2017a). Maximization of platoon formation through centralized routing and departure time coordination. *Transportation Research Record*, 2667(1), 10-16.
- Sokolov, V., Larson, J., Munson, T., Auld, J., & Karbowski, D. (2017b). Platoon formation maximization through centralized routing and departure time coordination. *arXiv preprint arXiv:1701.01391*.
- Speranza, M. G. (2018). Trends in transportation and logistics. *European journal of operational research*, 264(3), 830-836.
- StadieSeifi, M., Dellaert, N. P., Nuijten, W., Van Woensel, T., & Raoufi, R. (2014). Multimodal freight transportation planning: A literature review. *European journal of operational research*, 233(1), 1-15.
- Sun, X., & Yin, Y. (2019). Behaviorally stable vehicle platooning for energy savings. *Transportation Research Part C: Emerging Technologies*, 99, 37-52.
- Talebpoor, A., & Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71, 143-163.
- The European Environment Agency. (2019). Greenhouse gas emissions from transport in Europe. Retrieved from [https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12#:~:text=EEA%20estimates%20show%20that%20emissions%20from%20transport%20\(including%20aviation\)%20increased,international%20aviation%20and%20international%20shipping](https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12#:~:text=EEA%20estimates%20show%20that%20emissions%20from%20transport%20(including%20aviation)%20increased,international%20aviation%20and%20international%20shipping).
- United States Environmental Protection Agency. (2020). Fast Facts on Transportation Greenhouse Gas Emissions. Retrieved from <https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions>
- US Department of Transportation. (2020). How Connected Vehicles Work. Retrieved from <https://www.transportation.gov/research-and-technology/how-connected-vehicles-work>
- Van de Hoef, S. (2016). *Fuel-efficient centralized coordination of truck platooning*. KTH Royal Institute of Technology.
- Van De Hoef, S., Johansson, K. H., & Dimarogonas, D. V. (2015). *Fuel-optimal centralized coordination of truck platooning based on shortest paths*. Paper presented at the 2015 American Control Conference (ACC).
- Wadud, Z., MacKenzie, D., & Leiby, P. (2016). Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice*, 86, 1-18.
- Wieberneit, N. (2008). Service network design for freight transportation: a review. *OR spectrum*, 30(1), 77-112.
- You, J., Miao, L., Zhang, C., & Xue, Z. (2020). A generic model for the local container drayage problem using the emerging truck platooning operation mode. *Transportation Research Part B: Methodological*, 133, 181-209.
- Zhang, L., Chen, F., Ma, X., & Pan, X. (2020). Fuel Economy in Truck Platooning: A Literature Overview and Directions for Future Research. *Journal of Advanced Transportation*, 2020.
- Zhang, W., Jenelius, E., & Ma, X. (2017). Freight transport platoon coordination and departure time scheduling under travel time uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 98, 1-23.

**Session 2.1 - Session spéciale 3 :
Surveillance intelligente des processus
et maintenance prédictive pour
l'industrie du futur**

Chairman : Christophe Simon - Magali Pralus

Des modèles d'apprentissage d'ensemble pour la prédiction de l'état des systèmes de production

AMAL GOUIAA¹, ROUA HAMMAMI²

¹ Laboratoire Pluridisciplinaire de Recherche en Ingénierie des Systèmes, Mécanique, Énergétique (PRISME)
8 Rue Léonard de Vinci, 45100 Orléans, France
amal.gouiaa-mtibaa@univ-orleans.fr

² BIWARE CONSULTING
1053, RUE DU LAC HURON, TUNIS, TUNISIE
rhammami@biware-consulting.com

Résumé – L'augmentation des données disponibles ainsi que leur hétérogénéité et leur évolution en temps réel dans presque tous les domaines soulève la nécessité d'utiliser des algorithmes pour l'analyse automatisée des données y compris le domaine de la maintenance (Maintenance prédictive). Notre objectif consiste à prédire l'état de la machine (En panne/fonctionnel) pour 3 horizons différents (24h, 48h et 72h), afin de planifier les actions de maintenance à l'avance. Dans cet article, nous avons étudié et comparé 4 différents modèles : Voting Ensemble, Stack Ensemble, Light Gradient Boosting Model et eXtreme Gradient Boost. L'analyse de la performance des modèles a été effectuée selon 2 métriques basées sur la matrice de confusion (Moyenne pondérée des scores de précision (MPSP), précision pondérée (PP)). La performance des modèles a dépassé dans certains cas les 90%.

Mots clés – Intelligence artificielle, maintenance prédictive, machine Learning, Méthodes d'ensemble, Industrie 4.0

Abstract – The increase of available data as well as its heterogeneity and real-time evolution in almost every domain raises the necessity to use algorithms for automated data analysis. This necessity is highlighted also in maintenance area (Predictive Maintenance). Our objective is to predict failures for 3 different horizons (24h, 48h et 72h) (in order to planify maintenance actions in advance). In this article, we have studied and compared 4 different models: Voting Ensemble, Stack Ensemble, Light Gradient Boosting Model and eXtreme Gradient Boost. Our objective is to predict the state of the machine (Down / functional) for 3 different horizons (24h, 48h and 72h). Model performance analysis was done using 2 metrics based on confusion Matrix (Weighted Average Precision Scores (MPSP), Weighted Precision (PP)). The performance of models exceeded in some cases 90%.

Keywords – artificial intelligence, predictive maintenance, machine learning, Ensemble method, Industry 4.0

1 INTRODUCTION ET ETAT DE L'ART

L'industrie du futur (Industrie 4.0), marquée principalement par les nouvelles technologies, notamment digitales et numériques, a modifié en profondeur le visage de l'industrie. Cette transformation impacte plusieurs niveaux de l'entreprise et implique la mise en place d'une nouvelle organisation du mode de production donnant une plus grande importance au réseau. Connectées entre elles, les machines d'une usine (ou de plusieurs sites) et des capteurs génèrent des informations de manière continue et instantanée. L'exploitation des données envoyées par ces différents outils et postes de travail intégrés dans les chaînes de fabrication et d'approvisionnement permet d'optimiser des process manufacturiers et d'améliorer la flexibilité afin de s'adapter à la demande en temps réel et de mieux satisfaire les besoins individuels de chaque client [Techniques de l'ingénieur, 2016].

L'évolution vers l'industrie 4.0 se fait à tous les niveaux, y compris la maintenance. Une activité cruciale dans la mesure où elle détermine la disponibilité et les performances des systèmes de production et la qualité des produits finis. En effet, la disponibilité optimale des machines et des équipements est une composante-clé de la bonne performance des industriels. Un arrêt dans la chaîne de production peut

impacter les performances et l'image de marque d'un industriel. Par conséquent, tous les métiers et secteurs d'activités sont désormais à la recherche d'amélioration de leurs processus pour gagner en agilité, en performance, en qualité de produit/service, et ce, tout en rationalisant leurs différents coûts.

De plus, la quantité de données existante actuellement a induit des changements fondamentaux dans le processus de prise de décision en entreprise pour plusieurs fonctions (maintenance, production, qualité, marketing...). C'est dans ce nouveau contexte que la notion de maintenance prédictive s'impose. En effet, traditionnellement, le plan des interventions de maintenance est basé principalement sur les recommandations du constructeur ou bien sur l'historique de l'exploitation du système de production (maintenance préventive).

Contrairement aux méthodes traditionnelles, la maintenance prédictive consiste à élaborer un plan de maintenance mais en se basant sur des données de la production, la qualité des produits finis et de données réelles collectées en continue en temps réelles à travers différents types de capteurs (acoustiques, thermiques, vibratoire...). Ces capteurs transmettent les données afin d'être explorées et analysées en se basant sur des outils puissants d'analyse de données et permettent de déterminer par exemple si une machine tombe

en panne ou pas (prédiction binaire), la probabilité d'un défaut sur une machine, le temps de fonctionnement restant et même le type de défaut possible afin d'anticiper une panne et prévoir l'intervention nécessaire.

Ces immenses quantités de données collectées et les progrès des performances de calcul des ordinateurs ont amené à la naissance des méthodes de fouille de données (Data Mining « DM » en anglais). Ces données collectées représentent une mine d'informations qui peut assister les utilisateurs pour la prise de décision. Autrement dit, ces outils ont comme objectif de donner un sens et à faire parler des données déjà disponibles dans les systèmes d'informations des entreprises afin de bien comprendre notre système et par conséquent prendre les bonnes décisions [Yahyaoui, 2015].

La science des données est un domaine très large incluant de domaines comme : l'analyse des données, la fouille des données, l'apprentissage automatique et plus généralement l'intelligence artificielle. L'application de ces axes permet l'extraction de la connaissance à partir des données. Il existe plusieurs types et techniques de fouilles de données et d'apprentissage : des techniques se basant sur un modèle unique : arbre de décision, réseau de neurone, K-Moyen, Random forest et d'autres techniques dites des méthodes d'ensemble comme : votingEnsemble, StackEnsemble, MaxAbsScaler LightGBM et StandardScalerWrapper, XGBoostClassifier. Les méthodes d'ensemble sont des techniques qui créent plusieurs modèles puis les combinent pour produire de meilleurs résultats. Les méthodes d'ensemble produisent généralement des solutions plus précises qu'un modèle unique [Marcello et al. 2020].

[Romanowski et Nagi, 1999] ont utilisé la méthode d'arbre de décision pour expliquer des pannes d'équipements à travers l'analyse des données de maintenance et en choisissant comme valeur cible, l'occurrence d'une panne sur un équipement ou pas.

[Susto et al. 2015] propose une méthodologie permettant d'adopter des règles de décision dynamiques pour la gestion de la maintenance pouvant être utilisée avec des problèmes de données de grande dimension et censurés. Ceci est réalisé en formant plusieurs modules de classification avec différents horizons de prédiction pour fournir différents compromis de performances en termes de fréquence des ruptures inattendues et de durée de vie inexploitée, puis en utilisant ces informations dans un système de décision de maintenance basé sur les coûts d'exploitation pour minimiser les coûts attendus. L'efficacité de la méthodologie est démontrée à l'aide d'un exemple simulé et d'un problème de référence de maintenance de fabrication de semi-conducteurs et évaluée en se basant sur les pannes non-prédites et les cycles de vies non-exploités.

Dans leur article, [Wu, 2017] prédit l'usure des outils lors des opérations de fraisage à partir des signaux de force de coupe, de vibration et d'émission acoustique en utilisant l'algorithme de forêt aléatoire (Random Forest (RF)). Les résultats expérimentaux ont montré que le modèle prédictif formé par les RF est très précis. L'erreur quadratique moyenne (MSE) sur les données d'usure de l'outil de test peut atteindre 7,67%. Le coefficient de détermination (R²) sur les données d'usure de l'outil de test peut atteindre 0,992. Selon les auteurs, l'algorithme de forêt aléatoire a été appliqué pour prédire l'usure des outils pour la première fois.

En deuxième temps, les performances des réseaux de neurone artificiels (ANN), des machines à vecteurs de support (SVM)

et de l'algorithme de forêt aléatoire (RF) sont comparées des points de vue de performance et de temps de calcul. Bien que la construction du modèle prédictif avec les soit la plus couteuse en temps de calcul, sa performance dépasse les modèles créés par les ANN et le SVM. Les RF sont alors les plus précis pour cet exemple d'application.

[Kroll et al., 2014] ont choisi de combiner deux modèles pour analyser et prédire le comportement d'un système complexe. Le premier modèle est basé sur une approche phénoménologique utilisant des algorithmes d'intelligence artificielle, classifiant le système à partir des mesures des capteurs. Le deuxième est un modèle de simulation du comportement du système pour estimer les mesures du système et le considère défaillant si la différence entre les résultats obtenus de la simulation et les valeurs réelles dépasse un seuil. La combinaison de ces deux modèles permet de générer un modèle hybride dont les prédictions sont plus précises.

Contrairement aux travaux susmentionnés qui se limitent à des approches supervisées, [Nagdev et Tarun, 2018] utilisent des algorithmes d'apprentissage non-supervisé pour classer le système en trois catégories ou clusters (Normal, Alarmant et défaillant) selon les mesures de vibration. Une réduction de dimension avec la méthode l'analyse des composantes principales PCA a précédé la création du modèle de clustering. Les algorithmes utilisés dans cette étude sont les K-moyennes (K-means), le Clustering Flou (Fuzzy C-means Clustering) et le Modèle de mélange gaussien (GMM).

En plus de ces modèles de base (modèle unique), d'autres travaux ont utilisé des méthodes dites d'ensemble. Les méthodes d'ensemble sont des techniques d'apprentissage automatique qui combinent plusieurs modèles de base afin de produire un modèle prédictif permettant d'avoir des meilleurs résultats [Marcello et al.2020].

[Marcello et al. 2020] ont présenté un modèle d'apprentissage d'ensemble pour estimer le taux de défaillance des équipements soumis à différentes conditions de fonctionnement. La méthode proposée permet aussi d'identifier les paramètres de fonctionnement les plus critiques et qui affectant ce taux de défaillance. Une étude d'un échantillon de 143 pompes centrifuges installées dans une raffinerie de pétrole a montré les potentialités et l'efficacité de la méthode proposée.

Plusieurs travaux ont été effectués et traitant cette thématique quel que soit de point de vue métier (maintenance) ou technique (Machine Learning).

Dans nos travaux, notre objectif est de prédire l'état d'une machine (en panne/fonctionnel) pour 3 horizons différents (24, 48 et 72 heures). Pour atteindre cet objectif, plusieurs modèles d'ensemble ont été utilisés. La performance de ces modèles sera comparée selon plusieurs indicateurs de performance (Moyenne pondérée des scores de précision (MPSP), précision pondérée (PP) et la matrice de confusion).

2 PRESENTATION DU PROJET

2.1 Acquisition et Exploration des données

Les données utilisées ont été récupérées sur la plate-forme Kaggle. Celles-ci sont relatives à 100 machines de 4 modèles

différents avec des âges différents allant de 0 (machine neuve) à 20 (20 ans d'utilisation).

Le jeu de donnée choisi est composé de 5 Data Tables

- Défauts : Enregistrements des défauts des machines sous forme de défauts de qualité.
- Défaillances : L'historique des pièces changées à la suite des pannes permettant d'extraire l'historique des pannes de façon générale et leur localisation de façon plus précise.
- Machines : Les caractéristiques des 100 machines ; leurs modèles et leurs âges.
- Maintenances : Les enregistrements des interventions de maintenance précisant la machine et le composant maintenu.
- Mesures : Les enregistrements des mesures des capteurs pour chaque machine : La tension, la rotation, la pression et la vibration.

Le set de données utilisé dans notre travail comporte 105132 observations et est le résultat des croisements des Data Tables « Défaillance » et « Mesure ». Dans nos travaux, nous nous concentrons nos études sur les machines de type 4. Ce choix a été fait en étudiant le nombre de panne par machine. La machine 4 est parmi les machines ayant le nombre de panne le plus élevé.

En effet, en croisant ces deux bases, on peut déterminer à quel instant la panne a eu lieu, sur quelle machine, ainsi que les enregistrements de capteur précédant la panne. Il est important de préciser que pour tous types de panne, la machine est à nouveau opérationnelle à 8h le lendemain matin.

Les données récupérées sont des données synthétiques. Néanmoins, certains traitements de nettoyage de données ont été réalisés à savoir l'imputation des valeurs manquantes, la correction des valeurs aberrantes et la normalisation des mesures améliorant la robustesse du modèle et réduisant sa complexité.

De plus, une étude de corrélation des mesures des capteurs a été réalisée vu que ces enregistrements représentent les entrées essentielles de notre modèle. Leur indépendance a été ainsi confirmée. Les résultats obtenus sont présentés dans le tableau suivant :

Tableau 1. Matrice de corrélation

	Tension	Rotation	Pression	Vibration
Tension	1	0.05	-0.01	- 0.05
Rotation	0.05	1	-0.03	0.04
Pression	0.01	- 0.03	1	0.001
Vibration	0.05	0.04	0.001	1

Composé de 3% de pannes, notre jeu de données est non-équilibré. Nous avons alors traité notre problématique autrement. Au lieu de répondre à la question « Y aura-t-il une panne dans le prochain cycle (1h) », nous avons orienté nos travaux pour apporter une réponse à une nouvelle question : « Y aura-t-il une panne dans l'horizon h, où h est soit 24h, 48h ou 72h. ». D'un point de vue technique, cet horizon nous permet d'enrichir notre base et de l'équilibrer étant donné que

nous aurons plus de réponses positives comparées aux réponses de la première question. D'autre part, d'un point de vue métier il permet d'anticiper les actions de maintenance.

Etant donné que notre objectif est de prédire une prochaine panne à partir des mesures des capteurs, nous fouillons alors ces données pour chercher les variations et les anomalies annonçant une défaillance. De ce fait nous considérerons comme entrée de notre modèle, la variation des mesures des capteurs dans l'horizon h procédant le point d'observation.

Afin d'évaluer efficacement le modèle, nous aurons recours à la validation croisée. Selon [Arlot et al. 2018] cette procédure est fréquemment utilisée en apprentissage supervisé et permet d'évaluer le risque des modèles et d'en choisir le plus optimal en se fondant sur le principe d'échantillonnage.

2.2 Algorithmes de machine learning utilisés pour la création du modèle de prédiction

2.2.1 Méthode d'ensemble basé sur le vote « VotingEnsemble »

Cette méthode est utilisée pour la classification. La première étape consiste à créer plusieurs modèles de classification à l'aide d'un ensemble de données d'entraînement. Chaque modèle de base peut être créé en utilisant différentes divisions du même jeu de données d'entraînement et du même algorithme, ou en utilisant le même jeu de données avec différents algorithmes, ou toute autre méthode. Il existe deux types de vote : Vote majoritaire et vote pondéré. Dans le cas du vote majoritaire, chaque modèle fait une prédiction (votes) pour chaque instance de test et la prédiction de sortie finale est celle qui reçoit plus de la moitié des votes. Si aucune des prédictions n'obtient plus de la moitié des votes, on peut dire que la méthode d'ensemble ne pourrait pas faire une prédiction stable pour cette instance. Dans ce cas, on peut essayer la prédiction la plus votée (même si cela représente moins de la moitié des votes) comme prédiction finale. Dans certains articles, cette méthode peut être appelée «vote à la pluralité». En ce qui concerne le vote pondéré et, contrairement au vote majoritaire, où chaque modèle a les mêmes droits, nous pouvons augmenter l'importance d'un ou plusieurs modèles. Dans le vote pondéré, on peut compter plusieurs fois la prédiction des meilleurs modèles. Il nous appartient de trouver un ensemble de poids raisonnable.

2.2.2 Empilement de plusieurs modèles d'apprentissage automatique « StackEnsemble »

Est une méthode d'ensemble efficace dans laquelle les prédictions, générées à l'aide de divers algorithmes d'apprentissage automatique, sont utilisées comme entrées dans un algorithme d'apprentissage dit de deuxième couche. Cet algorithme de deuxième couche est formé pour combiner de manière optimale les prédictions du modèle pour former un nouvel ensemble de prédictions. Par exemple, lorsque la régression linéaire est utilisée comme modélisation de deuxième couche, elle estime ces poids en minimisant les erreurs des moindres carrés. Cependant, la modélisation de deuxième couche ne se limite pas aux seuls modèles linéaires ; la relation entre les prédicteurs peut être plus complexe, ouvrant la porte à l'utilisation d'autres algorithmes d'apprentissage automatique.

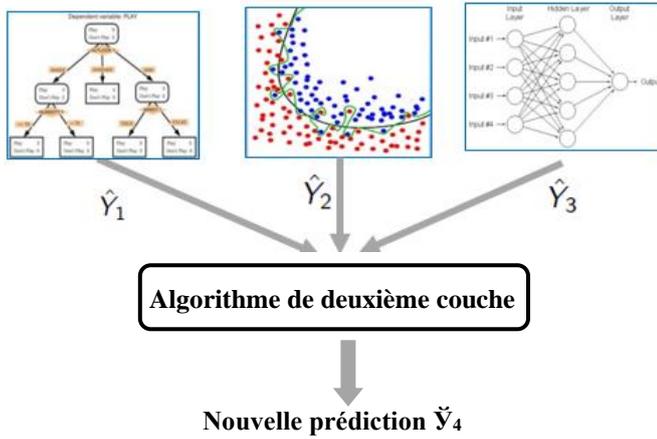


Figure 3. Algorithme StackEnsemble

2.2.3 Light Gradient Boosting Model

[Ke et al. 2017] ont présenté le Light GBM (Light gradient boosting model) comme une version améliorée de l'algorithme GBDT (Gradient boosting decision tree). Le LightGBM est un algorithme d'apprentissage automatique construit à partir de l'algorithme arbre de décision entraîné séquentiellement. A chaque itération, il choisit la feuille avec une perte de delta maximale pour se développer.

Il construit alors l'arbre verticalement donc par feuille et non pas par niveau. En effet, Lors de la croissance de la même feuille, l'algorithme par feuille peut réduire plus de perte qu'un algorithme par niveau.

Cet algorithme est connu non seulement pour sa performance mais également pour son optimisation de calcul surtout pour les bases de grande taille.

La figure suivante explique la mise en œuvre de cet algorithme

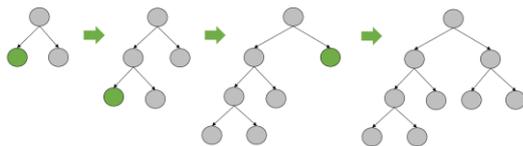


Figure 2. Algorithme Light Gradient Boosting

2.2.4 XGBoostClassifier

Proposé par [Chen et Guestrin. 2016], eXtreme Gradient Boost, XGBoost, est un algorithme d'apprentissage automatique basé également sur l'algorithme d'arbre de décision. Il s'agit également d'une version améliorée de l'algorithme GBDT (Gradient Boosting Decision Tree). En effet, Il rajoute l'élagage (tree puning) et la régularisation pour éviter le surapprentissage. De plus, l'XGboost est capable de traiter les valeurs nulles de lancer le calcul en parallèle.

Dans nos tests nous avons considéré le MinMaxScaler comme fonction de régularisation.

2.3 Mesure de performances

Afin de quantifier la qualité des classificateurs et comparer leurs performances, le score de précision ou la justesse est généralement utilisée. Pour notre problématique, étant donné que notre base reste, malgré les améliorations, non-équilibrée, ce score n'est pas significatif. En effet, puisqu'il est donné par le pourcentage des bonnes prédictions, il suffit que le modèle prédise toutes les observations comme la classe la plus présente pour avoir un bon score de précision.

Nous utiliserons alors comme critère d'évaluation des métriques plus adaptées permettant de prendre en considération le déséquilibre de la base de données à savoir la moyenne pondérée des scores de précision (MPSP) et le score de précision pondérée (PP). Nous accordons également une grande importance à la matrice de confusion pour analyser la performance.

2.3.1. Moyenne pondérée des scores de précision

La moyenne pondérée des scores de précision, comme son nom l'indique, est une moyenne des scores de justesse de chaque classe pondérée par le taux de présence de celle-ci. La prise en compte du taux de présence dans le calcul de cet indicateur permet d'ajuster le déséquilibre de la base de données. Cette moyenne est donnée par la formule suivante (1) :

$$\sum T_n \times P_n \quad (1)$$

Avec

T_n : Le taux de présence de la classe n

P_n : Le score de précision de la classe n

2.3.2. Score de précision pondéré :

Le score de précision pondéré permet d'évaluer la performance d'un modèle de classification non équilibré. Il est donné par la formule suivante :

$$\text{Précision pondérée} = \frac{\text{sensitivité} + \text{spécificité}}{2} \quad (2)$$

Avec

$$\text{sensitivité} = \frac{VP}{VP + FN} \quad (3)$$

Et

$$\text{Spécificité} = \frac{VN}{FP + VN} \quad (4)$$

Où

Pour une classification binaire :

VP : nombre de prédiction Vrai Positive

VN : nombre de prédiction Vrai Négative

FP : nombre de prédiction Faux Positive

FN : nombre de prédiction Faux Négative

2.3.3 Matrice de confusion :

Pour analyser la performance des différents modèles, nous nous basons aussi sur la matrice de confusion. Cette matrice est illustrée dans le tableau suivant :

Tableau 2. Matrice de confusion

	Positive	Négative
Prédiction Positive	Vrai Positive VP	Faux Positive FP
Prédiction Négative	Faux Négative FN	Vrai Négative VN

prédiction de panne est plus critique que celle du bon fonctionnement, nous donnerons plus d'importance à la précision pondérée. Nous sélectionnerons alors les modèles suivants par horizon :

Tableau 4. Les modèles sélectionnés

Horizon	24h	48h	72h
Modèle	Light GBM	XGBoost	Voting Ensemble

Nous pouvons également conclure que les horizons 48h et 72h ont de meilleurs résultats. En effet, étant donné que nous prédisons les pannes à partir des variations des mesures des capteurs, augmenter relativement l'horizon permet de détecter de mieux détecter les anomalies.

Le choix de l'horizon dépend également du besoin de l'entreprise et des délais nécessaire pour l'intervention.

$$\text{Taux de Réussite} = \frac{VP + VN}{\text{Total d'Observation}} \quad (5)$$

3 RESULTAT

Les résultats de nos travaux sont résumés dans le tableau suivant (Tableau 3):

Tableau 3. Performance des modèles

	24h		48h		72h	
	MPSP (%)	PP (%)	MPSP (%)	PP (%)	MPSP (%)	PP (%)
Voting Ensemble	92,0	52,9	84,4	96,9	79,6	92,1
Stack Ensemble	90,9	55,0	83,6	95,7	79,4	91,5
Light GBM	91,4	54,6	84,4	96,6	78,4	92,7
XGBoost	93,3	51,1	83,8	97,0	79,9	91,6

En plus des mesures de performances, une étude d'importance a été menée afin d'évaluer le poids de chaque entrée dans le modèle. Il est important de rappeler que l'algorithme d'arbre de décision et tous algorithmes dérivés ont la capacité de sélectionner les entrées qui expliquent le mieux le résultat. Cette caractéristique, ajoute de l'importance et de l'attractivité à ces modèles.

Notre étude a permis de confirmer que, pour les trois horizons choisis, l'importance des mesures des capteurs est très peu significative comparée à celle des variations de ces mesures. Ce résultat permet d'approuver le résonnement effectué au niveau de la création de la base de données. Nous rappelons que nous avons calculé les variations des mesures des capteurs sur les trois horizons considérés et utilisé les champs calculés comme entrées de nos modèles.

Afin de pouvoir sélectionner le meilleur modèle par horizon, il faudrait analyser les deux critères tout en tenant compte de la matrice de confusion. Vu que pour notre problématique, la

4 CONCLUSION ET PERSPECTIVES

Dans ce papier, nous avons étudié et comparé 4 différents modèles : Voting Ensemble, Stack Ensemble, Light Gradient Boosting Model et eXtreme Gradient Boost, pour 3 différents horizons 24h, 48h et 72h. Notre objectif est de prédire la probabilité de panne durant l'horizon donné et donc de décrire l'état du système (fonctionnel/En panne). Pour ce faire, nous avons considéré les variations des mesures des capteurs pour comme entrées pour créer et entraîner nos modèles. Dans ce cas de classification binaire, les résultats obtenus ont montré que les performances des algorithmes susmentionnés sont comparables. Nous avons effectué notre choix en analysant des métriques adaptées au problème de données non-équilibrées ; la moyenne pondérée des scores de précision et la précision pondérée. Nous avons également pris en considération la matrice de confusion pour appuyer notre analyse des métriques. Ce premier résultat présente un élément de départ pour un travail de recherche plus enrichi.

Il est fortement constaté en industrie qu'il existe une relation fondamentale entre la dégradation avancée des systèmes de production et la qualité des produits finis. En effet, l'une des causes possibles de la perte de qualité réside, souvent, à l'état de dégradation du système de production. Le set de données utilisé dans notre travail est le résultat des croisements des 2 Data Tables: Défauts et Mesures ce qui nous a permis de déterminer à quel instant la panne a eu lieu pour chaque machine. Le jeu de donnée choisi contient à la base 5 Data Tables décrivant les caractéristiques des machines, les composants défaillants, les actions de maintenance réalisées en plus des mesures des capteurs et de l'état des machines. Notre objectif final serait d'inclure ces données afin d'enrichir notre base. Ainsi nous prendrons en considération les améliorations de performance du aux interventions ainsi que les caractéristiques de chaque machine afin de prédire avec précision le composant défaillant.

5 REFERENCES

- Agard. B, Kusiak. A (2005). Exploration des bases de données industrielles à l'aide du data mining – Perspectives.
- Cornuéjols. A, Miclet. L et Haton. J. P (2010). Apprentissage Artificiel : Concepts et algorithmes (2^{ème} édition).
- Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785–794 (ACM, 2016).
- Guolin. K, Meng. Q, Finley. T, Wang. T, Chen. W, Weidong Ma. W, Q. Ye, T. Liu. T.Y, LightGBM: A Highly Efficient Gradient Boosting Decision Tree , NIPS 2017
- Hand. D, Mannila H et Smyth. P (2001). Principles of Data Mining, Massachusetts Institute of Technology.
- Kroll. B, Schaffranek. D, Schriegel. S et Nigge-mann. O (2014). System modeling based on machine learning for anomaly detection and predictive maintenance in industrial plants," Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), Barcelona, pp. 1-7, doi: 10.1109/ETFA.2014.7005202.
- L'usine 4.0, c'est quoi ? Techniques de l'ingénieur (2016)
- Marcello. B, Castellano. D, Frosolini. M, Gabbrielli. R, Marrazzini. L, Padellini. L (2020). An ensemble-learning model for failure rate prediction, Procedia Manufacturing, 42, 41-48.
- Nagdev. A, Tarun. G. (2018). A Research Study on Unsupervised Machine Learning Algorithms for Fault Detection in Predictive Maintenance. 10.13140/RG.2.2.28822.24648.
- Paolanti. M, Romeo. L, Loncarski. J (2018)., Machine Learning approach for Predictive Maintenance in Industry 4.0, 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA).
- Romanowski. C. J et Nagi. R (1999). Improving preventive maintenance scheduling using data mining techniques, chez Industrial Engineering Research Conference, Phoenix AZ.
- Susto. G.A, Schirru. A, Pampuri. S, McLoone., Beghi. S. A. (2015). Machine Learning for Predictive Maintenance: A Multiple Classifiers Approach. IEEE Transactions on Industrial Informatics, 11(3), 812-820. 10.1109/TII.2014.2349359
- Sylvain. A. Validation croisée. Myriam Maumy-Bertrand; Gilbert Saporta; Christine Thomas-Agnan. Apprentissage statistique et données massives, Editions Technip, 2018, 9782710811824. (hal-01485508)
- Wu. D, Jennings. C, Terpenney. J, Gao. R. X, Kumara. S (2017). A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests. Journal of Manufacturing Science and Engineering, Vol. 139
- Yahyaoui. H (2015). Méthode d'analyse de données pour le diagnostic a posteriori de défauts de production - Application au secteur de la microélectronique. Autre. Ecole Nationale Supérieure des Mines de Saint-Etienne. Français. (NNT : 2015EMSE0795). (tel-01282255)

Extraction et Exploitation de Règles Interprétables de Maintenance Cas d'un Four de Brasage

EMNA HAFDHHELLAOUI¹, MOHAMED-SALAH OUALI², AMEL JAOUA¹

¹ LR-OASIS, École Nationale d'Ingénieurs de Tunis,
Université El Manar, Tunis, Tunisie.

emna.hafdhellaoui@etudiant-enit.utm.tn; amel.jaoua@enit.utm.tn

² Département de mathématiques et de génie industriel, Polytechnique Montréal

BP6079, Succursale Centre-Ville, Montréal (Québec), Canada.

Membre du CIRRELT (cirrelt.ca) et IVADO (ivado.ca)

mohamed-salah.ouali@polymtl.ca

Résumé – Cet article propose une approche d'exploration des données de production et d'historique de maintenance afin de découvrir des règles interprétables de maintenance. L'approche tire profit des méthodes de classification par apprentissage automatique issue de l'intelligence artificielle. Dans cette approche, les données de production sont combinées avec les événements de maintenance de façon chronologique pour une meilleure représentabilité de l'état d'un système en exploitation au cours du temps. Les règles de maintenance sont extraites par type d'interventions et sont caractérisées, à la fois, par les variables de production et celles de maintenance. Les règles sont validées par la mesure de qualité du classement et l'avis de l'équipe de maintenance.

Abstract – This article presents a data-driven approach which combines production data and maintenance history to discover interpretable rules of maintenance. The approach takes advantage of machine learning classification methods derived from artificial intelligence. In this approach, production data and maintenance events are merged in a timestamp logs for better representability of the operating system state over time. The maintenance rules are extracted by intervention types and are characterized by both production and maintenance variables. The rules are validated by the ranking quality measurement and the maintenance team opinion.

Mots clés - Découverte de connaissances, Maintenance intelligente, Apprentissage automatique, Règles de décision.

Keywords – Knowledge discovery, Intelligent maintenance, Machine Learning, Decision rules.

1 INTRODUCTION

Durant les dernières années, l'exploration de données d'exploitation à l'aide des techniques d'apprentissage automatique (Machine Learning ML) pour améliorer la maintenance des systèmes de production est devenue très attrayante. Ce succès est principalement attribuable à la capacité de ces techniques de ML à traiter des données volumineuses provenant de diverses sources et de prédire l'état d'un système avec une certaine précision (Carvalho et al., 2019 ; Accorsi et al., 2017). Mathew et al. (2017) ont mené une étude comparative de différentes techniques de ML, telles que Support Vector Machine (SVM), Random Forest (RF) et k-Nearest Neighbours (KNN) pour prédire la durée de vie utile restante d'un moteur turboventilateur. Lesobre (2015) utilise une approche de type « Model-Based » exploitant les données de défaillance et d'intervention de maintenance. De manière générale, les interventions de maintenance sont planifiées sur la base de règles de décision faisant intervenir principalement l'âge. Ces règles préconisent des périodicités d'intervention sur les composants d'usure selon un usage nominale prévu par les constructeurs des machines. Elles sont ensuite adaptées par les gestionnaires de maintenance selon l'usage qu'ils font de leurs machines. Il en reste que ces règles ne font pas intervenir explicitement ni l'historique de performance de la production, ni la qualité des produits, ni les données environnementales. Avec de telles règles, les gestionnaires ne sont pas en mesure d'adapter la périodicité des interventions de maintenance d'une manière dynamique qui réponde à des besoins d'usage spécifiques des machines.

Le présent travail est réalisé dans le cadre d'un projet d'innovation, en Industrie 4.0, « KTA 8320 » lancé par

LOGIDAS Éditeur et Intégrateur de solutions ERP en Tunisie. Spécifiquement, LOGIDAS focalise son développement sur la maintenance 4.0 de machines pilotées par un système d'exécution de la production implanté chez Karmex-Tunisie, une entreprise de fabrication de pièces automobiles à haut standard de qualité destinées à l'exportation. Le travail concerne l'implantation d'une stratégie de maintenance intelligente, alignée sur les principes de la digitalisation des processus de maintenance, qui prend en compte l'analyse avancée des données de production et de maintenance pour la détection précoce des éventuelles défaillances des machines critiques. Dans ce contexte, nous profitons des nouvelles techniques de ML pour l'identification des règles de décision et leurs périmètres de validité pour mieux comprendre le comportement des modes de défaillance dominants et mieux planifier les interventions de maintenance. Plus particulièrement, il s'agit de développer une méthode permettant de prédire l'état d'une machine critique à l'aide des règles d'intervention extraites des données de production et de maintenance. Pour y arriver, une démarche systématique débutant par la préparation des données de qualité est réalisée. Ensuite vient la caractérisation des différents états de la machine à l'aide des variables de conception, de production et de maintenance sous la forme de modèles de classification interprétable et leur validation par les mesures appropriées de qualité. En troisième lieu vient l'implémentation et le test des modèles obtenus. Il s'agit particulièrement d'analyser les résultats obtenus, de les valider avec les experts de maintenance et d'émettre des recommandations pratiques pour améliorer la performance globale de la machine. La méthode est appliquée sur un four à gaz. Les données utilisées sont composées de l'historique

d'exploitation sur une période de 3 ans. Spécifiquement, les caractéristiques des produits fabriqués, les relevés de défaillances survenues et des interventions de maintenance pendant la même période sont considérés.

Le reste de l'article est organisé en quatre sections. La Section 2 propose une revue de littérature des approches de modélisation de l'état d'un système, les techniques de classification automatique supervisée et les métriques d'évaluation de la qualité de la classification. La Section 3 présente le cas d'étude ainsi que la préparation des données de modélisation et les tests préliminaires. La Section 4 extrait les règles d'intervention selon une démarche systématique à l'aide des techniques de ML, valide les modèles, interprète les règles obtenues et recommande des actions spécifiques pour une meilleure implantation des résultats. La Section 5 conclue l'étude et présente quelques perspectives de développement.

2 REVUE DE LITTÉRATURE

La revue de littérature présente et discute les approches de modélisation existantes en lien avec l'évaluation de performance d'un système et les interventions de maintenance et les méthodes de classification automatique supervisée dans le contexte de la maintenance ainsi que les principales métriques de validation des modèles de classification obtenues.

2.1 Approches de modélisation

Basé sur les données d'exploitation d'un système, il existe principalement deux approches de modélisation de son état de performance : utilisant des modèles prédéfinis (statistiques ou physiques) (model-based) et basées sur l'exploration des données (data-driven).

- *Approches basées sur des modèles prédéfinis*

Alaswad et Xiang (2017) proposent un modèle de régression qui permet de caractériser chaque processus de détérioration sous-jacents (discrets et continus) en fonction des données de maintenance conditionnelle. La méthode prend en compte plusieurs facteurs tels que la fréquence, le type et l'efficacité des inspections prédictives pour détecter les mécanismes de détérioration. Park et al. (2016) utilisent les paramètres du procédé de moulage par injection pour évaluer l'état des systèmes critiques et établir les règles d'inspection selon une stratégie de maintenance conditionnelle. Des modèles statistiques de régression ont permis d'identifier les paramètres ayant le plus d'influence sur la qualité des produits et sur l'occurrence des défauts du procédé. Harrou et al. (2018) développent une méthode de détection et de diagnostic de défauts de panneaux photovoltaïques (PV). La méthode exploite les valeurs des résidus de courant, de tension et de puissance du réseau de sortie du PV pour détecter et isoler les défauts de court-circuit, circuit ouvert et ombrage. Les résultats montrent la capacité de la méthode à surveiller le fonctionnement du PV et à détecter même l'ombrage partiel. Aivaliotis et al. (2017) présentent une stratégie de maintenance prédictive d'un système industriel en utilisant les techniques de pronostic et de gestion de la santé des équipements (PHM). Basée sur des modèles de simulation physiques, une méthode de calcul de la durée de vie résiduelle (RUL) est proposée. La prédiction du RUL permet de sélectionner les interventions pertinentes de maintenance.

Wang et al. (2020) présentent une méthode intégrée de diagnostic et de pronostic de défaillances d'éoliennes, en particulier, dans les cas où les données de dégradation sont limitées. Un indice d'état de santé est déterminé à partir de compagnes d'analyse vibratoire et les avis d'experts. La méthode permet de prédire le RUL à l'aide de filtres de fréquences de manière récursive. Cauchi et al. (2017) présentent une méthode pour quantifier l'inconfort causé par la régulation

non optimale de la température dans un bâtiment en raison d'un fonctionnement dégradé de la chaudière. Cet inconfort est incorporé dans un algorithme de programmation dynamique qui détermine la fréquence optimale d'entretien de la chaudière. La stratégie d'entretien minimise les coûts totaux d'opération sur le long terme.

- *Approches basées sur l'exploration des données*

La littérature distingue les approches d'exploration supervisée et non supervisée des données. L'exploration non supervisée permet de regrouper des observations selon une mesure de la distance (Bumblauskas et al. (2017)). Alors que l'exploration supervisée permet de caractériser un regroupement d'un autre. Les méthodes supervisées peuvent faire de la classification ou de la régression et ce dépendamment de la variable d'état (discrète ou continue) utilisée dans la labélisation des observations collectées (Lade et al., 2017). Dans ce contexte, les travaux suivants ont été recensés.

Zenisek et al. (2019) proposent une méthode de détection et de prévision des dérives d'un procédé par apprentissage automatique non supervisé. La méthode utilise la régression multivariée de séries chronologiques comme moyen de modélisation de l'évolution des variables d'état du procédé au cours du temps. La précision des modèles fait en sorte que la méthode est capable de détecter le début de la séquence de la dérive du procédé et de prédire avec une certaine incertitude sa tendance. Amruthnath et Gupta (2018) présentent une méthode permettant la détection de différents niveaux de défauts de fonctionnement à partir de quantité limitée de données. La méthode teste plusieurs algorithmes de classification non supervisé (Principle component analysis PCA, regroupement hiérarchique, K-Means et clustering Fuzzy C-Means) sur un échantillon limité de données de vibrations relevées sur un ventilateur d'échappement. L'approche est adéquate pour des équipements observés sur une courte période de temps. Jiménez et al. (2020) développent une méthode de classification non supervisée des défauts de délaminage observés sur les pales d'éoliennes. La méthode propose un algorithme de classification autorégressive non linéaire qui détermine les caractéristiques de chaque regroupement de défauts de délaminage d'aubes d'éolienne. Une fois ces caractéristiques déterminées, six méthodes de classification supervisées (Decision Trees, Discriminant Analysis, Quadratic Support Vector Machines, Nearest Neighbours et Ensemble Classification) sont utilisées pour tester la précision de classification de chacun des défauts de délaminage.

Par ailleurs, il existe d'autres travaux qui utilisent de l'apprentissage supervisé. Citons à titre d'exemples les travaux de Calabrese et al. (2020) qui proposent une méthode de classification supervisée appliquée à une ligne de sciage du bois. Les probabilités de défaillance ainsi que le RUL des machines de la ligne sont estimées à l'aide d'une variété d'arbres de classification de forêts aléatoires (Gradient Boosting, Random Forest and Extreme Gradient Boosting). Les résultats sont utilisés dans l'amélioration des pratiques de maintenance déployées sur la ligne. Ferreira et al. (2017) ont décrit la mise en œuvre en cours d'un pilote de maintenance proactive dans l'industrie 4.0 pour une presse plieuse, dans le cadre du projet MANTIS. Le projet MANTIS vise la réduction des coûts de maintenance grâce à des techniques de surveillance basées sur l'apprentissage automatique supervisé. Les données de la machine sont collectées au moyen de capteurs qui font partie des mécanismes de contrôle de la machine ou de capteurs qui ont été ajoutés spécifiquement à des fins de maintenance.

Les approches précédentes présentent des avantages et des inconvénients lorsqu'elles sont appliquées à une problématique

de modélisation de l'état de dégradation d'un système. Dragomir *et al.* (2009) ont démontré que l'approche basée sur un modèle fournit un aperçu des facteurs qui affectent le plus le processus de détérioration mais ce modèle n'évolue pas au rythme du système. Les résultats sont limités par la précision des modèles, les incertitudes sur les paramètres et par la puissance de l'outil de simulation. Les systèmes étudiés sont complexes et non linéaires, les phénomènes de dégradation sont en partie aléatoires et les modèles ne sont pas toujours généralisables. En contrepartie, comme indiqué dans (Bjaili et Rushdi 2020 ; Hasan *et al.*, 2015) l'approche basée sur l'intelligence artificielle peut modéliser des relations inconnues, complexes et non linéaires entre les entrées et les sorties. Elle peut être utilisée lorsque les données sont imprécises, incomplètes et subjectives. Pour cette fin (Falarzi *et al.*, 2018) précisent qu'il est nécessaire de sélectionner un échantillon d'apprentissage pertinent et choisir correctement les paramètres à surveiller.

À travers cette brève revue, l'approche basée sur l'exploration des données paraît plus avantageuse et répondra mieux au besoin de l'entreprise partenaire de cette étude. En effet, les méthodes basées sur des modèles prédéfinis sont très sensibles aux bruits dans les données et ne permettent pas une interprétation directe des probabilités. C'est pourquoi nous avons décidé d'utiliser l'apprentissage automatique pour traiter les données imprécises et incomplètes de l'entreprise dans le but de développer des règles d'intervention efficaces et interprétables.

2.2 Méthodes utilisées pour l'extraction des règles de maintenance

La Figure 1 présente la démarche générale pour la mise en œuvre d'un processus de classification utilisant les méthodes d'apprentissage automatique. La démarche consiste à : identifier le problème, collecter et préparer les données, choisir la méthode appropriée au problème étudié, entraîner la méthode sur un échantillon représentatif et évaluer les résultats.

La classification est généralement utilisée pour les données de maintenance qui peuvent être catégorisées ou libellées. La maintenance prédictive a une classification multi classes car il existe plusieurs catégories ou raisons pour lesquelles un équipement échouera. Les algorithmes couramment utilisés dans les techniques de classification comprennent les réseaux de neurones artificiels (RNA), la régression logistique (RL), l'arbre de régression et classification (Classification and Regression Tree CART), l'arbre de décision (AD), l'arbre C5.0 et les forêts aléatoires (Random Forests RF) ; (Bukhsh *et al.*, 2019)

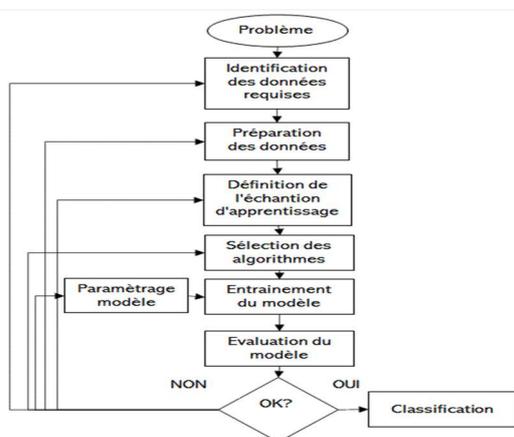


Figure 1. Démarche de classification (Osisanwo *et al.*, 2017)

Plus particulièrement, la méthode CART est non paramétrique. Elle divise les ensembles de données en groupes distincts. Il s'agit d'une procédure de partitionnement récursif binaire capable de traiter les attributs continus et nominaux comme cibles et prédicteurs. Les données sont traitées dans leur forme brute, aucun regroupement n'est requis ou recommandé (Steinberg et Colla, 2009). Les modèles CART présentent plusieurs avantages supplémentaires par rapport aux autres techniques, tels que par exemples les données d'entrée n'ont pas besoin d'être distribuées normalement ; il n'est pas nécessaire que les variables soient indépendantes ; les relations non linéaires entre les variables prédictives et les données observées peuvent être modélisées (Bel *et al.*, 2009). Les RF sont une méthode ensembliste fondée sur la génération aléatoire d'arbres. La méthode a été élaborée pour combler les lacunes de la méthode traditionnelle de l'arbre de classification et de régression (CART). Elle se compose d'un grand nombre d'apprenants de l'arbre de décision faible, qui ont été construit en parallèle pour réduire le biais et la variance du modèle en même temps (Ahmad *et al.*, 2018). Les RF sont de bonnes prédictions. Toutefois, l'interprétation est un peu difficile du fait que les variables peuvent apparaître à plusieurs niveaux de l'arbre final de prédiction. La méthode C5.0 possède de nombreuses fonctionnalités telles que la facilité d'interprétation des règles de classification, le traitement du bruit et des données manquantes, le sur-apprentissage et l'étalement par erreur résolu et l'effet explicite des variables sur la classification (Ortiz, 2019).

Pour l'évaluation de la qualité du classement on s'est basé sur les travaux de (Bukhsh *et al.*, 2019). Spécifiquement, cette évaluation se fait à l'aide de 5 critères principaux : la justesse (J), le F1_Score, le Kappa, Biais et la variance (Kautz *et al.*, 2017). La justesse désigne la proportion des prédictions correctes effectuées par le modèle de classification. Elle est définie par la proportion du nombre d'observations correctement prédites par rapport au total des observations prédites (Équation 1).

$$J = \frac{\text{Nombre de prédictions correctes}}{\text{Nombre total de prédictions}} \quad (1)$$

Le F1-Score est une mesure qui combine la précision du classement et sa sensibilité (Équation 2). Ce critère est compris entre 0 et 1, plus il est proche de 1 plus le modèle est performant (Kautz *et al.*, 2017).

$$F1_Score = \frac{2 \cdot \text{Prévision} \cdot \text{Rappel}}{\text{Prévision} + \text{Rappel}} \quad (2)$$

Le Kappa de Cohen est un critère toujours inférieur ou égal à 1. Les valeurs de 0 ou moins indiquent que le classificateur est inutile. Il n'y a pas de manière standardisée d'interpréter les valeurs prises par ce critère. Toutefois, selon Landis et Koch (1977) qui fournissent un moyen de caractériser les valeurs du Kappa, une valeur inférieure à 0 indique l'absence de concordance, entre 0 et 20% comme faible, entre 21 et 40% comme juste, entre 41 et 60% comme modéré, entre 61 et 80% comme substantiel (abondant) et enfin entre 81 et 100% comme accord presque parfait. Le Biais est l'erreur provenant d'hypothèses erronées dans l'algorithme d'apprentissage. Un biais élevé peut être lié à un algorithme qui omet certaines relations pertinentes entre les données d'entrée et la sortie prévue (sous-apprentissage). La Variance est l'erreur due à la sensibilité aux petites fluctuations de l'échantillon d'apprentissage. Une variance élevée peut entraîner un sur-apprentissage. Généralement, l'entraînement des modèles se fait avec 80% des données disponibles et le reste est gardé pour la validation croisée. C'est ce que nous avons appliqué pour tous les modèles proposés.

3 PRÉSENTATION DU CAS D'ÉTUDE ET TRAITEMENT DES DONNÉES

Avec l'aide du personnel de maintenance chez Karmex, le four de brasage a été sélectionné comme machine pilote. Ce four est considéré comme assez problématique. Il fonctionne 24 heures par jour, 5 jours par semaine. Sa capacité est limitée. Une quelconque défaillance ou arrêt non programmé engendre des coûts de rebuts très élevés. Le four FCN E20.35.75/ 225 GX, de hauteur 2,2 m et de longueur 22 m, est conçu pour effectuer le traitement de soudo-brasage et le recuit des pièces en acier au carbone. Le four est présenté dans la Figure 2.



Figure 2 – Four de soudo-brasage

Le four fait l'objet de 4 types d'intervention de maintenance :

- Type 1 (T1) : Maintenance de 1^{er} niveau. Elle concerne des tâches régulières de nettoyage, de lubrification, etc.
- Type 2 (T2) : Maintenance curative réalisée suite à une panne. Elle comporte le diagnostic, la réparation et le test. À noter que ce type d'intervention est considéré comme une réparation minimale qui n'affecte pas l'âge global du four.
- Type 3 (T3) : Maintenance de mise à niveau qui compense quelque part une certaine dégradation.
- Type 4 (T4) : Maintenance systématique qui s'attarde sur les remplacements préventifs de composants d'usure.

L'analyse préliminaire des données montrent des durées d'arrêts assez élevées lors des interventions curatives (T2) engendrant des pertes en disponibilité et en qualité. Pour intervenir sur ces pannes, les règles de décision sont basées seulement sur des variables issues de l'historique de maintenance, des documents du constructeur et du savoir-faire de l'équipe de maintenance. Alors que d'autres variables peuvent influencer la décision d'intervention et le comment en particulier des variables de production (atelier en job shop). Afin d'arriver à une planification efficace, il est important d'envisager des liens de causalité clairs entre variables caractérisant la situation du four et la décision prise. L'objectif est de rendre la maintenance dépendante des variables autant issues de l'historique de maintenance que de l'exploitation pour mieux gérer la maintenance du four. Ainsi, les nouvelles règles de décision vont dépendre à la fois, des variables de production et des variables de maintenance. Il sera possible d'identifier toutes les variables importantes de production et de maintenance qui influent le fonctionnement du four, et ensuite d'extraire des règles de décision plus explicites.

L'échantillon des données compte 20164 observations enregistrée entre 2017 et 2020 avec une trentaine de variables réparties entre des variables de maintenance telles que l'âge du

four, le type d'intervention, l'endroit d'intervention (localisation) dans le four, la durée d'intervention et le diagnostic (sous forme de texte), et les variables de production telles que la référence de l'article fabriqué, le code d'arrêt, le nombre de lots fabriqués, etc.

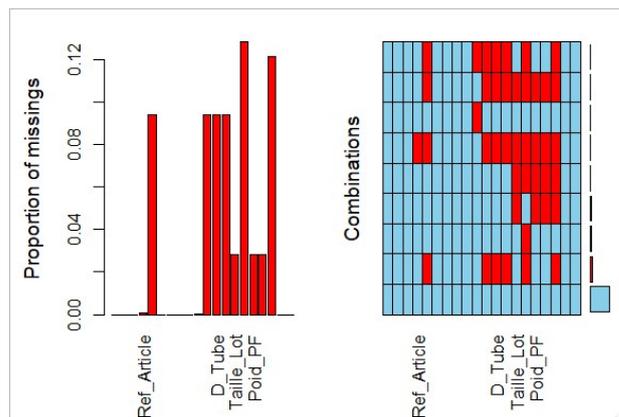


Figure 3 – Visualisation des valeurs manquantes

Le fichier des données collectées comporte 15.26% de valeurs manquantes (Figure 3) et plusieurs valeurs aberrantes. Comme ce pourcentage dépasse 5%, il est préférable de procéder avec une imputation multiple pour les traiter. Pour y arriver, la méthode CART a été utilisée puisqu'elle peut traiter les variables quantitatives et qualitatives. La méthode des RF n'était pas adéquate puisque certaines variables qualitatives ont plus de 250 modalités (types de produits). En appliquant l'imputation multivariée avec la méthode « CART » du package R « mice », nous avons réussi à obtenir un échantillon de données exempt de valeurs manquantes.

Concernant les valeurs aberrantes, nous avons utilisé la routine Box-plot avec R. Il s'agit d'une technique de visualisation utile pour se faire une idée de la diffusion des données. Elles mettent notamment en évidence le biais de la distribution et indiquent la présence de valeurs aberrantes dans les données tel qu'indiqué dans la Figure 4.

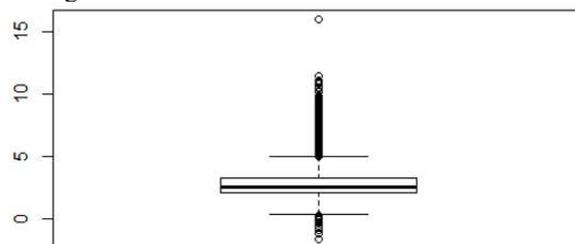


Figure 4 – Valeurs aberrantes de la Duree_Attente

Pour toutes les valeurs aberrantes, l'algorithme de Winsorization remplace chaque valeur détectée par le 5^e et 9^e centiles dans R. La Figure 5 présente le résultat de la Duree_Attente après le traitement de ses valeurs aberrantes.

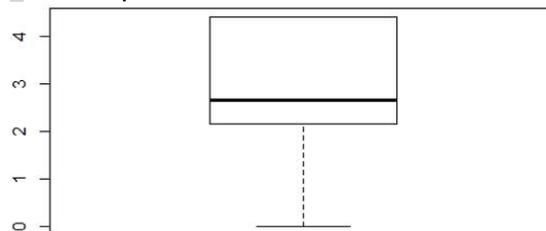


Figure 5 – Résultat après traitement des valeurs aberrantes

4 DÉCOUVERTE ET EXPLOITATION DES RÈGLES D'INTERVENTION

4.1 Tests préliminaires

L'objectif est de caractériser chaque type (T1 à T4) d'intervention de maintenance sur le four à brasage. En supposant que les observations sont indépendantes, nous avons appliqué les méthodes de RF, CART, SVM, ET et RL pour extraire les règles de classification des codes d'arrêt de maintenance. Nous avons divisé les données en un échantillon d'apprentissage (70% des observations) pour l'entraînement et le reste des données pour la validation. Le Tableau 1 présente les résultats avec les critères de qualité suivants : Accuracy (Acc.), Sensitivity (Sens.), Specificity (Spéf.) et Recall (Rec). Les meilleurs modèles sont donnés par RF, SVM, ET et C5.0. Pour sélectionner le meilleur modèle en termes de performances les quantités F1 Score, Kappa, Bias et Variance sont calculées. Les résultats sont présentés dans le Tableau 2.

Tableau 1. Les performances des méthodes de classification supervisée testées

	RF	AD	SVM	ET	RL	C5.0
T1	-	-	-	0	-	-
T2	100	100	75	78.57	80	100
T3	50	0	100	0	100	50
T4	50	0	100	100	66.6	100
Préc.(%)	98.9	99.1	98.8	98.8	98.2	99.2
Sens.(%)	42.3	23.9	40.8	31.3	26.7	46.8
Spéf.(%)	99.9	99.9	99.1	99.8	95.7	99.9
Rec.(%)	36.8	23.9	40.8	30	30.7	36.7

Tableau 2– Comparaison des méthodes de classification supervisée testées

	F1 Score	Kappa	Bias	Variance
RF	42.56	0.76	0	0.9
SVM	42.94	0.69	0.13	1.13
ET	42.20	0.71	-	-
C5.0	55	0.82	0.61	0.75

Les résultats montrent que la méthode de Boosting (C5.0) est la plus performante. Elle a un Kappa de 0.82 qui démontre que le modèle est en accord presque parfait avec les valeurs prédites. De plus la méthode C5.0 permet d'obtenir des résultats interprétables tel que souligné dans la revue de littérature. Ainsi, la méthode C5.0 est choisie comme le modèle de classification pour notre problématique.

4.2 Règles de décision obtenues par C5.0

La méthode C5.0 permet de caractériser la maintenance curative (T2) des autres types. La règle de classification s'énonce comme suit (Figure 6).

```
Rule 14: (34/6, lift 474.6)
  vitesse_Tapis > 4
  Type_Activite = S
  Local_interv in {A, C, E, F, non specifie}
  -> class 57 [0.806]
```

Figure 6 – Règle d'intervention de la maintenance curative

Si Vitesse Tapis > 4 m/s et Type Activité : Service (S) et la Localisation de l'intervention de maintenance dans les zones {A, C, E, F, et X pour non spécifique} alors l'intervention sera classée comme de la maintenance curative dans 81 % des cas. Bien que l'algorithme C5.0 ait de très bonnes métriques, pas de

sur-apprentissage (Overfitting) ni un sous-apprentissage (Underfitting), les règles de décision fournies ne sont pas exploitables facilement. En fait, la règle n'arrive pas à démêler les localités d'intervention et aucune variable caractérisant la charge de production n'y apparaît. Pour y remédier, nous avons apporté des modifications à la base de données afin de mieux caractériser la charge de production au moment de l'arrivée d'un événement de maintenance en calculant par exemples le cumul des pièces fabriquées et la vitesse moyenne de fonctionnement du tapis entre deux interventions successives de maintenance. Le Tableau 3 présente la qualité de la classification à l'aide des méthodes C5.0 et également les méthodes CART et RL en guise de comparaison. Il est clair que la méthode C5.0 dépasse ses concurrentes.

Tableau 3. Résultats des méthodes de classification après transformation de certaines variables

	C5.0	CART	RL
T1	0	-	-
T2	81.8	71.4	80
T3	100	33.3	50
T4	66.7	-	66.7
Accuracy	0.8	0.6	0.7
F1 Score	62.1	30	42.1
Kappa	0.6	0.2	0.5
Variance	31.57%	42.10%	36.84%
Bias	16.66%	44.44%	12.96

4.3 Importance des variables

Attribute usage:

```
85.11% Local_interv
57.45% Age_Four
38.30% vitesse_Moyenne
34.04% Duree_de_maintenance
21.28% cumul_Taille_Lot
```

Figure 6. Importance des variables utilisées par la méthode C5.0

Les résultats montrent clairement l'importance de l'âge du four, la vitesse moyenne du tapis, la durée d'exécution de l'intervention de maintenance et la charge appliquée au four (nombre de pièces fabriquées) dans la classification des interventions de maintenance.

4.4 Exploitation des résultats

Le Tableau 4 résume toutes les règles validées par l'équipe de maintenance. À noter que le type d'intervention de 1^{er} niveau est la classe de référence par rapport à laquelle les autres types d'intervention sont caractérisés.

Tableau 4 : Caractérisation des règles d'intervention

Règles d'intervention	
T2	Rule 2: (15, lift 1.8) Vitesse_Moyenne > 27.98 Local_interv = X -> class 57 [0.941]
	Rule 3: (13, lift 1.8) Duree_de_maintenance > 141.87 Local_interv = X -> class 57 [0.933]
	Rule 4: (7, lift 1.7) Cumul_Taille_Lot <= 33323 Age_Four > 68 Age_Four <= 88 Local_interv in {C, E, F} -> class 57 [0.889]
T3	Rule 5: (3, lift 4.2) Cumul_Taille_Lot > 33323 Age_Four <= 88 Local_interv in {A, E} -> class 58 [0.800]
	Rule 6: (7/1, lift 4.1) Age_Four <= 68 -> class 58 [0.778]
T4	Rule 7: (10/1, lift 3.9) Age_Four > 88 Local_interv in {A, B, C, D, E, F} -> class 59 [0.833]

Le Tableau 5 propose des pistes d'exploitation des résultats.

Tableau 5 : Exploitation des résultats

Recommandation d'exploitation	
T1	La durée de maintenance dure dans 80% des cas 2h 20 minutes.
T2	<ul style="list-style-type: none"> - en pratique, il faut surveiller le four suite au passage d'un ensemble de produits qui nécessitent une vitesse élevée du tapis du four. - la planification intégrée maintenance et production est à adopter. - il est recommandé de munir de capteurs le tapis transporteur et groupe d'entraînement (zone d'intervention C), l'installation électrique (zone E) et le générateur de gaz endothermique (zone F)
T3	<ul style="list-style-type: none"> - la planification intégrée maintenance et production est à adopter. - il est essentiel de mettre en place des capteurs dans la partie électrique du four car la zone E apparait comme la partie qui subit les pannes fréquemment.
T4	Plus le four est âgé, plus il est urgent de planifier plus d'interventions de maintenances systématiques.

4. CONCLUSION

La détection des pannes est l'un des éléments critiques de la maintenance prédictive ; elle est nécessaire pour que les industries puissent détecter les pannes rapidement et avec précision. Grâce à ce projet, nous sommes arrivés à extraire des règles interprétables d'intervention de maintenance. Ces règles dépendent des variables de production et des variables de

maintenance. Nous avons proposé l'installation de capteurs dans certaines parties du four afin de mieux caractériser les règles d'intervention. L'équipe de maintenance a la possibilité de converger vers une planification intégrée de maintenance et de production.

L'algorithme d'arbre de décision C5.0 est bien adapté à une problématique de maintenance. L'ensemble des résultats trouvés témoignent de l'applicabilité d'une approche basée sur les techniques d'apprentissage automatique à des problématiques de maintenance. De plus, l'implantation d'une maintenance 4.0 sur un cas réel n'est pas sans embûches. Il est important de s'assurer de la qualité des données et d'effectuer plusieurs tests avant de choisir la méthode d'extraction des connaissances adéquate.

Remerciements : Les auteurs souhaitent vivement remercier le personnel de Karmex et LOGIDAS pour les fructueuses discussions tout au long de ce travail ainsi que leur implication dans ce projet de recherche et développement.

RÉFÉRENCES

- Accorsi, R., Manzini, R., Pascarella, P., Patella, M., & Sassi, S. (2017). Data mining and machine learning for condition-based maintenance. *Procedia Manufacturing*, 11, 1153-1161.
- Ahmad, M. W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy system+ of cleaner production, 203, 810-821.
- Aivaliotis, P., Georgoulas, K., & Chryssoulouris, G. (2017). A RUL calculation approach based on physical-based simulation models for predictive maintenance. *International Conference on Engineering, Technology and Innovation (ICE/ITMC)*.
- Alaswad, S., & Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157, 54-63.
- Amruthnath, N., & Gupta, T. (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. *5th International Conference on Industrial Engineering and Applications (ICIEA)*.
- Ashjaei, M., & Bengtsson, M. (2017). Enhancing smart maintenance management using fog-computing technology. *International Conference on Industrial Engineering and Engineering Management (IEEM)*.
- Bel, L., Allard, D., Laurent, J. M., Cheddadi, R., & Bar-Hen, A. (2009). CART algorithm for spatial data: Application to environmental and ecological data. *Computational Statistics & Data Analysis*, 53(8), 3082-3093.
- Bjaili, H. A., & Rushdi, A. M. (2020). Prognostics and Health Monitoring Methodologies and Approaches: A Review. *Journal of Engineering Research and Reports*, 30-50.
- Bumblauskas, D., Gemmill, D., Igou, A., & Anzengruber, J. (2017). Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics. *Expert systems with applications*, 90, 303-317.
- Bukhsh, Z. A., Saeed, A., Stipanovic, I., & Doree, A. G. (2019). Predictive maintenance using tree-based classification techniques: A case of railway switches. *Transportation Research Part C: Emerging Technologies*, 101, 35-54.
- Calabrese, M., Cimmino, M., Fiume, F., Manfrin, M., Romeo, L., Ceccacci, S., & Kapetis, D. (2020). SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0. *Information*, 11(4), 202.
- Cauchi, N., Macek, K., & Abate, A. (2017). Model-based predictive maintenance in building automation systems with user discomfort. *Energy*, 138, 306-315.
- Carvalho, T. P., Soares, F. A., Vita, R., Francisco, R. D. P., Basto, J. P., & Alcalá, S. G. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024.

- Dragomir, O. E., Gouriveau, R., Dragomir, F., Minca, E., & Zerhouni, N. (2009, August). Review of prognostic problem in condition-based maintenance. European Control Conference (ECC).
- Falamarzi, A., Moridpour, S., Nazem, M., & Cheraghi, S. (2018). Development of a random forests regression model to predict track degradation index: Melbourne case study. In Australian Transport Research Forum.
- Ferreira, L. L., Albano, M., Silva, J., Martinho, D., Marreiros, G., Di Orio, G., & Ferreira, H. (2017). A pilot for proactive maintenance in industry 4.0. In 2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS).
- Harrou, F., Sun, Y., Taghezouit, B., Saidi, A., & Hamlati, M. E. (2018). Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. *Renewable energy*, 116, 22-37.
- Hasan, M. (2015). Deterioration prediction of concrete bridge components using artificial intelligence and stochastic methods.
- Jiménez, A. A., Zhang, L., Muñoz, C. Q. G., & Márquez, F. P. G. (2020). Maintenance management based on Machine Learning and nonlinear features in wind turbines. *Renewable Energy*, 146, 316-328.
- Lade, P., Ghosh, R., & Srinivasan, S. (2017). Manufacturing analytics and industrial internet of things. *IEEE Intelligent Systems*, 32(3), 74-79.
- Lesobre, R. (2015). Modélisation et optimisation de la maintenance et de la surveillance des systèmes multi-composants-Applications à la maintenance et à la conception de véhicules industriels, Thèse de doctorat, Automatique/Robotique. Université Grenoble Alpes.
- Mathew, V., Toby, T., Singh, V., Rao, B. M., & Kumar, M. G. (2017). Prediction of Remaining Useful Lifetime (RUL) of turbofan engine using machine learning. *International Conference on Circuits and Systems (ICCS)*.
- Ortiz, G. B. (2019). Multiclass classification using ensemble-boosting technique in R for the software quality data group development in IBM RPG.
- Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine-learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48(3), 128-138.
- Park, C., Moon, D., Do, N., & Bae, S. M. (2016). A predictive maintenance approach based on real-time internal parameter monitoring. *The International Journal of Advanced Manufacturing Technology*, 85(1), 623-632.
- Kautz, T., Eskofier, B. M., & Pasluosta, C. F. (2017). Generic performance measure for multiclass-classifiers. *Pattern Recognition*, 68, 111-125.
- Starr, A., Al-Najjar, B., Holmberg, K., Jantunen, E., Bellew, J., & Albarbar, A. (2010). Maintenance today and future trends. In *E-maintenance* (pp. 5-37). Springer, London.
- Steinberg, D., & Colla, P. (2009). CART: classification and regression trees. *The top ten algorithms in data mining*, 9, 179.
- Tran, H. D., Perera, B. J. C., & Ng, A. W. M. (2010). Markov and neural network models for prediction of structural deterioration of storm-water pipe assets. *Journal of Infrastructure Systems*, 16(2), 167-171.
- Wang, J., Liang, Y., Zheng, Y., Gao, R. X., & Zhang, F. (2020). An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples. *Renewable energy*, 145, 642-650.
- Yan, J., Koc, M., & Lee, J. (2004). A prognostic algorithm for machine performance assessment and its application. *Production Planning & Control*, 15(8), 796-801.
- Zenisek, J., Holzinger, F., & Affenzeller, M. (2019). Machine learning based concept drift detection for predictive maintenance. *Computers & Industrial Engineering*, 137, 106031.

CIGI 2021

Processus de surveillance autonome des machines tournantes.

XAVIER CHIEMENTIN¹, HASSANE HOTAIT¹, LANTO RASOLOFONDRAIBE²

¹ ITheMM, Institut de Thermique, Mécanique, Matériaux
URCA Reims - Moulin de la Housse - BP 1039 - 51687 Reims CEDEX 2
xavier.chiementin@univ-reims.fr, hassane.hotait@univ-reims.fr

² CReSTIC, Centre de Recherche en Sciences et Technologies de l'Information et de la Communication
URCA Reims - Moulin de la Housse - BP 1039 - 51687 Reims CEDEX 2
lanto.rasolofondraibe@univ-reims.fr

Résumé – Pour répondre à une concurrence de plus en plus accrue, l'avènement de l'industrie 4.0 offre des perspectives prometteuses en termes de compétitivité. Dans ce contexte, la maintenance prédictive, ou maintenance 4.0, devient un axe de réflexion pouvant s'avérer décisif. L'objectif de ce papier est de développer une méthodologie de la maintenance 4.0 pour briser les compromis des anciennes stratégies en permettant aux entreprises de maximiser la vie utile de leurs équipements de production tout en évitant les temps d'arrêt imprévus, par combinaison de l'analyse vibratoire et la méthode de reconnaissance de forme pour détecter, et suivre la dégradation d'un défaut. Une procédure est proposée, Dynamic-OPTICS (Dynamic Ordering Points to Identify the Clustering Structure), basée sur la méthode de reconnaissance de formes non supervisée. Cette méthode, autonome, vise à déceler les dysfonctionnements des machines tournantes et de suivre leur évolution. Cette démarche commence par l'extraction de trois types d'indicateurs (scalaires, temps-fréquence, temps-échelle), suivie par une réduction de dimension dans un espace visuel, représentant la présence de défauts dans le roulement. L'étape suivante consiste à suivre l'évolution de la dégradation en utilisant cinq paramètres : Davies-Bouldin index, Calinski-index, le saut du centre de la classe, la densité, le contour.

Mots clés – Surveillance, Roulement, Classification, OPTICS.

Keywords – Monitoring, Bearing, Clustering, OPTICS

1 INTRODUCTION

La concurrence de plus en plus accrue entre les entreprises, a fait naître deux défis jumelés que sont (i) la quête de la performance et (ii) la réduction des coûts de production. Pour y répondre, la maintenance se porte garante de la production. Sa fonction essentielle dans l'entreprise lui confère 4 niveaux d'analyse : la maintenance curative qui n'intervient qu'en cas de panne avérée, la maintenance préventive qui fait appel à une planification des interventions, la maintenance conditionnelle qui résulte d'une surveillance de non-dépassement de seuils et, enfin, une maintenance prédictive qui a pour objectif majeur l'estimation de la durée des équipements pour planifier ni trop tôt ni trop tard les interventions. Ce dernier niveau de maintenance reste un élément clé et admet un intérêt particulier de la communauté scientifique et des industriels. La maintenance prédictive, appelée également maintenance 4.0, souhaite se développer avec l'appui des outils de l'industrie 4.0: l'intelligence artificielle (IA), le jumeau numérique, la cybersécurité, la big data... (Liu et al., 2018; Rabah et al., 2018). D'après une enquête réalisée auprès de 220 industriels 91% plébiscite un intérêt envers la maintenance 4.0 (KPMG, 2017). Le cabinet McKinsey déclare que la maintenance 4.0 permettrait de faire un gain de 630 milliards de dollars d'ici 2025 ; de diminuer les coûts de production de 10 à 20%, de diviser par 2 le nombre de pannes et baisser les investissements de 3 à 5% (McKinsey Global Institute, 2015).

Parmi les applications visées par la maintenance 4.0, les machines tournantes possèdent une haute légitimité. Ces machines sont le plus souvent incontournables dans la performance d'une entreprise et peuvent subir de fortes sollicitations amenant des défaillances. On peut noter la prépondérance des défaillances électriques et mécaniques (Bonnett et Yung, 2008; Jelbaoui et El Menzhi, 2018). Une avarie mécanique peut se traduire par des arrachements de matières entraînant un écaillage sur les pistes de roulements ou sur les corps roulants, un balourd... Dans ce contexte de la maintenance 4.0, la surveillance en temps réel des systèmes c'est à dire la détection et le suivi de l'état pour l'estimation d'une durée de vie résiduelle reste déterminante pour assurer une maintenance optimale.

Cette surveillance est le plus souvent réalisée par le suivi de l'évolution d'un indicateur temporel (Jardine et al., 2006). Cette estimation permet de planifier un arrêt programmé de la machine pour le remplacement du (ou des) composant(s) (Koujok, 2010). Cependant, en raison de la variabilité permanente et irrégulière des paramètres de fonctionnement (Vachtsevanos et al., 2006) des machines tournantes en conditions non stationnaires, le diagnostic basé sur l'évolution de ces indicateurs pris séparément devient insuffisant. Par conséquent, un nombre important de méthodes ont été introduites pour prendre des décisions fiables et rapides sur la santé des machines en fonctionnement (Jiang et al., 2012;

Yiakopoulos et al., 2011). Le défi consiste également à développer des méthodes de diagnostic capables de détecter à un stade précoce l'apparition des défauts avec plus de précision que celles préexistantes à l'heure actuelle. Depuis quelques années, le diagnostic des machines tournantes emploie la complémentarité des indicateurs. Ce type de diagnostic passe par 3 étapes essentielles : (i) extraction des indicateurs par des techniques récentes de traitement du signal, (ii) sélection des paramètres pertinents (par projection ou par sélection), et (iii) classification des données.

Les méthodes de classification appartiennent à deux grandes catégories. Les méthodes de classification supervisée et non supervisée. Les Réseaux de Neurones Artificiels (RNA), les machines à vecteurs supports (Support Vector Machines SVMs), les arbres de décision, et les K plus proches voisins (Kppv), sont les méthodes de classification supervisée les plus connues et les plus utilisées. Elles sont souvent utilisées dans la détection et le diagnostic des défauts des machines tournantes. (Dellomo, 1999; Jack et Nandi, 2000; McCormick et Nandi, 1997). Les méthodes supervisées nécessitent la connaissance du nombre de classes à surveiller. Sans cette information les méthodes non supervisées sont à privilégier. Parmi ces méthodes, on peut noter les algorithmes DBscan et OPTICS, basée sur la densité des nuages des points. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) permet l'identification de classes, c'est à dire le regroupement des objets d'une base de données en sous-classes significatives. Elles présentent beaucoup d'avantage dans le diagnostic des défauts des machines tournantes. Cependant, on peut noter que ces techniques présentées offrent des classifications statiques et ne sont pas destinées à surveiller l'évolution d'un système. C'est pourquoi certaines études s'intéressent à la classification en temps réel afin d'améliorer le diagnostic. (Kerroumi et al., 2014) a proposé une méthode de classification dynamique basée sur le principe de l'algorithme DBSCAN pour la détection et la surveillance de l'état de dégradation des roulements. DBSCAN présente un inconvénient sur le choix du paramètre rayon ϵ . Si sa valeur est trop faible, la majorité des individus seront considérés comme des bruits et, si sa valeur est grande, les différents groupes peuvent être regroupés. L'Algorithme OPTICS (Ordering Points To Identify the Clustering Structure) a été conçu pour contourner cette difficulté. (Benmahdi et al., 2019) et (Hotait et al., 2020) ont proposé des méthodes combinant les avantages de la méthode OPTICS et la nature temps réel de la surveillance. Ces méthodes distinguent les variations des observations liées au changement du mode de fonctionnement du système de type variation de vitesse ou de charge et les variations liées aux défauts mécaniques. Cependant, ces méthodes ne proposent pas une stratégie complète, optimale et autonome de classification pour la maintenance conditionnelle. Certains éléments comme la projection des individus sur des composantes linéaires (ou non linéaires), le suivi des variables statistiques par rapport à un état de référence, l'automatisation des paramètres sont manquants.

Ce papier propose une procédure autonome, Dynamic-OPTICS, de surveillance de l'état de roulement, de la détection au suivi de la défaillance. Cette procédure propose également des nouveaux paramètres géométriques permettant le suivi des défaillances. La section 2 présente la méthode de classification OPTICS qui est le cœur du processus. La section 3 présentera l'ensemble du processus de surveillance ainsi que le principe d'automatisation. La section 4 est une validation

expérimentale du processus sur un banc de fatigue permettant la visualisation d'un endommagement naturel d'un roulement.

2 METHODE OPTICS

Cette section introduit la méthode de classification OPTICS (Ordering Points to Identify Clustering Structure) (Ankerst et al., 1999). Elle est le cœur du processus proposé. L'application de cette méthode n'est pas limitée à un seul domaine. Il a été utilisé en biologie, en astronomie, en topologie et récemment pour la détection du défaut des roulements dans les machines tournantes. Cette méthode est une méthode de partitionnement basé sur la densité des points. Elle est capable de déceler des différences significatives entre des nuages qui sont spatialement lointains. Sa force réside dans sa capacité à être robuste face à du bruit. Ainsi l'idée principale de l'algorithme OPTICS est que pour chaque point du nuage, son voisinage dans un rayon donnée (ϵ) doit contenir un minimum de points $MinPts$ où ϵ et $MinPts$ sont les paramètres d'entrées. Le concept de l'algorithme consiste à ajouter les points de façon itérative pour former les classes finales. Les formes de nuages peuvent prendre des formes complexes. La définition des classes passent par deux paramètres, le « core distance », C_d , and le « reachability distance », R_d , équations (1-2)

Si le nombre de points à proximité d'un objet, $N_\epsilon(p)$, est inférieur à $MinPts$, C_d est la distance entre p et $MinPts^{ième}$ voisin, noté $MinPts_{dist}(p)$. Dans ce cas, p est un « core-object ». Le « reachability distance » d'un objet o , R_d , est le maximum entre le « Core Distance » de p et la distance euclidienne entre o et p , figure 1.

$$C_d(\epsilon, MinPts(p)) = \begin{cases} Undefined & \text{si } N_\epsilon(p) < MinPts \\ MinPts_{dist}(p) & \text{sinon} \end{cases} \quad (1)$$

$$R_d(\epsilon, MinPts(p, o)) = \begin{cases} Undefined & \text{si } N_\epsilon(p) < MinPts \\ \max(C_d; dist(o, p)) & \text{sinon} \end{cases} \quad (2)$$

Le nombre de classes est déterminé par un graphique nommé « reachability plot », figure 1. Il correspond au nombre de vallées obtenues par la représentation des R_d en fonction de l'ordre des objets o .

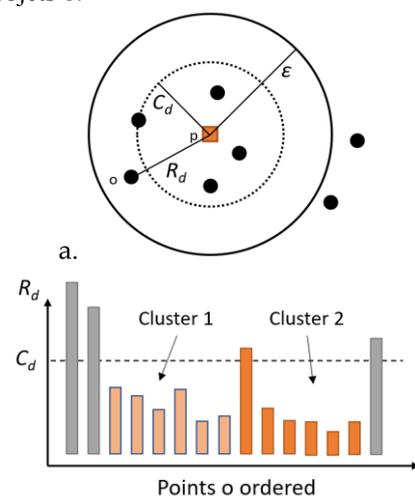


Figure 1. (a) Représentation de « Core Distance », « Reachability distance » pour $MinPts = 4$. (b) « Reachability plot »

3 PROPOSITION

Le processus proposé est itératif et possède la propriété de s'enrichir au cours du fonctionnement du système, figure 2. Il emploie principalement la méthode de partitionnement par densité, OPTICS. L'architecture du processus de surveillance repose sur trois boucles : initialisation, détection et suivi de la défaillance.

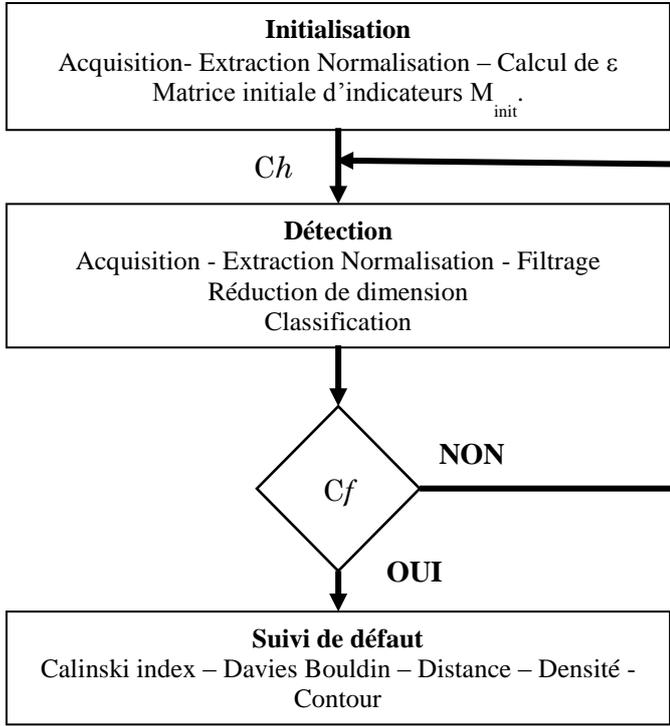


Figure 2 : Organigramme simplifié de la méthodologie de détection et de suivi d'un défaut mécanique.

La première boucle est exécutée pour une durée T_h , qui est supposée être une phase saine du roulement. Cette boucle forme la base initiale d'indicateurs. Pour chaque itération k , n signaux sont collectés, les indicateurs ($p = 17$) sont extraits dans les domaines temporel, fréquentiels et/ou temps-échelle, tableau 1. L'utilisation d'indicateurs provenant de différents domaines dans la détection des roulements défectueux peut offrir une détection plus précoce pour tout type de conditions de fonctionnement, vitesse et charge variables. En sortie de la boucle, une classe dite « saine » est formée. Le paramètre ε est calculé après une réduction de dimension. ε correspond à la distance maximale entre le centre de la classe saine Ch et le $MinPts^{ème}$ voisin, équation (3). $MinPts$ est automatisé en le liant aux nombres de mesures N pris à chaque itération. Des simulations ont montré que la valeur optimale était $MinPts = N/2$.

$$\varepsilon = distance(c_h, MinPts^{ième} \text{ voisin}) \quad (3)$$

Tableau 1. Indicateurs calculés pour la méthode Dynamic-OPTICS

Domain	Indicators
Time	Root Mean Square, Kurtosis, Peak to Peak, Crest Factor, Skewness, Impulse Factor, Standard deviation, Talaf, Tihkat
Frequency	Frequency Root mean square, Frequency Root mean square weighted, Frequency

	Center, Standard deviation frequency weighted, Mean, Power envelope
Time Scale	Effective value of the frequencies, Average value of the envelope amplitudes

La deuxième phase est une étape de détection de la défaillance mécanique. L'objectif de cette phase est de détecter un nouvel état appelé classe défectueuse, noté C_f . A chaque nouvelle itération de temps, les indicateurs sont extraits, normalisés par la méthode Z-scores, puis réduits et sélectionnés par la matrice des indicateurs comme dans la phase précédente. La nouvelle matrice d'indicateurs sélectionnée est réduite et classifiée puis testée par la méthode OPTICS pour détecter ou non une seconde classe. Si une seule classe est obtenue, ce qui correspond à l'état sain, l'algorithme reste en phase de détection, ces nouvelles données alimentent l'état sain. Si deux classes sont détectées, cette nouvelle classe C_f , est obtenue dans un plan B_f , qui sera conservé pour la phase de suivi.

La troisième étape est réalisée dans le plan B_f déterminé lors de la phase précédente. Il est important de conserver le même plan afin de visualiser l'évolution des indicateurs. Ce plan est le meilleur plan pour suivre l'évolution de la défaillance du roulement. A chaque nouvelle série de données, les indicateurs sont extraits, normalisés et projetés dans ce plan B_f . A partir de ces indicateurs, cinq paramètres géométriques GV_i sont calculés pour suivre l'évolution dans le temps : l'indice de Calinski-Harabasz, l'indice de Davies Bouldin, la distance, la densité et le contour.

L'indice de Calinski-Harabasz, (GV_1), également appelé critère de rapport de variance, est une mesure basée sur la dispersion des classes (Lukasik et al., 2016), équation (4).

$$GV_1 = \frac{\sum_i p d^2(c_f, c_h)/(n_c - 1)}{\frac{\sum_{x \in C_h} d^2(x, c_h)}{p - n_c} + \frac{\sum_{x \in C_f} d^2(x, c_f)}{p - n_c}} \quad (4)$$

p est le numéro des indicateurs. n_c est le nombre de classes. d est la distance euclidienne entre x, c_i .

L'indice Davies-Bouldin, (GV_2) mesure la moyenne des similarités entre chaque groupe, (Petrovic, 2006). Un indice plus faible signifie une meilleure configuration des classes. R_{ij} est la mesure de similarité de deux classes i et j , équation (5).

$$GV_2 = \frac{1}{n_c} \sum_{i=1}^{n_c} R_i \quad (5)$$

Avec $R_i = \max_{j=1 \dots n_c, i \neq j} (R_{ij}), i = 1 \dots n_c$

La distance (GV_3) correspond à la distance du centre de classe C_f entre deux temps successifs t et $t+1$, (Éq.3.2). $Cent_{2,t}$ est le centre de classe au temps t .

$$GV_3 = \left(d^2 (cent_{2,t}; cent_{2,t+1}) \right)^{\frac{1}{2}} \quad (6)$$

La densité (GV_4) est le nombre de points du groupe, N_f , pour une zone A_f .

$$GV_4 = \frac{N_f}{A_f} \quad (7)$$

Le contour (GV_5) de la classe est calculé à partir d'une enveloppe convexe qui correspond au plus petit ensemble convexe qui contient les points. L'algorithme de l'enveloppe convexe est intégré à Matlab par le moyen d'une fonction `convhulln`.

Les cinq paramètres possèdent le rôle de surveiller l'évolution de l'équipement dégradé avec le temps, ce qui nous donnera une idée générale du moment où il faudra changer l'équipement et le remplacer par un nouveau.

4 VALIDATION EXPERIMENTALE

4.1 Banc d'essais

Une validation est réalisée sur un banc de fatigue de roulement à butées à billes, figure 3. Un moteur synchrone entraîne un arbre en rotation à 1800 tours/minute. Un palier reçoit l'une des deux bagues de la butée examinée de référence RSB 51207. La deuxième bague de la butée, est chargée par un vérin à 30kN. Un accéléromètre est disposé radialement sur le châssis logeant une des butées. Le système est entraîné en mouvement par le moteur jusqu'à ce qu'un défaut d'écaillage apparaisse sur la piste de la bague de butée. À cette phase la taille du défaut d'écaillage est contrôlée régulièrement de façon visuelle. La butée testée comporte 12 billes et a une charge dynamique ISOC de 39000N. Les signaux sont collectés avec 8192 points et une fréquence d'échantillonnage de 51.2khz pour obtenir un nombre de cycles significatif. Après 50 heures de fonctionnement un signal était enregistré toutes les 3 heures jusqu'à 98 heures de fonctionnement. De 107 heures de fonctionnement jusqu'à 115 heures, un signal est enregistré toutes les heures. La fréquence d'enregistrement fut accélérée après 115 heures ; un signal est enregistré toutes les 30 minutes puis toutes les 5 minutes. Lors de la totalité de l'essai de fatigue, 5 mesures de défauts furent réalisées ce qui a nécessité un arrêt de fonctionnement temporaire. La mesure de la surface du défaut d'écaillage fut réalisée par microscope. Les résultats sont notés dans le tableau 2. La figure 3 montre le banc d'essais et l'évolution de défauts après 9 heures et 118 heures.

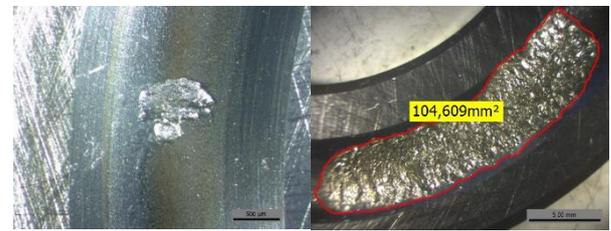


Figure 3. Banc d'essais SURVIB et exemple de défauts d'écaillage.

4.2 Application

La méthode dynamique automatisée Dynamic-OPTICS est représentée dans la figure 4. Après initialisation avec la matrice initiale ($T_{ini} = 8$ instants), la boucle de détection démarre, cette détection est faite pour le signal obtenu à 11h de fonctionnement, les paramètres d'entrée sont $MinPts = N/2 = 4$ et $\epsilon = 0.219$. La méthode proposée nous permet de détecter le défaut d'initialisation avec une précision de 100%.

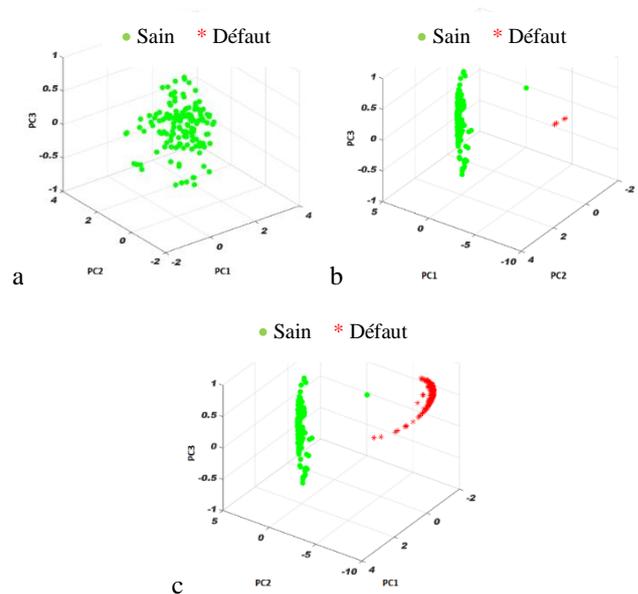
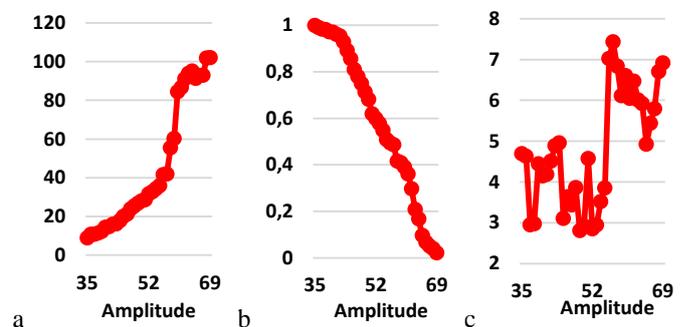


Tableau 2. Surface du défaut d'écaillage lors de l'essai de fatigue.

Durée (h)	114,0	116,0	116,3	117,1	118,5
Surface(mm²)	2.47	22.78	28.38	53.5	104.6

Figure 4. Méthodologie Dynamic-OPTICS, Validation expérimentale (a) Signal n°136-Sain (111 h), (b) Signal n°140-Défaut (112h), (c) Signal n°272-Défaut (118h25)



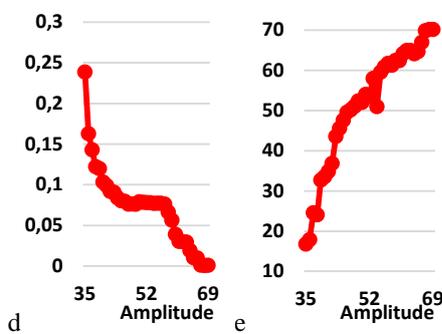


Figure 5. (a) Calinski-Index. (b) Davis-Bouldin Index, (c) Distance. (d) Densité, (e) Contour.

Le suivi commence à partir de 111h de fonctionnement (itération 35) jusqu'à l'arrêt correspond à 118.5h (itération 70) pour les 5 paramètres : indice de Calinski et Davies-Bouldin, densité, contour, distance, figure 5. Le comportement de ces variables est différent et reste similaire aux comportements établis lors de la simulation. L'indice de Calinski augmente avec la taille des défauts, avec une évolution exponentielle sous la forme mathématique $GV1 = 848.13e^{0.076k}$ ($R^2 = 0.975$), figure 5 (a). L'indice de Davies-Bouldin diminue proportionnellement au défaut, figure 5 (b). Dans ce cas, nous pouvons proposer une régression linéaire $GV2 = -0.0315k + 1.15$ ($R^2 = 0.975$). La densité diminuait avec l'augmentation des valeurs d'amplitude pour se situer autour de zéro du signal numéro trente-cinq à cinquante-huit, le modèle mathématique est $GV3 = 0.319 \cdot \exp(-0.0108k)$ ($R^2 = 0.628$), figure 5 (c). Le contour montre deux tendances, figure 5 (d). Le contour évolue proportionnellement avec la taille de défauts, $GV4 = 1.41k + 25.997$ ($R^2 = 0.934$). La courbe de distance augmente avec l'augmentation des valeurs d'amplitude de façon non monotone. L'évolution est exponentielle $GV5 = 3.323e^{0.018k}$ ($R^2 = 0.357$), figure 5 (d). Toutefois, la monotonie n'est pas pertinente. Il y a beaucoup de variabilité autour de la tendance moyenne, figure 5 (e).

En comparant l'évolution de ces indicateurs, les indices de Calinski et Davies Bouldin montrent leur pertinence, et indiquent le stade de dégradation dans le roulement. L'indice de Davies est Davies Bouldin l'indice le plus corrélé au nombre d'itérations ($R^2 = 0.975$). En général, ces indicateurs permettent de faire un pronostic sur l'évolution de ces paramètres avec les itérations. La surveillance des machines par analyse vibratoire occupe une place prépondérante dans le diagnostic et le pronostic de défauts des machines tournantes. Le pronostic et le suivi de défauts dans le cadre de la surveillance donne une idée générale sur le fonctionnement des équipements du système et modèle d'évolution de la dégradation.

5 CONCLUSION

Ce papier propose une méthodologie en ligne automatique de suivi des roulements à billes en automatisant et optimisant les paramètres internes de la méthode OPTICS. La surveillance dynamique Dynamic-OPTICS se divise en trois phases : l'initialisation, la détection et le suivi du défaut. Chaque phase fut explorée pour trouver les meilleures configurations en termes d'algorithmes. Les paramètres ϵ et $MinPts$ sont automatisés en considérant le contexte de l'analyse vibratoire. Cinq indicateurs géométriques sont évalués pour assurer un suivi de la défaillance. La méthodologie est éprouvée sur une évolution naturelle de défauts de type écaillage. La détection atteint une précision de 100%. Le suivi est assuré par des

valeurs géométriques dont la tendance suit des modèles mathématiques linéaires ou exponentiels avec des coefficients de corrélation atteignant 0,994.

Les perspectives de ce papier concernent la durée de vie résiduelle. Ce processus doit désormais permettre de définir les durées de vie résiduelles et donc intégrer des modèles de prédictions en s'appuyant sur les indicateurs géométriques testés.

6 REFERENCES

- Ankerst, M., Breunig, M. M., Kriegel, H. P. et Sander, J. (1999). OPTICS: Ordering Points to Identify the Clustering Structure. *SIGMOD Record (ACM Special Interest Group on Management of Data)*. 10.1145/304181.304187
- Benmahdi, D., Rasolofondraibe, L., Chimentin, X., Murer, S. et Felkaoui, A. (2019). RT-OPTICS: real-time classification based on OPTICS method to monitor bearings faults. *Journal of Intelligent Manufacturing*, 30(5), 2157- 2170. 10.1007/s10845-017-1375-6
- Bonnett, A. H. et Yung, C. (2008). Increased efficiency versus increased reliability. *IEEE Industry Applications Magazine*, 14(1), 29- 36. 10.1109/MIA.2007.909802
- Dellomo, M. R. (1999). Helicopter Gearbox Fault Detection: A Neural Network Based Approach. *Journal of Vibration and Acoustics*, 121(3), 265. 10.1115/1.2893975
- Hotait, H., Chimentin, X. et Rasolofondraibe, L. (2020). Monitoring of ball bearing based on Improved Real-Time OPTICS clustering. *Journal of Signal Processing Systems*, 93, 221-237.
- Jack, L. B. et Nandi, A. K. (2000). Genetic algorithms for feature selection in machine condition monitoring with vibration signals. *IEE Proceedings - Vision, Image, and Signal Processing*, 147(3), 205. 10.1049/ip-vis:20000325
- Jardine, A. K. S., Lin, D. et Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7), 1483- 1510. 10.1016/j.ymsp.2005.09.012
- Jelbaoui, Y. et El Menzhi, L. (2018). *Electromechanical Conversion Chain Fault Diagnostic -State of Art*. IPCO (dir.), .
- Jiang, L., Fu, X., Cui, J. et Li, Z. (2012). Fault detection of rolling element bearing based on principal component analysis. *Proceedings of the 2012 24th Chinese Control and Decision Conference, CCDC 2012*, 2944- 2948. 10.1109/CCDC.2012.6243071
- Kerroumi, S., Chimentin, X. et Rasolofondraibe, L. (2014). Online classification for spalling detection and vibratory behavior monitoring. *Mechanics & Industry*, 15(6), 517- 524. 10.1051/meca/2014058
- Koujok, M. El. (2010, 17 septembre). *Contribution au pronostic industriel : intégration de la confiance à un modèle prédictif neuro-flou*. [Université de Franche-Comté]. <https://www.hal.inserm.fr/ENSMM/tel-00542230v1>
- KPMG. (2017). *2eme baromètre KPMG-Usine Nouvelle de la transformation industrielle*.
- Liu, Z., Meyendorf, N. et Mrad, N. (2018, 20 avril). *The role of data fusion in predictive maintenance using digital twin* (vol. 1949), AIP Conference Proceedings, 2018. 10.1063/1.5031520
- Lukasik, S., Kowalski, P. A., Charytanowicz, M. et Kulczycki,

- P. (2016, 14 novembre). *Clustering using flower pollination algorithm and Calinski-Harabasz index* (p. 2724- 2728). 10.1109/CEC.2016.7744132
- McCormick, A. C. et Nandi, A. K. (1997). Classification of the rotating machine condition using artificial neural networks. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 211(6), 439- 450. 10.1243/0954406971521845
- McKinsey Global Institute. (2015). *The Internet of things: Mapping the value beyond the hype*.
www.mckinsey.com/mgi.
- Petrovic, S. (2006). A Comparison Between the Silhouette Index and the Davies-Bouldin Index in Labelling IDS Clusters. *Computer Science*.
- Rabah, S., Assila, A., Khouri, E., Maier, F., Ababsa, F., Bourny, V., Maier, P. et Mérienne, F. (2018). *Towards improving the future of manufacturing through digital twin and augmented reality technologies* (vol. 17, p. 460- 467). 10.1016/j.promfg.2018.10.070
- Vachtsevanos, G., Lewis, F. L., Roemer, M., Hess, A. et Wu, B. (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems - Vachtsevanos, George, Lewis, Frank L., Roemer, Michael, Hess, Andrew, Wu, Biqing - Livres* (Wiley). Wiley.
- Yiakopoulos, C. T., Gryllias, K. C. et Antoniadis, I. A. (2011). Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert Systems with Applications*, 38(3), 2888- 2911. 10.1016/J.ESWA.2010.08.083

CIGI 2021

Dynamic grouping maintenance considering quality degradation

HAI-CANH VU, AMELIE PONCHET-DURUPT, NASSIM BOUDAUD, ZOHRA CHERFI-BOULANGER

LABORATOIRE ROBERVAL, FRE UTC-CNRS 2012
UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE, ALLIANCE SORBONNE UNIVERSITE
RUE ROGER COUTTOLENC, 60200 COMPIEGNE, FRANCE

Résumé – Le regroupement dynamique est l'une des approches de maintenance les plus populaires développées pour les systèmes à composant multiple. Selon cette approche, les actions de maintenance sont exécutées conjointement pour réduire les coûts de maintenance. Compte tenu d'un certain nombre de modèles de regroupement présentés dans la littérature, la dégradation de la qualité de la production n'a pas encore prise en compte pour le regroupement dynamique. Pour cette raison, dans cet article, la dégradation de la qualité est d'abord modélisée en fonction de l'âge des composants. Un nouveau modèle de regroupement dynamique tenant compte de la dégradation de la qualité est ensuite développé. L'efficacité du modèle proposé et l'impact de la dégradation de la qualité sur la prise de décision en maintenance sont ensuite soulignés à travers différents exemples numériques.

Abstract – Dynamic grouping maintenance is one of the most popular maintenance approaches developed for multi-component systems. According to this approach, maintenance actions are jointly executed to reduce the maintenance cost. Given a number of grouping models presented in the literature, the degradation of product quality has not considered yet for the grouping maintenance. For this reason, in the paper, the degradation of production quality is first modelled as a function of the component age. A new dynamic grouping model considering the quality degradation is then developed. The effectiveness of the proposed model and the impact of the quality degradation on the grouping maintenance are then underlined through different numerical examples.

Mots clés – Optimisation de la maintenance, Qualité des produits, Systèmes à composant multiple.

Keywords – Maintenance optimization, Production quality, Multi-component systems.

1 INTRODUCTION

Grouping maintenance approach has been recently developed and becomes an interesting approach for the maintenance optimization of multi-component systems [De Jonge and Scarf, 2019]. By assuming that some activities can be shared among different preventive maintenance (PM) actions (opening machine, system downtime, spare part transportation, etc.), the main idea of this approach is to jointly execute several PM actions to reduce the maintenance cost. Among many grouping maintenance strategies, the dynamic grouping developed by [Wildeman et al., 1997] has drawn much attention thanks to its ability to consider different online information (e.g. a varying deterioration of components, unexpected opportunities, changes of utilization factors of components). This approach has been then extended by many authors to allow it to be applied in the different contexts, e.g. predictive maintenance [Van Horenbeek et Pintelon, 2013], complex multi-component systems [Vu et al., 2014], condition-based maintenance [Nguyen et al., 2015], maintenance durations [Vu et al., 2018], geographically dispersed production systems [Nguyen et al., 2019], stochastic dependence [Vijayan et Chaturvedi, 2020]. Given a number of grouping models presented in the literature, the degradation of the production quality over time has not considered yet for the dynamic grouping maintenance as well as for the maintenance optimization in general.

The quality degradation will lead the number of non-conforming products and related penalty costs to increase. The production system is unable to produce enough conforming

products to respond to the customer's need. One of the main reasons for the quality degradation is system aging. For this purpose, maintenance is a key solution to improve the system function and production quality. The consideration of the quality degradation in the maintenance decision-making is necessary.

For the above reason, we focus on developing a dynamic grouping model considering the quality degradation. The main contributions of the paper are:

- A long-term model of quality degradation. The quality degradation is modelled as a function of the component/system age;
- An age-based maintenance model at the component level considering the quality degradation;
- A dynamic grouping maintenance model at the system level considering the quality degradation and the maintenance opportunities triggered by production stops

This paper is organized as follows. Section 2 is devoted to present the general assumptions and the development of the quality degradation model. The problem of grouping maintenance considering quality degradation is introduced in Section 3. Section 4 focuses on the mathematical development of the problem. Several numerical examples are realized in section 5 to show the advantages of the proposed model. Finally, the last section presents the conclusions drawn from this work.

2 GENERAL ASSUMPTIONS

2.1 Quality degradation

Consider a production system consisting of n critical components that need to be maintained regularly. We assume that when these components operate perfectly, the system produces Q products per time unit with zero rate of non-conforming products (percentage of non-conforming products in Q products). However, in reality, they degrade over time and therefore do not operate correctly. The malfunctioning state of these components may lead the non-conforming rate to increase. This rate is modeled as a function of components' age.

$$ncr_i(x_i) = 1 - e^{-\left(\frac{x_i}{\mu_i}\right)^{\varepsilon_i}} \quad (2.1)$$

where, x_i is the age (total operational time) of the component i ($i = 1, \dots, n$); ncr_i is the non-conforming rate due to the malfunctioning state of the component; and ε_i , μ_i are the two constants.

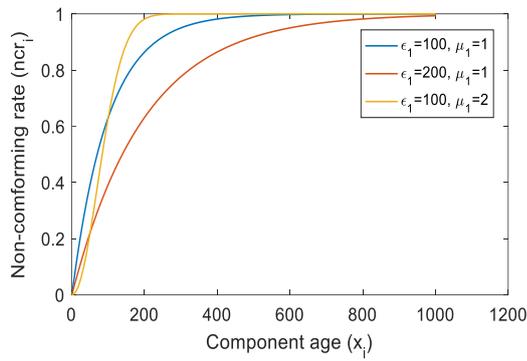


Figure 1. Non-conforming rate versus different values of ε_i and μ_i

Figure 1 shows that the proposed model ensures that the non-conforming rate is an increasing function of the component age and its value is between 0 and 1. Moreover, the different impacts of malfunctioning state of the component on the non-conforming rate can be modeled by varying the ε_i and μ_i values. It should be also noted that other model can be developed in real and complex applications by using machine learning approaches.

The penalty cost related to the non-conforming products per time unit can be calculated as

$$ncc_i(x_i) = ncr_i(x_i) \cdot Q \cdot c_{nc} \quad (2.2)$$

where, $ncr_i(x_i) \cdot Q$ is the number of non-conforming products per time unit due to the degradation of component i ; c_{nc} is the penalty cost paid for a non-conforming product.

2.2 Component failure

The component state is either operational or failed. It is assumed that the failure rate of component i denoted fr_i , is described by the Weibull probability density function with scale parameter $\lambda_i > 0$, and shape parameter β_i [Vu et al., 2014],

$$fr_i(x_i) = \frac{\beta_i}{\lambda_i} \left(\frac{x_i}{\lambda_i}\right)^{\beta_i-1} \quad (2.3)$$

The shape parameter β_i should be bigger than 1 to model the increasing behavior of the failure rate.

We assume that a prior analysis of the history data allows estimating the two parameters (λ_i and β_i) as well as the constants (ε_i and μ_i) in equation 2.1.

2.3 Maintenance activities

In order to prevent the system from failure and high non-conforming rate, each component is preventively maintained after a certain time. After a preventive action, the maintained component becomes new. The duration of a preventive maintenance (PM) activity of component i is considered as a constant and denoted by ω_i .

Each time when a PM activity is performed, the following costs must be paid:

- a fixed setup cost denoted S , represents the common costs for all preventive maintenance activities. This setup cost can be composed of the cost of crew traveling and preparation costs (e.g. erecting a scaffolding or opening a machine) [Wildeman et al., 1997];
- a specific cost denoted c_i^p , is related to the specific characteristics of component i such as spare part costs, specific tools and replacement procedures, etc.;
- a downtime cost rate denoted by c_d , paid for the system shutdown during a time unit. It may contain production loss costs, restart costs, etc. We assume that a PM of any component will lead the system to shut down.

As a consequence, the total costs of a PM action can be written as

$$C_i^p = S + c_i^p + c_d \cdot \omega_i \quad (2.4)$$

Between two PM actions, if the component fails, a minimal repair with negligible time is immediately performed to restore the component to the state that it had just before the maintenance. When the corrective action is carried out, it requires a corrective maintenance (CM) cost

$$C_i^c = S + c_i^c + c_u \quad (2.5)$$

where, c_i^c is a specific cost that depends on the specific characteristics of the component such as specific tools and repair procedures, etc.; c_u is unplanned shutdown cost that is an additional cost relying on production quality loss due to the failure or/and restart system costs, etc.

3 PROBLEM FORMULATION

The paper focuses on how to establish an effective PM plan for the system with the consideration of the following aspects:

3.1 Quality degradation

Given the quality degradation phenomenon (see subsection 2.1), the system should be maintained to meet the client demands. As mentioned above, the impact of the component degradation on the product quality is not the same (depending on ε_i and μ_i). We must therefore determine which components should be preventively maintained and at what time.

3.2 Maintenance opportunities

Two types of opportunities are considered: (1) production stops. When a PM action of a component is carried out at a system

stop, the downtime cost related to this action ($c_d \cdot \omega_i$) can be saved; (2) a PM of a component in the system. When a PM of a component is carried out at the same time with that of the others, the downtime cost, as well as the setup cost (S), can be shared among these actions.

In conclusion, a maintenance decision should be made with regarding the quality degradation and to the possibility of taking advantages of the maintenance opportunities to reduce the maintenance cost as much as possible. For this purpose, a dynamic grouping maintenance model considering quality degradation will be presented in the next section.

4 DYNAMIC GROUPING MAINTENANCE CONSIDERING QUALITY DEGRADATION

The dynamic grouping model is developed based on the grouping maintenance approach presented in [Vu et al., 2015]. The developed model contains two main steps:

- Maintenance optimization at the component level: in this step, the PM is planned separately for each component in an infinite planning horizon considering the quality degradation;
- Grouping maintenance optimization at the system level: the PM is planned jointly in short-term horizons. First, all the PM dates within the considered horizon are identified based on the individual PM plan obtained in the previous step. They are then updated to take advantage of the maintenance opportunities.

The detailed development of each step will be presented in the next subsections.

4.1 Maintenance optimization at the component level

To establish a PM plan for a component i in infinite horizon, the age-based maintenance policy is usually applied. For more detail, the component will be preventively maintained each time when its age reaches a threshold x_i^0 . The component will be minimally repaired if it fails between two consecutive PM actions. x_i^0 is determined by minimizing a long-run cost rate, denoted by $\theta_i(x_i^0)$. By applying the renewal theory [Smith, 1958], $\theta_i(x_i^0)$ can be calculated as

$$\theta_i(x_i^0) = \lim_{t \rightarrow \infty} \frac{C_i(t)}{t} = \frac{E[C_i(T_i)]}{E[T_i]} \quad (4.1)$$

where, $C_i(t)$ is the total operational cost of the component i at time t containing the maintenance cost (PM and CM), and the penalty cost related to the quality degradation; $E[x]$ is the mathematical expectation of x ; T_i and $C_i(T_i)$ denote the renewal cycle and the total maintenance within the cycle.

Since the PM actions are perfect, the renewal cycle T_i is the time between two consecutive PM actions. We have therefore

$$E[T_i] = T_i = x_i^0 + \omega_i \quad (4.2)$$

The total operational cost within the renewal cycle T_i is

$$E[C_i(T_i)] = C_i^p + C_i^c \cdot E[N_i(x_i^0)] + C_i^q(x_i^0) \quad (4.3)$$

where, $E[N_i(x_i^0)]$ is the mean number of failures within a renewal cycle. Since the component is minimally repaired at its

failures, $E[N_i(x_i^0)]$ can be calculated as

$$E[N_i(x_i^0)] = \int_0^{x_i^0} f r_i(x_i^0) dx_i^0 = \left(\frac{x_i^0}{\lambda_i}\right)^{\beta_i} \quad (4.4)$$

$C_i^q(x_i^0)$ is the penalty cost related to the quality degradation in x_i^0 time units. Thanks to equation 2.2, $C_i^q(x_i^0)$ can be written as

$$C_i^q(x_i^0) = \sum_{k=0}^{x_i^0} ncc_i(k) = \sum_{k=0}^{x_i^0} ncr_i(k) \cdot Q \cdot c_{nc} \quad (4.5)$$

Finally, we can calculate the long-run cost rate as

$$\theta_i(x_i^0) = \frac{E[C_i(T_i)]}{E[T_i]} = \frac{C_i^p + C_i^c \left(\frac{x_i^0}{\lambda_i}\right)^{\beta_i} + \sum_{k=0}^{x_i^0} ncr_i(k) \cdot Q \cdot c_{nc}}{x_i^0 + \omega_i} \quad (4.6)$$

By minimizing the above function, we can determine the optimal value of x_i^0 , denoted by x_i^* , and the long-run cost rate is then $\theta_i^* = \theta_i(x_i^*)$.

4.2 Grouping maintenance optimization at the system level

In this subsection, the PM is planned for all the components at the same time and in a finite horizon. The planification in finite horizons allows updating the maintenance plan to take advantage of the maintenance opportunities. To do this, we first identify all the PM actions within the finite horizon based on the x_i^* obtained from the previous step; evaluate the economic profit of the different updated solutions; optimize the update process. All these tasks will be presented in the next paragraphs.

4.2.1 Identification of PM actions within the planification horizon

Let consider a finite horizon $FH = [a, b]$ ($b > a \geq 0$). The PM dates of all the components can be determined as follow

$$t_{i1} = x_i^* - x_i(a) + D_0^1 \quad (4.7a)$$

$$t_{ij} = t_{i,j-1} + \omega_i + D_{j-1}^j, j > 1, t_{ij} \leq b \quad (4.7b)$$

where, t_{i1} and t_{ij} are the first and the j^{th} PM dates of component i within the FH ; $x_i(a)$ is the age of the component at time a ; D_0^1 and D_{j-1}^j are respectively the cumulative system downtimes in $[a, t_{i1}]$ and $[t_{i,j-1}, t_{ij}]$. All these identified PM dates are then modified (advanced or delayed) to consider the different maintenance opportunities.

4.2.2 Economic profit evaluation

As mentioned above, the PM of a component can be considered as an opportunity to carry out the PM of the other components. It means that when we jointly maintain several components at the same time, the maintenance cost can be saved. We therefore, try to modify the PM dates identified in the previous step to form the different groups of PM actions and evaluate the profit of the grouping.

Assume now a group of n_v PM actions denoted G_v , are simultaneously performed at time t_{G_v} . The corresponding economic profit of this group can be divided into the following parts:

- Since the execution of a group of maintenance actions requires only one setup cost, the group G_v yields a cost reduction:

$$\Delta H_{G^v}^1 = (n_v - 1) \cdot S \quad (4.8)$$

- When the PM actions are grouped, they will be executed at the same time at t_{G^v} instead of at their initial dates t_{ij} . This modification of the execution dates will lead to the following additional cost

$$\Delta H_{G^v}^2 = \sum_{ij \in G^v} C_i^c \cdot \left(\frac{x_i^* + \Delta t_{ij}}{\lambda_i} \right)^{\beta_i} - C_i^c \cdot \left(\frac{x_i^*}{\lambda_i} \right)^{\beta_i} - \Delta t_{ij} \cdot \theta_i^* \quad (4.9)$$

where, $\Delta t_{ij} = t_{ij} - t_{G^v}$.

- The grouping of the PM actions can save the system downtime cost because the system has to be shut down only once instead of several times when they are maintained separately.

$$\Delta H_{G^v}^3 = c_d \cdot \left(\sum_{ij \in G^v} \omega_i - \max_{ij \in G^v} \omega_i \right) \quad (4.10)$$

$\max_{ij \in G^v} \omega_i$ is the execution time of group G^v because we assume that the number of repair teams is enough to maintain all the components at the same time.

- Finally, the modification of the execution date of the PM actions also leads to the change of the number of non-conforming products. Let $\Delta H_{G^v}^4$ denote the penalty related to this change. We have

$$\Delta H_{G^v}^4 = \sum_{ij \in G^v} \left(\sum_{k=0}^{t_{G^v}} ncr_i(k) \cdot Q \cdot c_{nc} + \sum_{k=t_{G^v} + \max_{ij \in G^v} \omega_i}^b ncr_i(k) \cdot Q \cdot c_{nc} - \sum_{k=0}^{t_{ij}} ncr_i(k) \cdot Q \cdot c_{nc} - \sum_{k=t_{ij} + \omega_i}^b ncr_i(k) \cdot Q \cdot c_{nc} \right) \quad (4.11)$$

Thanks to equation from 4.8 to 4.11, we can deduce the expression of the economic profit of group G^v as

ss

$$EP(G_v, t_{G^v}) = \Delta H_{G^v}^1 - \Delta H_{G^v}^2 + \Delta H_{G^v}^3 - \Delta H_{G^v}^4 \quad (4.12)$$

It should be noted that until now, we have considered only the first type of maintenance opportunities (PM of a component). As mentioned above, in this paper, a production stop is also considered as an opportunity to do the PM at lower costs. To consider this kind of maintenance opportunity, the downtime periods of the system are modeled as fiction PM actions. For more detail, let 0^j ($j=1, 2, \dots$) denote the j^{th} downtime period in FH . 0^j starts at t_{0j} and finishes at $t_{0j} + \omega_{0j}$. If group G^v contains 0^j means that the components in the group are opportunistically maintained at 0^j . In this case, the system has already shut down, the penalty cost related to the system shutdown has not to be paid during the occurrence of the maintenance opportunity. In this case, we should modify the equation 4.10 as

$$\Delta H_{G^v}^3 = c_d \cdot \left[\sum_{ij \in G^v} \omega_i - \max \left(0, \max_{ij \in G^v} \omega_i - \omega_{0j} \right) \right] \quad (4.13)$$

4.2.3 Problem formulation and optimization

We define first a grouping solution, denote GS , is a collection of ng mutually exclusive groups $GS = \{G^1, \dots, G^{ng}\}$ which cover all N PM actions (real PM actions and fiction PM actions) in the planning horizon FH .

The total economic profit of the grouping solution is

$$EP(GS) = \sum_{v=1}^{ng} EP(G_v, t_{G^v}) \quad (4.14)$$

The problem of PM planification at the system level becomes the problem of finding the optimal grouping structure (denoted $GS^* = \{G^{1*}, \dots, G^{ng*}\}$ and the optimal execution time of each group $t_{GS^*} = \{t_{G^{1*}}^*, \dots, t_{G^{ng*}}^*\}$ which maximize the total economic profit.

$$\{GS^*, t_{GS^*}\} = \operatorname{argmax} [EP(GS)] \quad (4.15)$$

and

$$t_{G^{v*}}^* = \operatorname{argmin} (\Delta H_{G^{v*}}^2 + \Delta H_{G^{v*}}^4) \quad (4.16)$$

because $\Delta H_{G^{v*}}^1$ and $\Delta H_{G^{v*}}^3$ do not depend on $t_{G^{v*}}$.

The finding of the optimal grouping solution which maximizes the economic profit (equation 4.15 and 4.16) and respects the production constraint (4.17) is an NP-complete optimization problem. The exact methods will take too much time to find the optimal solution. Genetic algorithm (GA) [Whitley, 1994], a popular meta-heuristic method, is therefore used in this paper. GA is able to provide a good grouping solution in a reasonable computing time. In addition, its flexibility allows considering a number of maintenance constraints (repair team number, maintenance opportunity, opportunity duration, spare part availability, etc.) [Vu et al., 2014; Do et al., 2015; Vu et al., 2020].

5 NUMERICAL EXAMPLES

5.1 Given data

To analyze the performance of the proposed maintenance model, we consider here a system containing 5 critical components. The data related to the failure characteristics, the maintenance cost, and the impacts on the product quality of each component are generated randomly and reported in Table 1.

Table 1. Components' data

Component	λ_i	β_i	μ_i	ε_i	ω_i	c_i^c	c_i^p
1	100	2.50	350	3	2	32	500
2	305	2.35	400	2	3	36	825
3	120	2.70	380	3	2	66	300
4	270	2.20	390	2	3	60	600
5	250	2.40	380	1	3	40	950

The other data are: the setup cost $S = 100$; the downtime cost rate $c_d = 20$; the unplanned shutdown cost $c_u = 60$; the total number of products produced per time unit $Q = 20$; the penalty cost paid for a non-conforming product $c_{nc} = 1$.

5.2 Maintenance planning at the component level

We consider here the maintenance decision-making in the infinite horizon for each component separately. The maintenance decision-making is done based on the optimal age threshold of each component x_i^* . By substituting the data given in subsection 5.1 to the model developed in subsection 4.1, x_i^* and θ_i^* of each component were calculated and then reported in Table 2.

The total operational cost rate of the system is

$$\theta_{sys} = \sum_{i=1}^n \theta_i^* = 36.48 \quad (5.1)$$

Table 2. Optimal parameters of the maintenance planning at the component level

Component	1	2	3	4	5
x_i^*	128	229	119	200	220
θ_i^*	7.90	6.65	5.61	5.91	10.41

Now to analyze the impact of the quality degradation on the maintenance planning at this level, we assume that the non-conforming rate is equal to zero. The same calculations were realized and the obtained results are shown in Table 3.

Table 3. Optimal parameters of the maintenance planning at the component level without the consideration of the quality degradation

Component	1	2	3	4	5
x_i^*	136	500	125	434	442
θ_i^*	7.64	3.20	5.45	3.17	4.26

The total operational cost rate of the system when the quality degradation is not considered equals

$$\theta_{sys} = \sum_{i=1}^n \theta_i^* = 23.72 \quad (5.2)$$

Tables 2 and 3 show that x_i^* is decreased when the quality degradation is considered. It means that the component is preventively maintained more frequently. It seems to be reasonable because the maintenance decision was made regarding both the maintenance cost and the quality degradation. The component should be maintained more frequently to reduce not only the system failure probability but also the additional non-conforming products. In addition, it is natural when the operational cost (θ_{sys}) increase when the quality degradation is considered instead of only the maintenance cost.

5.3 Maintenance planning at the system level

We now determine the maintenance plan at the system level. The individual PM actions are grouped to reduce the operational cost. To do this, we first identify all the PM actions within a finite horizon. Without loss of generality, a horizon from 0 to 242 was chosen to cover at least one PM action of all components. In addition, a system stop is planned from $t_{01} = 125$ to $t_{01} + \omega_{01} = 125 + 3 = 128$. All the PM actions (5 real actions and 1 fiction) are shown in Table 4 and in Figure 2.

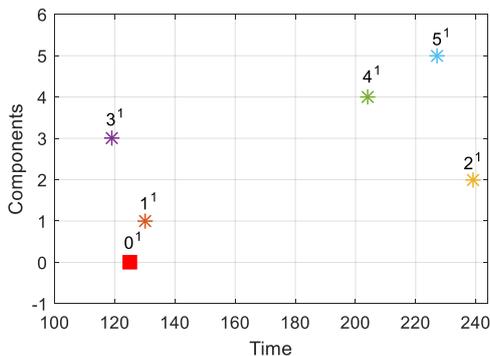


Figure 2. Preventive maintenance plan before the grouping

Table 4. Individual PM actions in the planning horizon

Component	0	1	2	3	4	5
t_{i1}	125	130	239	119	204	227

The individual PM actions are then grouped. The optimal grouping plan is found by minimizing the total economic profit calculated by equation 4.14 and reported in Table 5 and Figure 3.

Table 5. Optimal grouping solution

Group (v)	PM actions	Execution time ($t_{G^v}^*$)	Duration
1	$0^1, 1^1, 3^1$	125	2
2	$2^1, 4^1, 5^1$	224	3

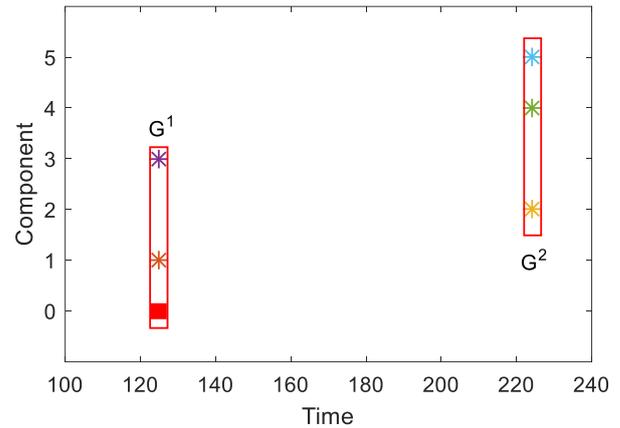


Figure 3. Preventive maintenance plan after the grouping

The optimal grouping solution contains two groups. The first group contains the fiction PM action (0^1) who represents the first system stop in the considered horizon. The group is executed at the occurrence time of the maintenance opportunity $t_{G^1} = t_{01} = 125$. It means that components 1 and 3 are opportunistically maintained at the production stop. The second group means that components 2, 4, and 5 are jointly maintained at $t_{G^2} = 224$. This grouping solution yields an economic profit that equals 394. When compared to the individual maintenance plan, the grouping plan helps to reduce 4.6% of the system operational cost rate.

Finally, to show the impact of quality degradation on the grouping maintenance, we repeat the above process but with zeros non-conforming rates. The obtained results, in this case, are presented in Table 6.

Table 6. Optimal grouping solution without the consideration of the quality degradation

Group (v)	PM actions	Execution time ($t_{G^v}^*$)	Duration
1	$0^1, 1^1, 3^1$	125	2
2	$2^1, 4^1, 5^1$	470	3

When the quality degradation is not considered, the execution time of the second group is delayed from 224 to 470, and this grouping plan helps to reduce only 3% of the system operational cost rate.

6 CONCLUSIONS

In this paper, the dynamic grouping maintenance approach has been developed to consider quality degradation. The quality degradation was modelled as a function of the component age and then integrated into the maintenance planning at both component and system levels. Numerical examples have shown that the consideration of the quality degradation is necessary and important for the accuracy of the maintenance plans. However, further studies are needed to improve the quality degradation modeling, to consider the maintenance imperfect, or to carry out the sensitivity analysis of the performance of the proposed model.

7 REFERENCES

- De Jonge, B., & Scarf, P. A. (2019). A review on maintenance optimization. *European Journal of Operational Research*.
- Wildeman, R., Dekker, R., & Smit, A. C. J. M. (1997). A dynamic policy for grouping maintenance activities. *European Journal of Operational Research*, 99(3), 530-551.
- Van Horenbeek, A., & Pintelon, L. (2013). A dynamic predictive maintenance policy for complex multi-component systems. *Reliability engineering & system safety*, 120, 39-50.
- Vu, H. C., Do, P., Barros, A., & Bérenguer, C. (2014). Maintenance grouping strategy for multi-component systems with dynamic contexts. *Reliability Engineering & System Safety*, 132, 233-249.
- Nguyen, K. A., Do, P., & Grall, A. (2015). Multi-level predictive maintenance for multi-component systems. *Reliability engineering & system safety*, 144, 83-94.
- Vu, H. C., Do, P., & Barros, A. (2018). A study on the impacts of maintenance duration on dynamic grouping modeling and optimization of multicomponent systems. *IEEE Transactions on Reliability*, 67(3), 1377-1392.
- Nguyen, H. S. H., Do, P., Vu, H. C., & lung, B. (2019). Dynamic maintenance grouping and routing for geographically dispersed production systems. *Reliability Engineering & System Safety*, 185, 392-404.
- Vijayan, V., & Chaturvedi, S. K. (2020). Multi-component maintenance grouping optimization based on stochastic dependency. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 1748006X20947511.
- Vu, H. C., Do, P., Barros, A., & Bérenguer, C. (2015). Maintenance planning and dynamic grouping for multi-component systems with positive and negative economic dependencies. *IMA Journal of Management Mathematics*, 26(2), 145-170.
- Smith, W. L. (1958). Renewal theory and its ramifications. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 243-284.
- Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and computing*, 4(2), 65-85.
- Do, P., Vu, H. C., Barros, A., & Bérenguer, C. (2015). Maintenance grouping for multi-component systems with availability constraints and limited maintenance teams. *Reliability Engineering & System Safety*, 142, 56-67.

- Vu, H. C., Do, P., Fouladirad, M., & Grall, A. (2020). Dynamic opportunistic maintenance planning for multi-component redundant systems with various types of opportunities. *Reliability Engineering & System Safety*, 198, 106854.

Session 2.2 : Management de la Qualité & Amélioration continue

Chairman : Georges Abdounour - Michel Tollenaere

Une démarche Lean dans la phase amont du cycle de vie produit – Application au processus opérationnel de l'Entreprise Stäubli Electrical Connectors SAS

SEBASTIEN MARANZANA¹ BERTRAND ROSE² JEAN-MICHEL LEIRITZ³

¹ ICUBE

300 bd Sébastien Brand – CS 10413 67412 Illkirch, France
sebastien.maranzana2@etu.unistra.fr

² ICUBE

300 bd Sébastien Brand – CS 10413 67412 Illkirch, France
bertrand.rose@unistra.fr

³ STÄUBLI EC SAS

4 rue de l'Industrie – 68220 Hémingue, France
jm.leiritz@staubli.com

Résumé – L'évolution des marchés, soumis à une forte pression concurrentielle et technologique, oblige les entreprises à évoluer pour gagner en performance en optimisant le triptyque coût-qualité-délai.

Les besoins des clients sont toujours plus complexes et les délais demandés plus courts. Dans ces conditions, le lancement d'un nouveau produit représente un défi pour les entreprises. Période-clé, la phase amont du cycle de vie produit repose principalement sur l'efficacité des processus opérationnels et des systèmes d'information de l'entreprise.

Le Lean a démontré son intérêt pour améliorer la performance industrielle dans le domaine de la production et est aujourd'hui décliné dans différents environnements. Le déploiement d'une méthodologie Lean dans la phase amont du cycle produit basée sur l'interface processus opérationnels/systèmes d'information reste cependant peu étudié.

La démarche mise en place par Stäubli EC SAS repose sur la suppression des Mudras du Lean et l'alignement des systèmes d'information aux processus opérationnels dans le but d'optimiser la chaîne de valeur.

L'amélioration de la performance, dans la phase de lancement d'un nouveau produit par une approche globale qui compile les principes et outils du Lean IT, du Lean Office et du Lean Product Development.

Abstract – The evolution of markets, subject to strong competitive and technological pressure, obligate companies to evolve in order to gain in performance by optimising the triptych cost-quality-delay.

Customers' needs are becoming increasingly complex and deadlines required are shorter and shorter. Under these conditions, the launch of a new product represents a challenge for companies. A key period, the upstream phase of the product lifecycle relies mainly on the efficiency of the company's operational processes and information systems.

Lean has demonstrated its interest in improving industrial performance in the field of production and is now used in different environments. However, the deployment of a Lean methodology in the upstream phase of the product cycle based on the operational process/information system interface remains not much studied.

The approach implemented by Stäubli EC SAS is based on the elimination of Lean Mudras and the alignment of information systems with operational processes in order to optimise the value stream.

Performance improvement, in the launch phase of a new product by using a global approach that compiles principles and tools of Lean IT, Lean Office and Lean Product Development.

Mots clés – Lean IT, Phase amont cycle de vie produit, Performance industrielle, Master Data.

Keywords – Lean IT, Upstream product life cycle, Industrial performance, Master Data.

1 CONTEXTE

1.1 Stäubli Electrical Connectors SAS

Stäubli Electrical Connectors SAS est une entreprise spécialisée dans la fabrication de connecteurs électriques.

Basée sur le site de Hémingue (68), l'entreprise appartient à la division Electrical Connectors du Groupe Stäubli. Cette division est composée de six sites dont trois historiquement implantés dans la Région des Trois Frontières (Allemagne, Suisse et France) articulées autour de la ville de Bâle (CH). Cette implantation est issue d'une volonté de favoriser les échanges entre les sites et d'assurer une représentation nationale.

Basés sur la technologie des contacts à lamelles Multilam®, les connecteurs électriques Stäubli EC possèdent une résistance de

contact très faible et maîtrisé permettant le passage de fort courant. Ces connecteurs sont voulus de haute qualité et conçus pour fonctionner dans des environnements exigeants, développés spécifiquement pour répondre aux besoins du client et garantir un maximum de performance.

Centre de compétence ferroviaire et E-mobility, Stäubli EC SAS est présent sur différents marchés en France et aussi à l'international via le site d'Allschwil (CH), centre logistique et de distribution internationale de la division EC. Le site de Hémingue possède un outil industriel complet : équipes projet et commerciales, bureau d'études, service achats et qualité, un outil de production (montage) moderne et un laboratoire d'essais complet. Ces moyens permettent une activité R&D mais également d'absorber une activité projet importante : concevoir

Et produire des solutions sur mesure pour des applications clients pointues.

L'éventail de produits développés et fabriqués par Stäubli EC SAS est très large et varié allant du produit standard présent au catalogue aux solutions spécifiques pour des applications clients, en passant par des développements permettant de répondre à des problématiques connues, issues de la R&D.

Ce mix produit important implique une grande variabilité en termes de moyens de production, de volumes et de délais requis et impose une grande complexité des modes de gestion de la Supply Chain.

1.2 Problématiques liées au Marché

Présent sur des marchés en constante évolution et soumis à une forte pression concurrentielle et technologique, Stäubli EC SAS doit continuellement évoluer pour gagner en performance en optimisant le triptyque coût-qualité-délai. Un des facteurs de réussite réside dans la capacité à développer et produire des solutions complexes et variées pour répondre aux besoins des clients, et ce dans des délais toujours courts.

En effet, depuis les années 2000 le marché des connecteurs électriques a vu s'établir des acteurs implantés dans des pays à faible coût de production et l'ouverture des marchés permet aux clients d'avoir accès à un large choix de fournisseurs chez lesquels s'approvisionner. Les acteurs historiques et leaders du secteur sont particulièrement sensibles à cette pression concurrentielle.

Parmi les facteurs clés coût-qualité-délai pris en compte par les clients, aujourd'hui la qualité des produits n'est plus une marque de démarcation. Dans les éléments décisifs, le prix est un facteur majeur, mais de plus en plus le délai de livraison est pris en compte par le client dans le choix du fournisseur.

De plus, on remarque une évolution du marché et des besoins, les clients sont davantage à la recherche de solutions globales assurant une fonction complète plutôt qu'un fournisseur de composants à intégrer.

Cette modification du mode de consommation entraîne une complexification des produits et des changements chez les fabricants (spécialisation et/ou changement de cœur de métier).

Le marché de la connectique électrique est caractérisé par une évolution rapide des technologies, des normes et des standards ainsi que l'arrivée constante de nouvelles applications. Il offre des opportunités intéressantes, mais pour en profiter il est capital pour l'entreprise d'avoir la capacité de constamment s'adapter au marché et d'évoluer rapidement pour répondre aux besoins des clients.

Ces constats amènent vers la problématique industrielle suivante à laquelle est actuellement confronté Stäubli Electrical Connector SAS :

Comment livrer des produits de plus en plus complexes dans des délais toujours plus courts ?

1.3 Problématiques opérationnelles

La réponse à la problématique issue du Marché repose sur la capacité de l'entreprise à s'adapter aux besoins des clients : Proposer les bons produits dans les bons délais. Pour arriver à cet objectif, l'entreprise doit améliorer sa performance industrielle et par conséquent sa réactivité durant la phase amont du cycle de vie produit ; démarrant à la définition du cahier des charges avant la conception du produit jusqu'aux premières livraisons chez le client.

Le lancement d'un nouveau produit représente une période-clé dont la réussite repose tout particulièrement sur la performance des processus opérationnels en place et la capacité des différents

Intervenants de l'entreprise (Projet, Bureau d'Etude, Méthodes, Achat, Planification, Production ...) à communiquer entre eux de manière efficace.

L'alignement entre les processus et les systèmes d'information afin de répondre aux besoins opérationnels, représente un défi déterminant pour les entreprises qui souhaitent améliorer leurs performances industrielles.

Cet alignement processus opérationnels/ systèmes d'information est d'autant plus complexe pour les filiales de Groupe : les systèmes d'information étant les plus couramment déployés de manière globale et les processus standardisés bien que les structures ne soient pas toujours identiques et que des spécificités peuvent faire varier de manière importante les besoins. Ce constat peut aussi être accentué par les différences culturelles ou encore réglementaires dans le cadre de Groupes Internationaux.

La performance industrielle d'une entreprise est donc basée sur sa capacité à adapter son organisation, ses processus opérationnels et son système de communication pour répondre aux besoins des clients.

Ce qui amène vers la problématique opérationnelle suivante :

Comment améliorer la performance industrielle durant la phase amont du cycle de vie produit ?

2 ORIGINALITES

2.1 Performance industrielle

Pourquoi être plus performant ? Pour se démarquer de la concurrence, se développer ou simplement survivre sur des marchés en constante évolution.

Il est important de définir cette notion de performance industrielle. D'après L.Berrah : « Un système industriel est performant s'il est efficace, efficient et effectif » [Berrah, 2013].

Historiquement taylorienne, la notion de performance repose sur la réussite économique et prend en compte l'efficacité correspondant au rendement des équipements et de la productivité de la Main d'œuvre. Aujourd'hui néo-taylorienne, la notion de performance devient organisationnelle et prend désormais également en compte les éléments systémiques suivants : L'efficacité qui représente la qualité dans l'atteinte des objectifs, le ratio entre les résultats et les moyens consommés pour y arriver. L'effectivité plus qualitative consiste à rapporter le résultat d'une action à l'intention qui la fonde (est-ce que l'action réalisée est celle qui avait été prévue ?).

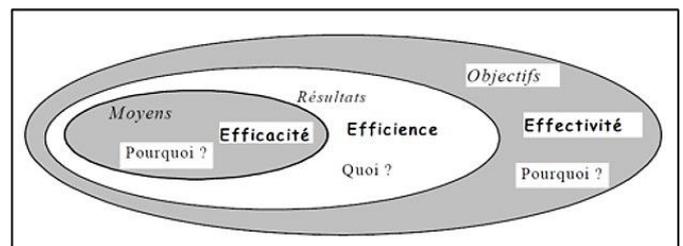


Figure 1. Efficacité/Efficience/Effectivité [Berrah, 2013 inspiré de Jacot, 1996]

La performance industrielle ne se résume pas uniquement au résultat financier. Pour assurer la pérennité de cette performance, l'entreprise doit prendre en considération l'aspect social : l'intérêt des salariés, des clients et des territoires, mais également de l'environnement : les ressources, les évolutions des pratiques ... [Galichet, 2018].

D'après H.Hamadmad, « la performance ne prend sens que si elle se mesure » Pour effectuer des mesures, il faut définir les principaux critères que l'on veut évaluer, généralement des fonctions que l'on veut optimiser.

Cinq critères principaux sont généralement utilisés pour caractériser la performance industrielle [Hamadmad, 2016] :

- Le coût
- La qualité
- Le délai
- L'innovation
- Le développement durable

Cette notion de performance industrielle est d'autant plus importante lors de la phase de lancement d'un nouveau produit, où sa réussite dépend de la capacité de l'entreprise à fournir le bon produit, au bon prix et dans le bon délai. La capacité de l'entreprise à maîtriser ces critères de performance est nécessaire pour trouver sa place sur le marché.

2.2 Phase amont du cycle de vie produit

Cette phase de lancement correspond à la première phase du cycle de vie produit.

La notion de cycle de vie produit apparue pour la première fois dans les années cinquante et soixante suite à l'analyse de l'évolution caractéristique des marchés des produits, depuis leur apparition jusqu'à leur disparition.

Le cycle de vie produit est couramment représenté par une « courbe de vie » : cette courbe en forme de cloche représente les trois phases caractéristiques de l'évolution du produit au cours de sa vie. La première représente une phase d'expansion correspondant à la mise sur le marché du produit de ses éventuelles adaptations aux besoins. La seconde phase prend la forme d'un plateau représentant la période d'utilisation du produit par les utilisateurs. La dernière et troisième phase : la récession qui correspond à l'obsolescence du produit principalement liée à l'évolution technologique ou à la modification des pratiques de consommation. La représentation habituelle montre l'évolution dans le temps de « l'indice de satisfaction » traduit en unité comptable : ventes, chiffre d'affaires ... [Bernat et Marcon, 2001].

La phase amont du cycle produit nous intéresse tout particulièrement, car elle condense de nombreux facteurs de la performance industrielle.

La « naissance » d'un nouveau produit est intimement liée au processus d'innovation utilisé par l'entreprise. La représentation par processus de la démarche d'innovation a pour but de décrire les différents fonctions et métiers mobilisés ainsi que le mécanisme et les différentes étapes/jalons permettant les prises de décision pour arriver à la mise au point d'un nouveau produit [Ferioli, 2018].

Le processus d'innovation utilisé par Stäubli EC présenté figure 2, se base sur les principes du modèle State-Gate [Cooper, 2001], modèle courant dans l'industrie, car offrant une représentation macro de l'ensemble du processus d'innovation.

Ce modèle divise le processus de développement de nouveaux produits en un ensemble d'étapes réalisées par une équipe projet et à l'entrée de chaque étape se trouve une porte (gate) qui représente des points de décision (go/no go). Les décisions sont prises à l'issue de revues faites par les décideurs de l'entreprise après analyse des livrables de la phase précédente.

Le modèle Stage-Gate System a pour but de décrire facilement l'organisation des activités durant l'ensemble du processus et de permettre la prise de décision. Chaque étape devient plus complexe et coûteuse au fur et à mesure que l'engagement de l'entreprise dans le développement du produit augmente.

Le bon déroulement de ce processus permet de garantir la maturité et la viabilité du projet et de son produit.

Néanmoins cette vision axée gestion de projet reste éloignée de l'activité opérationnelle de l'entreprise.

Le lien entre l'activité « projet » et l'activité opérationnelle est très important, car souvent les mêmes acteurs interviennent sur plusieurs projets de manière simultanée en plus de leurs activités opérationnelles hors des phases projet.

De même il est nécessaire pour les Responsables de projet d'avoir conscience des tâches opérationnelles à tous les niveaux du cycle de vie projet.

C'est ce processus opérationnel souvent oublié et pourtant essentiel qui nous intéresse pour cette étude. En effet les actions permettant le lancement d'un produit reposent sur le processus opérationnel et ses acteurs. Son lien avec la performance industrielle est donc direct.

2.3 Système d'information

Ce processus opérationnel repose sur un ensemble de systèmes d'information qui permettent la communication entre les différents intervenants de l'entreprise.

Un système d'information (SI) est un ensemble organisé de ressources organisationnelles et techniques permettant l'acquisition, le stockage et la diffusion de l'information afin de fournir des services opérationnels aux acteurs de l'organisation [El Mahou, 2018].

Chaque système d'information intervient sur un périmètre défini et dans le cadre d'une entreprise, il s'agit de faire interagir les différents SI entre eux afin d'assurer la transmission des bonnes informations.

L'importance de l'alignement des SI a depuis longtemps été étudiée d'un point de vue stratégique [Weil et Broadbent, 1997], mais rarement d'un point de vue opérationnel.

Cet alignement des SI est d'autant plus difficile dans les grandes structures ou les entreprises appartenant à un groupe ayant une cartographie SI complexe [Adam et al, 2003].

Pour assurer un service performant à l'entreprise, les SI doivent être en adéquation avec la stratégie (business plan) et les processus de l'organisation [Gmati et Nurcan, 2007].

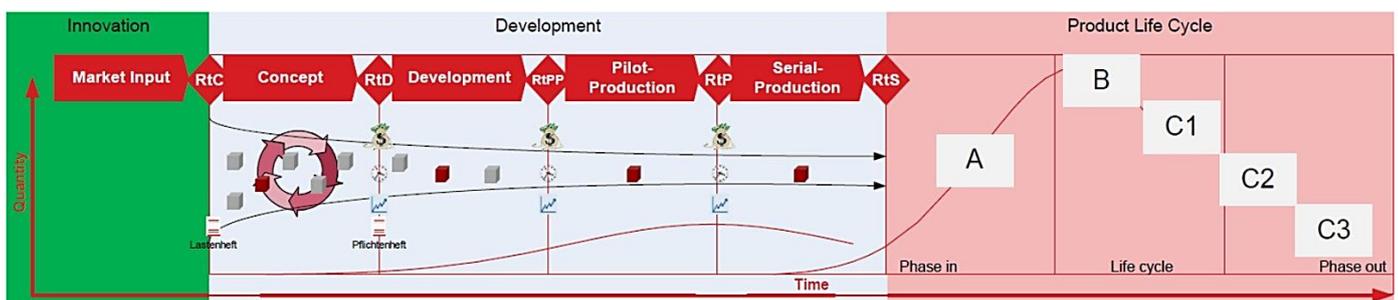


Figure 2. Stage Gate ECS [Stäubli EC]

Différents modèles ont été développés concernant l'alignement des SI au marché ou à la stratégie de l'entreprise : parmi les plus connus le modèle ARIS : Architecture of Integrated Information System [Scheer et Nuttgens, 2000] et le modèle SAM : Strategic Alignment Model [Henderson et Vendatraman, 1999]. Toutefois peu de travaux portent sur le lien entre les SI et les processus de l'organisation. Cette réflexion de déploiement stratégique est classiquement définie et mise en application par une entité appelée Direction des Systèmes d'information (DSI).

Loin de la vision stratégique, les entreprises appartenant à un grand groupe sont souvent confrontées à une vision terrain : comment assurer le lien entre les différents SI en place et les processus opérationnels et répondre aux différents besoins ?

Afin de répondre à cette problématique, la division EC de Stäubli a déployé une DSI, voir figure 3, en créant sur chaque site du Groupe une fonction appelée Master Data et établi une organisation permettant de les coordonner.

Les missions de Master Data étant :

- Etablir les processus en interaction avec les SI.
- Définir les règles d'utilisation des SI.
- Maintenir et garantir la justesse des données des SI.
- Veiller à ce que les SI répondent aux besoins opérationnels.

Les acteurs Master Data de chaque site sont coordonnés par une entité Groupe ayant pour mission d'établir les règles globales d'utilisation des SI afin de standardiser les pratiques. Mais également de définir, organiser et accompagner les sites dans les évolutions des SI.

Le but de l'organisation Master Data chez Stäubli EC est de garantir la juste transmission des informations au sein de l'ensemble de la division durant l'ensemble du cycle de vie des produits, de leur création à leur fin de vie.

2.4 Le Lean

Depuis son apparition à la fin des années 1970, le Lean a su démontrer son efficacité en matière d'amélioration des performances industrielles des entreprises manufacturières. La définition du Lean varie selon les auteurs, mais il peut être défini comme une philosophie ou un ensemble de pratiques

visant la réalisation des améliorations en suivant les voies les plus économiques tout en se focalisant sur la réduction des gaspillages [Elrhanimi et al, 2015].

L'ouvrage « Lean Thinking » [Womack et Jones, 1996], étude rétrospective sur la pratique du Lean, synthétise la philosophie Lean en cinq principes :

- Identifier précisément la valeur ajoutée des produits, d'un point de vue du client.
- Déterminer la chaîne de valeur des produits.
- Etablir des flux continus de valeur.
- Tirer les flux par le client.
- Viser l'excellence.

Le Lean prend racine dans les pratiques développées par l'entreprise automobile TOYOTA et son système de production : le Toyota Production System (TPS). Ces pratiques ont été analysées et transposées dans l'ouvrage « The Toyota Way » [Liker, 2004]. Le TPS repose sur la chasse à la non-valeur ajoutée. Les gaspillages (Mudas) identifiés sont au nombre de huit : surproduction, surstock, transports inutiles, processus inutiles, mouvements inutiles, erreurs et rebuts, temps d'attente et sous-utilisation des compétences des collaborateurs.

Deux autres typologies de non-valeur ajoutée ont également été définies : la variabilité (Mura) et la surcharge de travail (Muri).

Initialement Lean Manufacturing, car issus de l'industrie manufacturière, les principes du Lean ont été transposés de la production à d'autres secteurs de l'entreprise et à d'autres activités économiques.

Ainsi au début des années 2000, le Lean sort des lignes de production et ses principes sont déclinés tout d'abord dans l'univers du management (Lean Management) et des Services (Lean Service) puis adaptés à de nombreuses activités : les services administratifs (Lean Office), le développement de produit (Lean Product Développement) ou encore le développement informatique (Lean Software développement) [Ughetto, 2012].

Nous nous intéresserons désormais davantage au Lean IT (Information Technologie), la philosophie Lean appliquée au Système d'Information.

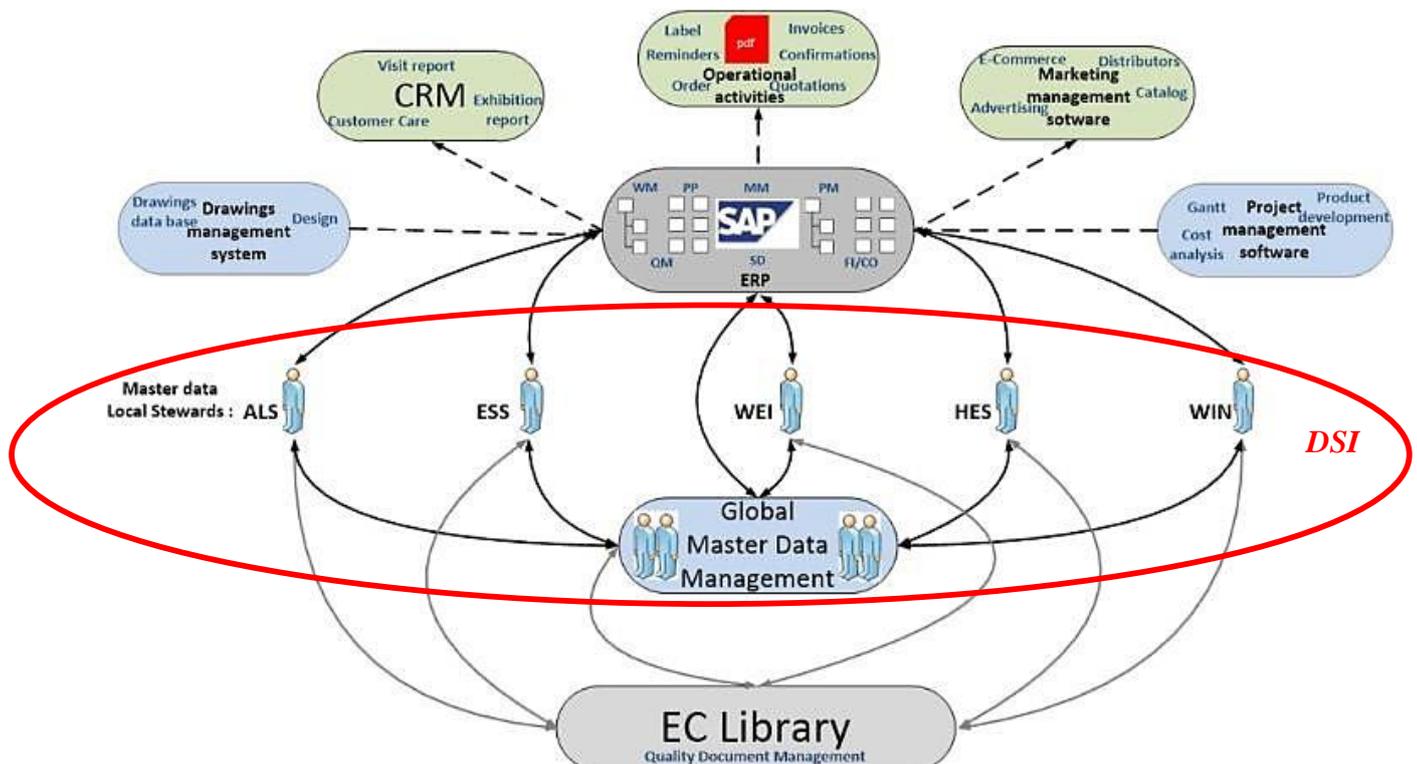


Figure 3. Organisation Master Data et l'écosystème SI de la division Stäubli EC [Stäubli EC]

J.Kobus s'est intéressé dans ses travaux à la définition et à la conceptualisation du Lean IT. Voici sa définition du Lean IT : "Lean IT is a holistic management system based on philosophy, principles and tools. Its aim is to systematically management continuous improvement by reducing waste and variability as enhancing value and flexibility in all function of an IT organization" [Kobus, 2016].

Bell et Orzen, dans leurs ouvrages « Lean IT : Enabling and sustaining your Lean transformation » [Bell et Orzen, 2011] se basent sur les cinq principes du Lean établis par Womack et Jones pour définir les cinq principes du Lean adaptés à l'activité IT (Tableau 1).

Tableau 1. Comparaison des principes du Lean et du Lean IT [Ferreira et al, 2018]

	<i>Lean (Womack and Jones, 1994)</i>	<i>Lean IT (Bell and Orzen, 2011)</i>
<i>1st Principle</i>	Specify Value: Set value according to the customer's perspective, for a specific product, with specific resources and over a period of time.	Foundation: Set value explaining a purpose through balanced leadership, with the participation of employees (respect) in order to contribute to the continuous improvement of the whole system (perfection).
<i>2nd Principle</i>	Identify the value stream: Identify the entire value stream for each product or family product and find ways to eliminate waste.	Behaviour: Invest in proactive continuous improvement, solving problems through discipline and responsibility.
<i>3rd Principle</i>	The value stream: Create value stream, ensuring that production takes place in a continuous stream. Respond to the real needs of workers in any situation.	Perspective: Focus on customer requirements, though an involvement with these. All imperfect work is removed and is not sent to the next stage. There should be a clear understanding of the overall value stream.
<i>4th Principle</i>	Flow: To draw and provide what the customer wants. Reversal of the productive flow; let the customer <i>pull</i> value.	Flow: The system runs according to Just-In Time. Waste elimination allows the connection between activities (pull mechanism).
<i>5th Principle</i>	Pull and perfection: Looking for better results, trying to achieve perfection and continually removing wastes as soon as they are identified.	Capstone: Maintain the values through attitudes and behaviours. The company is nothing more than the collective capacity of creating value.

De la même manière Ferreira et al en 2018 synthétisent la vision de différents auteurs de recherche sur le Lean IT, de l'adaptation des Mudras du Lean dans l'environnement IT (tableau 2).

Cette présentation bilatérale des Mudras du Lean IT, selon la relation par rapport au SI est particulièrement intéressante. En effet elle sépare en deux les Mudras du Lean IT, d'un côté vu par les développeurs informatiques et de l'autre ceux auxquels les utilisateurs de système d'information sont confrontés.

Les principaux ouvrages portant sur le Lean IT, « Lean IT : Enabling and sustaining your Lean transformation » [Bell et Orzen, 2011], « La pratique du Lean management dans l'IT » [Ignace et al, 2012] ou encore « The Lean IT Expert » [Loader, 2018] portent sur la pratique du Lean dans le domaine de l'informatique. La littérature se focalise sur l'application de la philosophie Lean à travers un regard d'informaticien (développement, réseaux ...).

Tableau 2. Mudras du Lean vs Lean IT [Ferreira et al, 2018]

Lean Mudras	Lean Thinking	Lean IT	
		<i>Developer views</i>	<i>User views</i>
<i>Inventory</i>	Excessive products in stock implying the existence of outdated products.	Programming unnecessary functions. Creation of software code without understanding the needs of the customer.	Excessive information causing unnecessary research, excessive delays and accumulation of work.
<i>Over production</i>	Excessive production of goods and services compared to what is needed, so production does not keep up with market demand.	Production of software code in advance.	Too many emails, reports, unread system alerts. Excessive data processing to meet customer needs. Duplication of information.
<i>Waiting / Delays</i>	Execution of the works on hold, for resources or for decision.	Lack of work for many factors, such as delay between code creation and testing or waiting for documents.	Unavailability of the system or its slowness. Time wasted waiting for additional information.
<i>Transports</i>	Unnecessary transport of materials from one place to another.	Transfer work from one team to another.	Transfer of information through multiple intermediaries and through multiple systems. Security barriers in the flow of information.
<i>Over processing</i>	Adding excess value without the client requesting, that is, doing more work than the client wants.	Development that does not add value directly to the end user.	Redundant data, unnecessary transaction and reporting, software features that users do not need.
<i>Motion</i>	Any motion according to the activities performed, that does not add value.	Exchange between team members who do not create value to the process.	Unnecessary individual work activities, including searching for tools and information, writing data, frequently changing priority.
<i>Rework/ Defects</i>	Defects that require corrections, reprocessing of work already done.	Software errors that need to be fixed. Bad software code.	Information that is incorrect, premature or confusing

En 2020 la majorité des utilisateurs et gestionnaires de système d'information en entreprise n'ont pas ou peu de connaissances ou de formations en informatique. Il s'agit dans ces travaux de montrer cette autre facette du Lean IT, mise en pratique dans les activités opérationnelles des entreprises industrielles : l'utilisation et la gestion de système d'information.

Une vision du Lean IT destinée aux utilisateurs et à la DSI permettant de gagner en performance dans les interactions processus opérationnelle/système d'information. Les processus administratifs sont souvent les plus concernés et ils représentent souvent un grand vivier d'amélioration de la performance des entreprises.

Le Lean est défini par une philosophie et un ensemble de pratiques et d'outils. La difficulté consiste à les mettre en œuvre de la bonne manière et au bon moment. « Il n'existe pas de modèle de réussite « prêt à consommer » pour appliquer les concepts du Lean » [Liker, 2004]. Pour cela beaucoup de travaux s'intéressent à la mise en pratique du Lean et à la définition d'un mode opératoire pour son déploiement.

C'est ainsi que les travaux de Hoppmann ou Reinhard portent sur la mise en place d'une roadmap de déploiement dans le premier cas issue d'une analyse des pratiques industrielles du Lean Product Development et dans le second, issue d'une mise en application industrielle du Lean Office [Hoppmann, 2009] [Reinhard, 2017].

3 APPLICATION CHEZ STÄUBLI EC SAS

3.1 Stratégie utilisée

Nous nous focaliserons maintenant à l'étude de la démarche mise en place au sein de l'entreprise Stäubli EC SAS.

La méthodologie scientifique utilisée pour cette mise en application est issue d'une approche croisée basée à la fois sur la méthode « Case study » [Yin, 2008] benchmark et analyse des bonnes pratiques industrielles et d'une démarche issue de la « Grounded theory » [Charmaz, 2003] reposant sur une approche issue de la bibliographie.

Pour déployer une démarche Lean dans l'activité opérationnelle de la phase amont du cycle de vie produit avec comme objectif d'améliorer la performance industrielle du site de Héringue, deux approches ont été mises en place :

- Une approche *Bottom Up* :

Pour cela il s'agit d'établir dans l'activité opérationnelle de l'entreprise, les bonnes pratiques et les outils d'amélioration continue. Le but étant d'instaurer cette mentalité à l'ensemble des collaborateurs par l'utilisation quotidienne et standardisée de méthodologie et d'outils issus de l'univers Lean.

Cette démarche n'est pas aisée, car il est difficile de changer les pratiques dans un environnement administratif, la résistance au changement est toujours importante. A titre d'exemple : la recherche des gaspillages dans les processus ou les pratiques administratives s'avère particulièrement délicate. Facilement ressentie comme de la surveillance ou une intrusion dans l'intimité professionnelle du collaborateur.

La première partie de cette approche *Bottom Up* est basée sur la communication :

Mise en place d'un rituel hebdomadaire ayant pour but de réunir l'ensemble d'un pôle afin d'évoquer son activité.

Ainsi un « Weekly Meeting » a été instauré, il s'agit d'un point hebdomadaire de 30 min regroupant l'ensemble d'un pôle (Responsable de pôle, Responsables de service et collaborateurs). L'objectif étant d'avoir l'ensemble de l'effectif réuni une fois par semaine afin de communiquer sur l'activité opérationnelle. Un tableau de communication a été mis en place afin d'apporter un support à ce point hebdomadaire. Ce tableau de communication permet de laisser une trace des éléments évoqués durant le rituel et de permettre à l'ensemble des collaborateurs d'y avoir accès tout au long de la semaine.

Le second moyen de développer cet état d'esprit est d'utiliser au quotidien les démarches de résolution de problèmes et d'outils du Lean afin de les rendre familiers auprès des collaborateurs. L'objectif de cette manœuvre est de les faire entrer dans les mœurs des collaborateurs et donc par extension dans ceux de l'entreprise.

Ainsi l'utilisation systématique de démarches de résolution de problèmes et d'outils du Lean (A3, PDCA, DMAIC, SIPOC, VSM, ...) a été faite lors de groupes de travail pour les projets d'amélioration continue et autres diverses réunions.

L'utilisation de ces outils sert de vecteur de déploiement des méthodologies auxquels ils sont associés.

L'objectif de l'approche *Bottom Up* est de rendre légitimes et incontournables les pratiques Lean au sein de l'entreprise par leur utilisation quotidienne par les collaborateurs.

- Une approche *Top Down* :

Avec la naissance de ces nouvelles pratiques tournées vers l'amélioration continue, il fallait également faire appel à la Direction afin de les soutenir, mais également pour leur donner un cadre et les piloter.

Jusqu'à présent les projets d'amélioration continue étaient répartis et affectés aux différents pôles de l'entreprise bien qu'ils soient transverses. Cela posait des problèmes de communication, de priorisation et d'implication.

Pour contrer ces différents problèmes et accompagner ces projets, Stäubli EC Héringue a mis en place en 2020 un plan d'action dédié à l'amélioration continue. Ce plan d'action, directement géré par la Direction, a pour but de piloter les projets d'amélioration continue du site.

Voici le rôle de ce plan d'action :

- Limiter les initiatives isolées et non coordonnées.
- Piloter l'Amélioration Continue :
Lister les projets en cours, gérer les priorités et les ressources, définir les prochains projets (impact/gain).
- Suivre l'avancement des projets et leurs résultats.
- Impliquer : responsabiliser les acteurs des projets.
- Informer l'ensemble de l'entreprise des actions en cours.
- Instaurer une culture de l'amélioration.

L'objectif d'une approche croisée *Bottom Up - Top Down* : développer une culture Lean dans l'ensemble de l'entreprise, déployée par les collaborateurs, soutenue et poussée par la Direction.

3.2 Les projets

La mise en place d'une culture Lean constitue un premier pas, il s'agit maintenant d'utiliser cette dynamique pour améliorer la performance des processus opérationnels et de l'alignement des SI de l'entreprise.

Le site de production de Héringue exerce le métier de concepteur-assembleur, en effet quasiment aucun composant n'est produit sur le site, l'activité se concentre sur le cœur de métier à savoir le développement de solutions de connexions électriques et leurs mises en œuvre (assemblage).

Aujourd'hui le principal objectif de l'entreprise est de réduire le lead time moyen lors du lancement d'un nouveau produit (temps entre la réception de la commande et la livraison client) long de quelques semaines et variant selon la typologie de produit, pour un temps de passage des produits physiques dans l'entreprise de quelques jours seulement.

Si nous reprenons la problématique industrielle, présentée plus haut : *Comment livrer des produits de plus en plus complexes dans des délais toujours plus courts ?*

Ce défi se focalise sur la phase amont de la supply chain (flux d'information avant le flux matière). C'est en effet, après analyse dans ce périmètre que se concentre le plus grand vivier d'améliorations. Notre travail va donc principalement porter sur les processus opérationnels administratifs de l'entreprise.

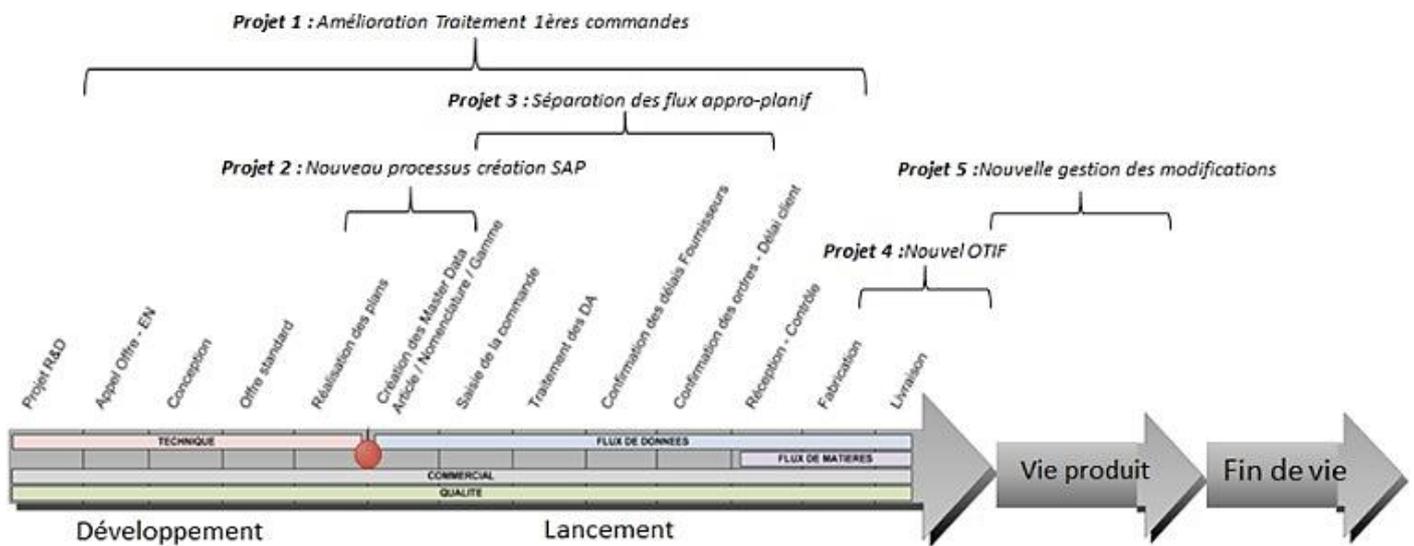


Figure 4. Disposition projets par rapport au cycle de vie produit [Stäubli EC]

L'objectif principal étant d'optimiser les processus opérationnels et l'interaction avec les SI afin de gagner en performance opérationnelle et donc de gagner en efficacité (réduction du Lead Time), mais également en efficacité (réduire les tensions et optimiser l'énergie mise dans chaque action).

L'objectif secondaire est d'instaurer une culture Lean au sein de l'entreprise et d'insuffler à l'ensemble des collaborateurs cette vision tournée vers la création de valeur.

Pour se rapprocher de ces objectifs, différents projets ont été lancés. Présentés dans la figure 4 ces projets portant principalement sur les premières phases du cycle de vie produit.

Projet n°1 : Amélioration du traitement des 1ères commandes

Objectifs :

- Réduire le temps de traitement administratif des 1ères commandes et améliorer l'efficacité du processus.
- Analyse de la situation actuelle et identification du flux de valeur. Identifier les actions à non-valeur ajoutée. Conception du processus idéal et mise en place d'un plan d'action structuré pour y arriver.

Défi :

- Utilisation d'outils nouveaux (VSM, SIPOC, ...).
- Recherche de la valeur ajoutée, basée sur le fonctionnement recherché et non les outils existants.

Projet n°2 : Nouveau processus de création SAP

Objectifs :

- Digitaliser le processus de création des articles SAP et standardisation des pratiques sur tous les sites de la division EC de Stäubli.

Défis :

- Standardisation des processus des différents sites du Groupe avec un outil commun.
- Optimisation du processus opérationnel et alignement des SI associés.
- Digitalisation d'un processus existant.

Projet n°3 : Séparation des flux produits

Objectif :

- Canaliser les flux par typologie de produit pour optimiser leurs gestions et dynamiser la communication interne dans la phase amont de la Supply chain. Créer des binômes gestionnaire approvisionnement-planification / commerciale pour chaque typologie de produit.

Défis :

- Identification les typologies et flux produits.
- Formation les collaborateurs (polyvalence).
- Communication interservices.

Projet n°4 : Nouvel indicateur OTIF

Objectif :

- Mettre en place un nouvel indicateur de performance de livraison client OTIF (On time, in full) plus fiable.

Défis :

- Définition d'un indicateur SMART et facilement maintenable.

Projet n°5 : Nouvelle Gestion des modifications

Objectif :

- Etablir un nouveau processus de gestion des modifications permettant de gagner en rapidité, sécurité et efficacité.

Défis :

- Création d'un processus transverse à la fois standardisé et adapté aux besoins.
- Définition de l'organisation et des SI nécessaires

3.3 Premiers constats

La méthodologie mise en place dans ces projets répartis sur l'ensemble de la phase amont du cycle de vie produit est principalement axée sur l'optimisation de la chaîne de valeur via la réduction des gaspillages : élimination des Mudras du Lean. Le tableau 3 montre pour chaque projet les différents Mudras identifiés par les équipes projets.

Chacun de ces projets et les actions qui en découlent reposent sur la volonté d'améliorer l'efficacité, dans notre cas, la réduction du Lead Time. Pour arriver à ce résultat, il est impératif de travailler sur l'efficacité de la structure en optimisant le ratio énergie / résultat et les ressources nécessaires.

La principale difficulté rencontrée est l'identification de la chaîne de valeur et des Mudras durant la phase amont du cycle de vie produit. Cette difficulté repose sur le simple fait que l'on intervient sur des éléments immatériels difficiles à observer.

Déterminer les actions à valeur ajoutée à travers des processus administratifs et des systèmes d'information est une tâche complexe. En effet, la vision des acteurs peut être formatée par les contraintes techniques et organisationnelles, la séparation entre actions à valeur ajoutée et celle requise par le fonctionnement sont minces.

Tableau 3. Récapitulatif des Mudras du Lean identifiés lors des projets [Stäubli EC]

Mudas du Lean	Projet n°1 : Amélioration du traitement des tâches commandées	Projet n°2 : Nouveau processus création SAP	Projet n°3 : Séparation des flux produits	Projet n°4 : Nouvel indicateur OTIF	Projet n°5 : Nouvelle Gestion des modifications
Non qualité, défauts	X	X		X	X
Surproduction			X		
Temps d'attentes	X	X	X		X
Déplacement, transports inutiles	X	X	X		
Stocks			X		X
Gestes, mouvements inutiles	X	X			X
Processus excessif	X	X		X	X
Potentiel humain non-exploité	X		X		X

Outre les Mudras, le Lean a identifié deux autres facteurs de non-valeur ajoutée à prendre en compte, car ils influencent la performance des organisations :

La variabilité (Mura), influe fortement sur la performance des organisations, des processus et outils. En effet pour optimiser un ensemble d'activités, il est impératif que celles-ci soient stables. Cette stabilisation passe donc par une standardisation des activités opérationnelles. La difficulté de cette étape est de ne pas créer de la non-valeur ajoutée en imposant des actions non-nécessaires.

La surcharge (Muri) au niveau de la phase amont du cycle de vie produit est difficile à appréhender, car souvent temporaire et ciblée. Ces goulets d'étranglement au sein de processus opérationnel se montrent complexes à identifier, il est en effet difficile de déterminer leurs causes (problèmes organisationnels, structurels, ou conjoncturels). Le défi pour l'entreprise est de pouvoir mesurer cette charge de travail avec des indicateurs factuels permettant de s'adapter aux besoins. Sans cela la notion de surcharge dans des activités immatérielles reposant sur l'Humain est liée à l'interprétation.

La méthodologie utilisée pour améliorer la chaîne de valeur pour les projets décrits ci-dessus est présentée dans la figure 5.



Figure 5 : Pyramide de déploiement

Cette méthodologie de création de processus se déroule en quatre étapes consécutives :

- Définir la stratégie pour répondre au besoin.
- Mettre en place l'organisation.
- Etablir le processus.
- Définir les outils.

Naturellement, la tendance pour résoudre une problématique est de mettre en place un outil (SI dans notre cas) ou de modifier l'existant pour y répondre.

La méthodologie proposée permet une démarche globale pour aligner le besoin aux actions mises en œuvre pour le combler. Cela permet également de s'affranchir de contraintes techniques liées aux outils (standards Groupe, customising coûteux ...), car ce sont avant tout le processus et l'organisation qui répondent au besoin.

On en déduit donc qu'une approche purement Lean IT reposant sur l'amélioration des SI n'est pas envisageable pour répondre à notre problématique. Une approche multifacette du Lean est nécessaire, moins technocentrée sur les SI et ouverte sur les différents environnements présents dans la phase amont du cycle de vie produit.

En effet, la phase amont du cycle de vie produit fait intervenir différents environnements opérationnels et stratégiques interconnectés. Une démarche Lean globale, utilisant les principes et outils adaptés aux problématiques industrielles, paraît plus appropriée pour gagner en performance (figure 6).

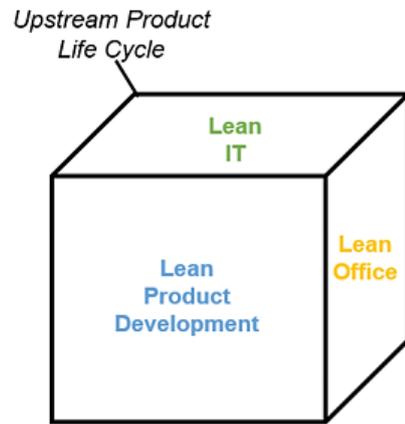


Figure 6 : Les composants du Lean dans la phase amont du cycle de vie produit

4 CONCLUSIONS ET PERSPECTIVES

4.1 Conclusions

Face à la complexification des contraintes liées aux processus et aux caractéristiques mêmes des produits, il convient de chercher des niveaux de performance encore plus hauts pour pouvoir rester compétitif.

Dans ce cadre d'étude, l'amélioration des performances industrielles de la phase amont du cycle de vie produit (du lancement de la conception jusqu'aux premières livraisons chez le client), sujet sur lequel nous nous sommes tout particulièrement focalisés, passe par l'optimisation à la fois des activités liées au développement produit, des tâches administratives et des supports de transfert d'informations.

Pour atteindre des objectifs de performance plus ambitieux lors de ces phases amont de cycle de vie, il convient de travailler sur l'écosystème d'information, et l'optimisation de ce dernier.

L'étude de l'état de l'art a pu montrer, notamment, un manque d'outils et de méthodes opérationnels relatif à cette thématique contrairement à l'aspect stratégique déjà étudié. Face à une digitalisation croissante des activités des entreprises dans ces phases amonts du cycle de vie, il convient donc de disposer d'outils et de processus appropriés d'un point de vue opérationnel. Dans une approche croisée (Top Down - Bottom Up), en analysant les Mudras du Lean instanciés dans un cas

d'étude industriel, les travaux présentés dans cet article démontrent, à travers la mise en œuvre de différents projets, le besoin pour atteindre notre objectif, d'une approche Lean globale qui compile les principes et les outils du Lean IT, du Lean Product Development et du Lean Office pour arriver à un niveau de performance supérieur et améliorer l'efficacité des processus de l'organisation.

4.2 Perspectives de recherche

L'ordre de traitement des projets présentés dans cette étude a été imposé par l'activité opérationnelle, l'utilisation des principes et des outils du Lean et découle dans notre cas de besoins industriels.

La prochaine étape consiste à analyser les principes et les outils du Lean IT et afin de les adapter aux besoins des utilisateurs opérationnels des SI et des processus liés.

L'objectif de ces travaux, à terme, est d'instancier de manière réfléchie en fonction de la maturité de l'organisation, la stratégie de déploiement des principes et outils du Lean IT tout en gérant la dette technique de l'entreprise [Caseau, 2020] et en intégrant une approche robuste de déploiement, en se basant notamment sur un retour d'expérience intégrant des concepts issus de l'Intelligence Artificielle dans une approche systémique [Mignot, 2019].

Il n'existe pas de méthodologie de déploiement universelle de cette approche Lean intégrée, mais il s'agit de la construire pour une typologie d'entreprise ayant des caractéristiques et des besoins proches. Dans notre cas, un site industriel et commercial faisant partie d'un Groupe et possédant un mix produit important ainsi que différents modes de gestion (Make to Order/Stock) donc avec un niveau de contraintes important.

5 REFERENCES

- Adams, R., Latimore, D., Wong, D., (2003). Business and IT operational models in financial services: Beyond strategic alignment. *IBM Institute for Business Value study*.
- Bell, C, S., Orzen, M, A. (2011). Lean IT: Enabling and Sustaining your Lean Transformation. *CRC Press*.
- Berrah, L. (2013). La quantification de la performance dans les entreprises manufacturières : de la déclaration des objectifs à la définition des systèmes d'indicateurs. *Université de Savoie*.
- Bernat, JP., Marcon, C. (2001). Cycle de vie et courbe d'apprentissage de produits complexes : le cas des outils coopératifs de management des connaissances. *VSSST 2001, Barcelone, Espagne*.
- Caseau, Y. (2020). L'approche Lean pour la transformation digitale. *Edition Dunod*.
- Charmaz, K. (2003). Handbook of interview research: context and method. *J. F. Gubrium & J. S. Holstein (Eds.)*.
- Cooper, R. G. (2001), Winning at New Products: Accelerating the Process from Idea to Launch. *3rd edition. Perseus Books, Reading, MA*.
- Elrhanimi, S., El Abbadi, L., Abouabdellah, A. (2015). Proposition d'un tableau de bord pour l'évaluation de l'impact du Lean manufacturing sur la performance globale de l'entreprise. *Xème Conférence Internationale : Conception et Production Intégrées, Tanger, Maroc*.
- Ferrioli, M. (2010). Phases amont du processus d'innovation : proposition d'une méthode d'aide à l'évaluation d'idées. *Institut National Polytechnique de Lorraine, France*.
- Ferreira, C. L., Almeida, M. H., Grilo, A. (2018). Lean IT adoption: success cases in Portuguese banks. *In Proceedings of the International Conference on Industrial Engineering and Operations Management, IEOM 2018*.
- Galichet, G. (2018). Proposition d'une méthodologie de transformation organisationnelle pour une performance industrielle et de la santé au travail. *Université Paris-Saclay, France*.
- Gmati I., Nurcan, S. (2007). Un cadre de référence pour analyser les exigences d'alignement métier/système d'information. *Ingénierie et gestion des processus d'entreprise. France*.
- Hamadmad. H. (2017). Définition d'une expression temporelle de la performance des entreprises manufacturières. *Gestion et management. Université Grenoble Alpes, France*.
- Henderson, J. C., Venkatraman, N. (1999). Strategic Alignment: Leveraging information technology for transformation organizations". *IBM Systems Journal, vol 38. N° S2&3*.
- Hoppmann, J., (2009). The Lean Innovation Roadmap – A Systematic Approach to Introducing Lean in Product Development Process and Establishing a Learning Organization. *Institute of Automotive Management and industrial Production Technical University of Braunschweig*,
- Ignace, M, P., Ignace, C., Médina, R., Contal, A. (2012). « La pratique du Lean management dans l'IT ». *Pearson*.
- Kobus, J. (2016). Demystifying Lean IT: Conceptualization and Definition. *Multikonferenz Wirtschaftsinformatik (MKWI)*.
- Liker, J, K. (2004). The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer. *McGraw-Hill Edition*.
- Loader, N. (2018). The Lean IT Expert: Leading the Transformation to High Performance IT. *Productivity Press*.
- Mignot, O. (2019). La transformation digitale des entreprises. *Edition Maxima*.
- Reinhard, E. (2017). Contribution méthodologique à l'introduction du lean office dans un service support de gestion des approvisionnements : analyse longitudinale par étude de cas dans une entreprise fournisseur du secteur de la santé. *Gestion et management. Université de Strasbourg*.
- Scheer, A, W., Nüttgens, M. (2000). ARIS Architecture and Reference Models for Business Process Management. *Institut für Wirtschaftsinformatik, Universität des Saarlandes, Im Stadtwald Geb. Saarbrücken*.
- Ughetto, P. (2012). Le lean : pensée et impensé d'une activité sans relâchement. @ctivité, Association Recherches et Pratiques sur les Activités, 9 (2).
- Weill, P., Broadbent, M. (1997). Management by Maxim: How business and IT managers can create IT infrastructures. *MIT Sloan Management Review, 1997. 38*.
- Wonack, J, P., Jones, D, T. (1996). Lean Thinking: Banish Waste and Create Wealth in your Corporation. *Taylor & Francis*.
- Yin, R, Y. (2008). Case Study Research: Design and Methods, *Fourth Edition, Applied Social Research Methods Volume 5, Sage Publications Incorporated*.

Rôle de la conduite du changement dans la mise en place des systèmes de management

MALIKA ALAOUI MDAGHRI¹, FATIMA BENNOUNA², DRISS AMEGOUZ³

¹ Univ. Sidi Mohamed Ben Abdellah
Fès, Maroc
malika.alaoui.mdaghri@gmail.com

² Univ. Sidi Mohamed Ben Abdellah
Fès, Maroc
fatima.bennouna@usmba.ac.ma

³ Univ. Sidi Mohamed Ben Abdellah
Fès, Maroc
driss.amegouz@usmba.ac.ma

Résumé – Que ce soit lors de la mise en place d'un système de management, ou lors de son maintien, les différents acteurs du système de management sont constamment amenés à gérer des changements, qu'ils soient planifiés, subis ou pilotés, et ce dans l'objectif d'atteindre les objectifs escomptés du système, qui s'inscrivent dans la vision stratégique de l'organisme. Face à l'absence de méthodologie de pilotage de changement pour la mise en place de système de management, nous avons analysé les liens entre les exigences des normes ISO relatives aux systèmes de management, et entre la conduite de changement. Cet article a, dans un premier temps, pour premier objectif de démontrer l'appétence des exigences ISO relatives aux systèmes de management à initier un changement et/ou à l'accompagner. Dans un second temps, ce travail de recherche vise à définir une méthodologie de pilotage de changement en fonction des étapes de mise en place d'un système de management et en fonction des différents profils rencontrés dans les organismes.

Abstract – Whether it is during the implementation of a management system, or during its maintenance, the various actors of the management system are constantly required to manage changes, whether they are planned, undergone or managed, and this in the objective of achieving the expected system objectives, which are part of the organization's strategic vision. Faced with the lack of a change management methodology for the implementation of a management system, we analyzed the links between the requirements of ISO standards relating to management systems, and between change management. The first objective of this article is to demonstrate the appetite of ISO requirements relating to management systems to initiate a change and / or to support it. Secondly, this research work aims to define a change management methodology according to the stages of setting up a management system and according to the different profiles encountered in the organizations.

Mots clés - (5 au maximum) **Système de Management, Conduite de changement, Méthodologie, Pilotage**
Keywords – **Management system, Change Management, Methodology**

1 INTRODUCTION

De plus en plus d'entreprises décident de franchir le pas de la mise en place d'un système de management certifié. En 2019, selon l'étude ISO, 1 357 241 certificats de systèmes de management ont été délivrés. La mise en place d'un système de management est généralement motivée par des objectifs d'atteinte de performances, qu'elles soient des performances de qualité, des performances environnementales, énergétiques ou autres [Rolland, 2009] ... Néanmoins, la mise en place d'un système de management peut aussi être justifié par un besoin organisationnel, une restructuration des processus [Lambert et Ouédraogo, 2010]. Au fil des années, nous avons assisté à l'évolution des normes ISO qui définissent les exigences liées aux systèmes de management. Pour rappel, selon ISO, un système de management est défini comme étant l'ensemble des processus par lesquels un organisme gère les éléments corrélés ou en interaction de ses activités afin d'atteindre ses objectifs. De plus en plus d'entreprises adhèrent à la démarche de mise en place un système de management [Boiral, 2002]. Face à un contexte qui ne cesse de changer et face à la nécessité d'intégrer le changement rapidement et efficacement, nous nous sommes penchés sur le rôle des systèmes de management dans la conduite des différents changements. Comment un système de management peut faciliter le processus de conduite de changement ? est ce qu'un système de management peut être considéré comme un outil de conduite de changement ? comment conduire le changement en se basant sur les exigences des normes ISO relatives aux systèmes de management ? afin de répondre à ces questions, cet article a été scindé en 5 sections. La première section traitera des évolutions qu'ont connues les normes ISO traitant les systèmes de management. La deuxième section soulignera comment la conduite de changement a été intégrée dans les évolutions des versions desdites normes ISO. La troisième section mettra en exergue les exigences des normes ISO permettant de conduire le changement. La quatrième section, quant à elle, décrira la méthodologie proposée pour piloter le changement à travers la mise en place d'un système de management. Et finalement, la cinquième section présentera les perspectives qu'offre ce travail pour une recherche plus approfondie et poussée sur le sujet. Afin de faciliter la lecture de cet article, nous utiliserons le terme « Normes ISO » pour désigner l'ensemble des normes ISO définissant les exigences relatives aux systèmes de management.

2 EVOLUTION DES NORMES ISO

La norme ISO 9001 a été publiée la première fois en 1987. A date, nous retrouvons 5 versions de cette norme, à savoir celle de 1987, celle de 1994, 2000, 2008 et enfin 2015. Chaque version de cette norme a connu des concepts clés que la figure 1 reprend :

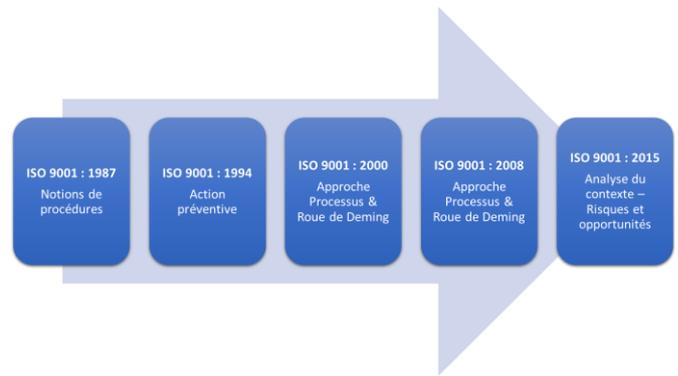


Figure 1. Evolutions de la norme ISO 9001

La figure 1 illustre l'évolution de la norme ISO 9001 relative aux systèmes de management de la Qualité (SMQ).

Les évolutions de cette norme nous permettent de conclure que les versions des normes mettent moins en exergue le côté documentaire du système de management (procédures, manuel qualité ...) et œuvrent plus à la promotion d'une approche de prévention, de proactivité et d'amélioration continue, et donc de changement continu aussi. Les principales évolutions concernent les concepts clés suivants :

2.1 Connaissances organisationnelles

Nous remarquons l'intégration d'un chapitre « Connaissances organisationnelles » (§7.1.6) dans la norme ISO 9001 version 2015, faisant référence au phénomène d'apprentissage organisationnel.

L'apprentissage organisationnel étant défini comme « processus de création de connaissances, de distribution de celles-ci au sein de l'organisation et de leur inscription dans les pratiques » [Bootz et Monti, 2008].

2.2 Représentant de la Direction : un temps révolu

Finie la notion du représentant de la Direction, celle-ci est intégrée dans le chapitre Leadership. Cette nouveauté annonce une nouvelle approche de mise en place et de déploiement d'un système de management : ce dernier n'est plus porté par un représentant, mais doit être porté par un leadership commun émanant de la Direction, mais aussi d'une approche participative où tous les acteurs d'un système se doivent de le porter et de le faire vivre.

2.3 Analyse du contexte

L'un des changements phares des nouvelles versions est l'analyse du contexte, qu'il soit interne et externe, et le lien établi avec les risques et opportunités. Par ce biais, la norme intègre la nécessité d'adopter une approche proactive, et de gérer le changement à travers des risques maîtrisés (changement géré) ou des opportunités saisies (changement suscité).

2.4 Structure HLS

Avec l'avènement de l'intégration des systèmes de management, ISO a développé la structure HLS (High Level Structure) et une annexe SL, permettant d'intégrer au moins deux systèmes de management [Tangen et Warris, 2012]. Selon Kopia, cette nouvelle structure a facilité l'intégration des systèmes de management [Kopia, 2016]. Ceci permet donc aux entreprises d'intégrer leurs systèmes de management à travers des processus régis par une seule politique [Domingues et al., 2017]. Cette structure permet de définir un cadre commun à

toutes les normes régissant les systèmes de management. Elle est applicable à la fois aux nouvelles normes ainsi qu'aux prochaines révisions des normes existantes. Les normes ISO ayant des exigences communes grâce à la structure HLS, nous analyserons ces exigences en se basant sur la structure de la norme ISO 9001 : 2015.

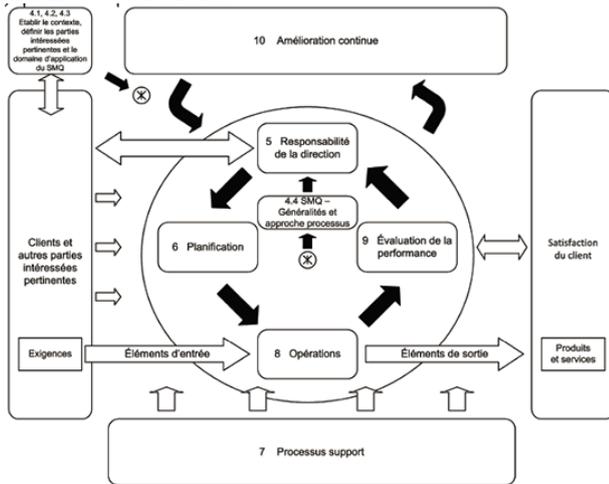


Figure 2. Schéma de la structure HLS

au contexte et à son évolution regroupant la mise en place d'un système sur 4 grandes rubriques, telles que décrites par Lecerf-Thomas [Lecerf-Thomas, 2009] :

- 1- La vision, les stratégies et le pilotage
- 2- Les procédures molles
- 3- Les procédures dures
- 4- Les modalités opérationnelles

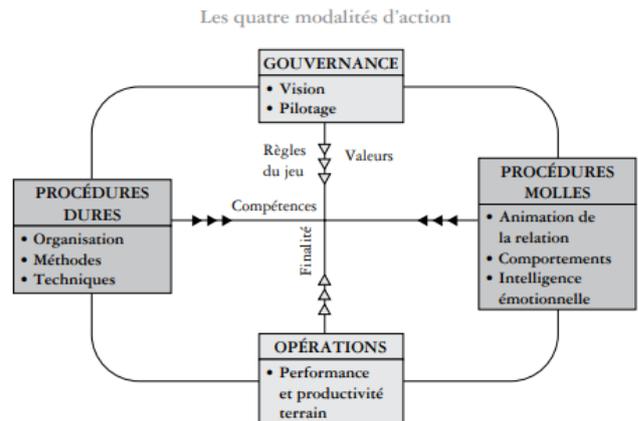


Figure 3. Les quatre modalités d'action de mise en place d'un système de management [Lecerf-Thomas, 2009]

3 INTEGRATION DU PILOTAGE DE CHANGEMENT DANS LES EXIGENCES DES NORMES ISO

Dans sa première version, la norme ISO 45 001 : 2018, relative à la Santé et Sécurité au Travail, intègre un sous-chapitre « Pilotage du changement » (§ 8.1.3) dans le chapitre « Réalisation des activités opérationnelles » (§ 8.1). Malgré la structure HLS précédemment citée, les normes ISO relatives aux systèmes de management continuent à évoluer via l'ajout de nouvelles exigences, l'exemple étant illustré avec la nouvelle norme ISO 45001 à travers son sous-chapitre « Pilotage du changement ».

Le fait que ce sous-chapitre soit inscrit dans le chapitre §8 démontre la nécessité de piloter le changement lors de la phase « DO » de la Roue de Deming. L'exigence y afférente précise la nécessité d'établir un ou des processus pour la mise en œuvre et la maîtrise des changements temporaires et permanents prévus ayant une incidence sur la performance Santé & Sécurité au Travail. Nous remarquons donc que l'exigence traite du changement dans 2 cas distincts :

- Cas 1 : Changements temporaires ayant une incidence sur la performance S&ST, caractérisé par une gestion de changement ponctuelle
- Cas 2 : Changements permanents ayant une incidence sur la performance S&ST, exigeant la nécessité de conduire un processus de changement qui peut durer dans le temps et qui doit être consolidé et maintenu après sa mise en place.

A noter que dans ces 2 cas, ces changements sont prévus et par conséquent planifiés. La norme ne traite donc que des changements anticipés, et n'intègre pas le pilotage des changements subis et imprévus.

L'exigence souligne aussi 2 étapes pour conduire le changement à savoir la mise en œuvre des changements et la maîtrise des changements.

4 EXIGENCES DES NORMES ISO EN LIEN AVEC LA CONDUITE DU CHANGEMENT

Les chapitres précédents nous ont permis de remarquer que les évolutions des systèmes de management ont tenu compte des changements des styles de management, des changements liés

L'axe vertical concerne le passage de l'intention à l'action, l'axe horizontal est l'axe des moyens d'action essentiels aux compétences ; d'un côté les procédures dures concernent l'organisation, les règles du jeu formalisées et les moyens techniques, et de l'autre les procédures molles concernent les comportements, les valeurs et les moyens relationnels. Ces quatre modalités d'action représentent des leviers permettant de piloter le changement en fonction des orientations stratégiques du top management de l'organisme. L'objectif de la méthodologie proposée est de faciliter et favoriser le passage de l'intention à l'action (axe vertical de la figure), et de capitaliser sur l'expérience à travers un apprentissage organisationnel, que ce soit via des procédures molles ou des procédures dures (axe horizontal de la figure). Pour ce, nous nous sommes basés sur les quatre modalités d'actions afin d'identifier les exigences des normes ISO permettant d'assurer une transition vers un changement géré et piloté. Nous avons donc établi une matrice regroupant les chapitres et sous-chapitres de la norme ISO 9001 : 2015, puis nous avons évalué l'appétence des exigences à initier et à accompagner le changement.

4.1 Méthodologie d'évaluation de l'appétence des exigences des normes ISO à piloter le changement

Le lien entre les exigences des normes ISO et la conduite du changement a été évalué selon une grille de notation conçue par nos soins. Elle est présentée dans le tableau 1 :

1. **Tableau 1. Grille de notation des exigences des normes ISO et leur lien avec la conduite du changement**

Critère	Notation
Lien direct avec les orientations stratégiques de l'entreprise	Oui = 5 Non = 0
Nécessité d'être porté par un pilote du changement	Oui = 5 Non = 0
Importance de l'adhésion des acteurs du système de management	L'exigence est non conforme si l'adhésion de l'ensemble des acteurs n'est pas intégrée = 10 Nécessité d'adhérer au moins une partie des collaborateurs pour le respect de l'exigence = 5 Exigence pouvant être respecté sans y impliquer l'ensemble des acteurs = 0
Type de ressources exigées	Ressources humaines et organisationnelles = 10 Ressources matérielles (technologiques et financières) = 0

6.2	Objectifs qualité et planification des actions pour les atteindre	B
6.3	Planification des modifications	A
7.1	Ressources	C
7.2	Compétences	B
7.3	Sensibilisation	B
7.4	Communication	B
7.5	Informations documentées	C
8.1	Planification et maîtrise opérationnelles	C
8.2	Exigences relatives aux produits et services	C
8.3	Conception et développement de produits et services	B
8.4	Maîtrise des processus, produits et services fournis par des prestataires externes	C
8.5	Production et prestation de service	C
8.6	Libération des produits et services	C
8.7	Maîtrise des éléments de sortie non-conformes	C
9.1	Surveillance, mesure, analyse et évaluation	C
9.2	Audit interne	C
9.3	Revue de direction	B
10.1	Généralités	C
10.2	Non-conformité et action corrective	C
10.3	Amélioration continue	A

Le tableau 1 nous permet de mettre un scoring face à chaque exigence de la norme ISO. Pour ce, nous avons classé la corrélation « exigences des normes » et « lien avec la conduite du changement » en 3 classes :

- Classe A : Lien direct ayant un fort impact sur le respect de l'exigence ;
- Classe B : Lien direct pouvant avoir un impact sur le respect de l'exigence ;
- Classe C : Lien indirect, voire inexistant.

Les résultats de cette évaluation sont décrits dans le point suivant (§4.2).

4.2 *Résultats de l'évaluation des exigences des normes ISO et leur lien avec la conduite du changement*

La grille de notation représentée dans le tableau 2 nous a permis d'évaluer l'appétence des exigences des normes ISO à servir de tremplin pour conduire le changement. Les résultats de cette évaluation ont été synthétisés dans le tableau 2 :

Tableau 2. Matrice des exigences des normes ISO et leurs liens avec la conduite du changement

N° chapitre	Nom du chapitre	Lien avec la conduite du changement
4.1	Compréhension de l'organisme et de son contexte	A
4.2	Compréhension des besoins et des attentes des parties intéressées	B
4.3	Détermination du domaine d'application du SMQ	C
4.4	SMQ et ses processus	C
5.1	Leadership et engagement	A
5.2	Politique	A
5.3	Rôles, responsabilités et autorités au sein de l'organisme	B
6.1	Actions à mettre en œuvre face aux risques et aux opportunités	A

Avec :

A : Lien direct ayant un fort impact sur le respect de l'exigence

B : Lien direct pouvant avoir un impact sur le respect de l'exigence

C : Lien indirect, voire inexistant

Les exigences faisant partie de la classe A ont donc un lien direct avec la conduite du changement. Se conformer à ces exigences signifie initier un changement, le planifier, puis le gérer en atténuant au maximum les résistances qui peuvent surgir.

La suite de l'article sera principalement focalisée sur les exigences de la classe A.

5 METHODOLOGIE DE CONDUITE DE CHANGEMENT VIA LE SYSTEME DE MANAGEMENT

Pour faire le lien avec la problématique de départ, le chapitre précédent nous a prouvé que le système de management est bel et bien un outil de conduite de changement. Ce chapitre tentera de répondre aux questions : Comment un système de management peut faciliter le processus de conduite de changement ? Et, comment conduire le changement en se basant sur les exigences des normes ISO relatives aux systèmes de management ?

Pour ce, la méthodologie proposée a été construite sur un processus de mise en place de système de management basé sur les étapes suivantes :

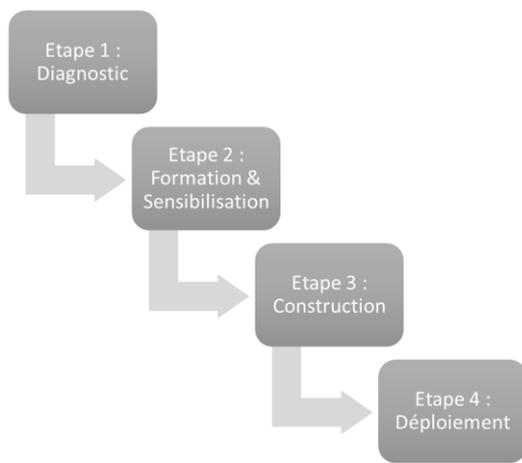


Figure 4. Etapes de mise en place d'un système de management

La méthodologie proposée se base sur les travaux de Kurt Lewin et de John Kotter pour le pilotage du changement. L'objectif de la méthodologie proposée est de trouver l'état d'équilibre du système en augmentant l'intensité et / ou le nombre des forces motrices favorables au changement, et en diminuant l'intensité des forces restrictives au changement [Burnes, 2004].

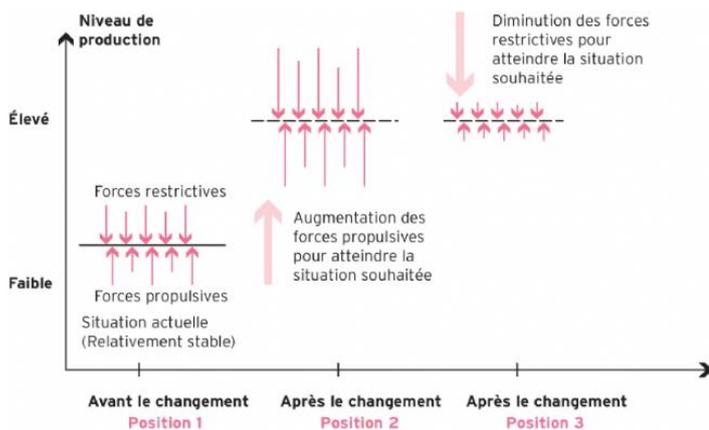


Figure 5. Modèle de changement de Kurt Lewin

Pour y arriver, nous nous sommes inspirés des 8 étapes du changement de Kotter, que l'on peut résumer à la préparation du changement, l'engagement du changement, la réalisation du changement et l'ancrage du changement [Kotter et Cohen, 2012].

La figure 5 permet de faire le lien entre les étapes du changement de Kotter et les étapes de mise en place du système de management :

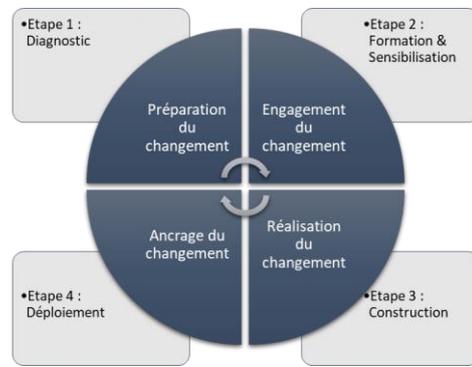


Figure 6. Lien entre les étapes du changement de Kotter et les étapes de mise en place du système de management

Afin de rentrer plus en détail dans la méthodologie de conduite de changement grâce à la mise en place d'un système de management, nous avons identifié l'ensemble des actions nécessaires à une conduite du changement efficace. Nous retrouvons ci-après le détail de la méthodologie préconisée :

5.1 Diagnostic

Initialement, cette étape a pour objet d'évaluer le niveau de conformité du système existant par rapport aux exigences des normes ISO applicables. Elle peut être portée soit par des collaborateurs internes, soit par des consultants externes à l'entreprise. En plus de l'aspect lié aux exigences ISO, cette étape constitue l'étape de base pour la construction de la démarche à adopter pour la mise en place du système. Elle a pour objectif d'identifier les différents acteurs du système, à savoir :

- Les leaders de l'entreprise : ce sont les acteurs du système qui influencent directement les orientations stratégiques de l'organisme et qui déterminent les lignes directrices vers lesquelles doit s'orienter le système de management. Ils font généralement partie du top management de l'entreprise, mais peuvent aussi faire partie de tous les niveaux hiérarchiques en cas de forte influence sur les pouvoirs (représentants du personnel, collaborateur réputé et ayant de l'ancienneté ...)
- Les résistants au changement : ce sont des acteurs exprimant des comportements d'opposition vis-à-vis de la démarche de mise en place de système de management. L'expression de l'opposition peut se faire explicitement à travers la communication verbale ou écrite, ou implicitement à travers le refus des démarches proposées.
- Les porteurs de changement : ce sont les acteurs qui sont enthousiastes à la démarche. On les reconnaît à leur force de proposition, leur curiosité par rapport à la nouveauté, leur motivation pour évoluer et développer leurs compétences.
- Les constructeurs du changement : sont ceux qui adhèrent aux démarches de construction du changement. Ils sont chargés de faire vivre le système, de le maintenir et de le déployer.

Il est à noter qu'un même acteur peut incarner différents rôles, nous pouvons retrouver par exemple des leaders qui s'opposent au changement, des porteurs du changement qui sont constructeurs aussi ...

La finalité de l'étape Diagnostic est donc d'identifier les acteurs clés du système et de déterminer le(s) rôle(s) qu'ils jouent dans le système et dans la démarche de mise en place du système.

5.2 Formation & Sensibilisation

Le premier objectif de cette étape est de former et de sensibiliser l'ensemble des acteurs du système sur les exigences des normes ISO prévues d'être déployées. Le second objectif consiste à :

5.2.1 Pour les leaders :

- Initier la réflexion quant au mode de gouvernance, les valeurs, et la vision souhaités dans le système de management
- Expliquer aux collaborateurs la raison de la mise en place de la démarche

5.2.2 Pour les résistants :

- Comprendre les causes de la résistance au changement
- Les conscientiser par rapport à leur potentiel de constructeur et de porteur de changement

5.2.3 Pour les porteurs :

- Initier la réflexion quant au mode de construction du système
- Les accompagner à identifier les constructeurs du changement et à les fédérer autour des différents sujets susceptibles de subir des changements

5.2.4 Pour les constructeurs :

- Intégrer l'ensemble des exigences de la norme
- Comparer les modes de travail actuels et les changements susceptibles d'interférer

5.3 Construction

5.3.1 Pour les leaders :

- Définir la vision et les valeurs que doit porter le système de management et formaliser les orientations stratégiques
- S'assurer de la compréhension des orientations stratégiques définies par les leaders

5.3.2 Pour les résistants :

- Transformer les résistants en porteurs et constructeurs du changement afin de donner du sens à la démarche et de neutraliser les causes de la résistance au changement

5.3.3 Pour les porteurs :

- Construire des équipes prêtes à piloter le changement
- Mener des brainstormings quant aux modes de construction
- Coordonner entre les constructeurs de changement
- Répondre aux différentes appréhensions des résistants et constructeurs du changements

5.3.4 Pour les constructeurs :

- Assurer la remontée d'informations en cas d'inefficacité d'outils, de procédures, ou autres
- Construire les outils et les nouveaux modes de travail

5.4 Déploiement

L'objectif du déploiement est de s'assurer que le système de management construit répond aux exigences de la norme, s'intègre aux orientations stratégiques définies et permet d'atteindre les résultats escomptés. Dans le processus de

conduite de changement, les missions de chacun des profils sont décrites comme suit :

5.4.1 Pour les leaders

- Assurer le suivi des performances du système de management
- Féliciter l'ensemble des acteurs du système pour les efforts déployés pour la construction du système
- Ne pas sanctionner les non-conformités remontées

5.4.2 Pour les résistants

Au stade du déploiement, l'objectif est d'avoir neutralisé les résistants au changement. Cependant, des signes d'opposition pourront apparaître. La conduite à tenir pendant ces étapes est l'écoute active de l'ensemble des doléances des résistants. L'écoute devra se faire par les autres profils. Par la suite, une revue de la pertinence de la doléance devra être faite. En cas de pertinence, il sera recommandé de faire un retour à la personne ayant assuré la remontée. Une des méthodes de neutralisation de la résistance de changement est aussi de transformer le résistant en constructeur en lui demandant de proposer une solution à la problématique soulevée.

5.4.3 Pour les porteurs

- Investiguer sur les causes de non atteinte des performances
- Etre à l'écoute des résistants si existants, et proposer des plans d'actions pour la neutralisation des causes de résistance de changement
- Encourager la remontée des non-conformités et des propositions d'amélioration

5.4.4 Pour les constructeurs

- Déployer les actions de mise en conformité
- Remonter les non-conformités et les propositions d'amélioration

6 CONCLUSIONS ET PERSPECTIVES

Le présent article nous permet de démontrer que la mise en place d'un système de management peut être un outil pour la conduite du changement, que ce soit dans une perspective de transition organisationnelle ou dans une perspective d'atteinte de performances.

Dans ce travail, nous avons traité le lien entre les exigences des normes ISO et la conduite du changement, et nous avons proposé une méthode pour piloter le changement en fonction des étapes de mise en place du système.

Cette méthode a été inspirée des travaux de change management initiés par John Kotter et Kurt Lewin, et sur lesquels plusieurs théories de change management ont été basés.

Ce travail de recherche permet d'initier la réflexion quant aux modes de gestion du changement grâce à une démarche itérative et systémique. Il permet aussi d'identifier les différents profils interagissant dans un système et comment capitaliser sur le potentiel de chaque profil afin d'augmenter les forces positives en faveur du changement.

D'autres pistes sont à explorer, notamment l'impact des styles de management des leaders dans le pilotage du changement ainsi que les modes de capitalisation sur les apprentissages organisationnels issus des changements, etc ...

7 REFERENCES

- Communication et Organisation*, 33, 30-39.
- Boiral, O. (2002). ISO 9000 : De l'impératif commercial au « Rite de passage » organisationnel. *Conférence de l'AIMS*.
- Boots, JP., Monti, R. (2008). Les démarches prospectives : de l'aide à la décision à la conduite du changement. *Revue Finance Contrôle Stratégie*.
- Burnes, B. (2004). Managing change : A strategic approach to organisational dynamics. *Pearson Education*.
- Burnes, B. (2004). Kurt Lewin and the planned approach to change : a re-appraisal. *Journal of Management studies*.
- By, R. T., (2005). Organisational change management : A critical review. *Journal of Change Management*. 369-380
- Domingues, P., Sampaio, P., Aarezes, P.M. (2017). Management systems integration : survey results. *Int. J. Qual. Reliab. Manag.* 34 (8), 1252-1294.
- Folke, C., Carpenter, S., Elmqvist, T., Gunderson, L., Holling, C.S., Walker, B. (2002). *Journal of the Human Environment*, 31(5), 437-440
- Lambert, G., Ouédraogo, N. (2010). Normes, routines organisationnelles et apprentissage d'entreprise. *Revue française de gestion*.
- Lecerf-Thomas, B. (2009). Neurosciences & management : Le pouvoir de changer. *Editions Eyrolles*.
- Kopia, J. (2016). Study on integration and leadership styles of management systems based on a high level structure. *International Conference on Management Leadership & Governance* , 431-441.
- Kotter, JP., Cohen, DS. (2012). The heart of change : Real-life stories of how people change their organizations.
- Rolland, S. (2009). Un bilan de 20 ans de certification des systèmes de management de la qualité : les apports perçus de la certification ISO 9000 par les managers. *Management Avenir*
- Tangen, S., Warris, A.M., (2012). Management Makeover - New Format for Future ISO Management SystemStandards.

Etudes ISO 2019, disponible sur :
<https://isotc.iso.org/livelink/livelink?func=ll&objId=18808772&objAction=browse&viewType=1>

Proposition d'une méthodologie d'amélioration continue adaptée à une situation de crise.

Application chez Safran Landing Systems pendant la crise du Covid-19

ALICIA MINOT-LENFANT^{1,2}, NICOLAS MARANZANA², ANGELINE EYMERY¹

¹ SAFRAN LANDING SYSTEMS
7 rue du Général Valérie André, 78140 Vélizy-Villacoublay, France
angeline.eymery@safrangroup.com

² ARTS ET METIERS SCIENCES ET TECHNOLOGIES, LCPI, HESAM UNIVERSITE
151 boulevard de l'Hôpital, 75013 Paris, France
nicolas.maranzana@ensam.eu

Résumé – Dans un contexte de crise, les entreprises sont face à un important défi. Elles doivent, dans l'immédiat, assurer leur survie, en se recentrant sur les activités essentielles, tout en anticipant la reprise d'activité de demain. Pour maintenir leur niveau de compétitivité les entreprises n'ont cessé de s'améliorer. Mais dans une période aussi incertaine qu'une crise, comment peut-on mener des projets d'amélioration ? Pour y répondre, nous proposons une démarche structurée et collaborative adaptée au contexte de crise, en termes de versatilité et de dispersion spatiale et temporelle des équipes. Nous démontrons la fiabilité de cette démarche par son application concrète : un projet d'amélioration mené au sein de *Safran Landing Systems* pendant la crise du Covid-19.

Abstract – In a context of crisis, companies are facing a major challenge. They must, in the immediate term, ensure their survival, by refocusing on core activities, while anticipating tomorrow's business recovery. To maintain their level of competitiveness, companies are constantly improving. However, in a period as uncertain as a crisis, how can improvement projects be carried out? To answer this question, we propose a structured and collaborative approach adapted to the crisis context, in terms of versatility and the spatial and temporal dispersion of teams. We demonstrate the reliability of this approach through its concrete application: an improvement project carried out within *Safran Landing Systems* during the Covid-19 crisis

Mots clés – amélioration continue, collaboration asynchrone, maîtrise des risques, gestion de crise

Keywords – continuous improvement, asynchronous collaboration, risk management, crisis management

1 INTRODUCTION

« Alors, étions-nous préparés à cette crise ? A l'évidence, pas assez, mais nous avons fait face en France comme partout ailleurs. Nous avons donc dû parer à l'urgence, prendre des décisions difficiles à partir d'informations partielles, souvent changeantes, nous adapter sans cesse ». Emmanuel Macron, Allocution télévisée enregistrée à l'Élysée, le 13 avril 2020.

Cette année 2020 restera marquée par la pandémie du Covid-19 qui a sévèrement ébranlé l'activité économique des entreprises, en France comme dans le monde. Le premier confinement, au printemps 2020 a mis à l'arrêt l'économie française, n'épargnant aucun secteur d'activité. Le déconfinement s'est accompagné d'une reprise loin des standards auxquels nous étions habitués. Au-delà de l'application des gestes barrières dans notre quotidien, notre vie professionnelle a également été modifiée en profondeur, avec notamment la réduction du temps de travail, via l'activité partielle mise en place grâce aux aides de l'Etat, et la généralisation du télétravail.

Si les propos du Président de la République, cités en introduction, faisaient référence au côté sanitaire de la crise, ils sont également applicables à la crise économique qui lui est désormais indissociable.

« Alors, étions-nous préparés à cette crise ? A l'évidence, pas assez [...] ». Cette crise a mis en évidence qu'une grande partie des entreprises n'était pas armée pour travailler à distance, que ce soit d'un point de vue matériel (informatique, communication, etc.) ou social (dépressions liées à l'isolement).

« [...] mais nous avons fait face en France comme partout ailleurs ». Faire face aujourd'hui pour se préparer au monde de demain, qui sera certainement plus compétitif, voilà le challenge que les entreprises doivent relever. Tout en se maintenant à flot en se recentrant sur les activités essentielles, elles doivent également anticiper la reprise d'activité. Ainsi, les projets d'amélioration continue doivent subsister et trouver de nouveaux moyens d'exister, afin d'être compatibles avec le temps qui peut leur être accordé.

« Nous avons donc dû parler à l'urgence [...] ». Les entreprises se sont en effet adaptées, dans l'urgence, pour travailler et communiquer à distance. L'utilisation de logiciels de conférence tels que Zoom® ou Teams® s'est démocratisé, tout comme la digitalisation en général. A titre d'exemple, au-delà des réunions, des audits ont également pu se réaliser à distance. En ce sens, la crise s'est révélée être une opportunité, un vecteur de changement permettant d'accélérer la digitalisation à travers l'acceptation (forcée) de nouveaux outils.

« Nous avons donc dû [...] prendre des décisions difficiles à partir d'informations partielles, souvent changeantes, nous adapter sans cesse ». Mais dans un contexte aussi versatile, comment anticiper l'avenir, comment mener malgré tous des projets d'amélioration ?

C'est précisément cette problématique, « **Comment mener un projet d'amélioration dans un contexte de crise ?** » que nous avons souhaité développer. Si plusieurs hypothèses peuvent être envisagées pour y répondre comme la variation des paramètres du triangle QCD – Qualité, Coûts, Délais – nous ne les avons pas retenues. En effet, dégrader la qualité du projet ou reporter sa mise en œuvre peuvent remettre en cause l'utilisation des livrables qui lui sont associés (qualité) ou leur pertinence (livraison trop tard). Nous avons plutôt choisi de proposer une démarche adaptée au contexte de crise pour mener à bien des projets d'amélioration. En effet, compte-tenu du contexte, ressources limitées en raison de l'activité partielle et équipes dispersées dû au télétravail, adapter la démarche de gestion de projet est essentiel pour travailler de manière efficiente. Dans cet article, après un état de l'art, nous proposerons une démarche structurée, supportée par un outil digital, permettant de prendre en compte les critères fondamentaux de réussite avant de l'illustrer par une application concrète au sein d'un grand groupe aéronautique de renommée internationale, Safran.

2 ETAT DE L'ART

L'amélioration est au cœur de la qualité et figure non seulement comme l'un des sept principes de la qualité mais également comme l'un des chapitres de la Norme ISO 9001 [AFNOR, 2015]. En effet, l'amélioration continue est une démarche essentielle pour garantir la pérennité des systèmes de management. Son succès repose sur l'utilisation d'une démarche structurée, pas à pas, dont les plus connues sont le DMAIC, pour *Define-Measure-Analyze-Improve-Control* [Shankar, 2009], et le PDCA, pour *Plan-Do-Check-Act* [Deming, 1982]. Ce dernier, également connu sous le nom de cercle vertueux de Deming [Chardonnet et Thibaudon, 2003] est composé de quatre phases distinctes. La phase *Plan* consiste à planifier les activités et à décider de la façon dont elles devront être menées. La phase *Do* permet de déployer les activités prévues dans la phase précédente. La phase *Check* vise à vérifier la conformité des réalisations par rapport aux actions prévues et à contrôler l'atteinte des objectifs par le biais des indicateurs. La phase *Act* est une étape d'analyse, de réflexion visant à prendre du recul par rapport à ce qui a été réalisé pour imaginer de nouvelles améliorations à implémenter dans un nouveau cycle. De manière intéressante, le PDCA peut également être utilisé comme un squelette sur lequel des spécificités, liées au contexte du projet, peuvent être greffées. A titre d'exemple, nous pouvons citer la démarche de digitalisation du système documentaire de Eymery et al. dans laquelle le PDCA est décliné en 6 lots de travaux incorporant une spécificité indispensable à la réussite d'un projet de

digitalisation, la conduite du changement [Eymery et al., 2019].

3 PROPOSITION D'UNE METHODOLOGIE

Pour répondre à notre problématique, nous avons choisi d'utiliser le PDCA comme méthode structurante en l'adaptant au contexte particulier de la crise en termes de versatilité d'une part et de dispersion spatiale et temporelle des équipes d'autre part (Figure 1).

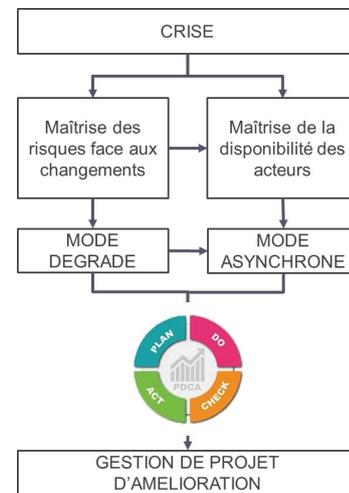


Figure 1. Proposition d'adaptation du PDCA au contexte de crise

Afin de **s'adapter aux conditions changeantes et de maîtriser les risques associés**, nous avons choisi de anticiper en mettant en place **un plan d'action en « mode dégradé »**. Cela consiste à imaginer différents scénarii possibles, ou alternatives, pour mener à bien une action, sans impacter la qualité du livrable. Ainsi, certaines exigences, ou objectifs, peuvent être redéfinis en fonction des risques identifiés afin de maîtriser l'avancée du projet et ainsi s'assurer de le terminer dans le délai imparti. De manière concrète, les différentes actions du projet sont planifiées selon un « *mode initial* », correspondant à ce qui devrait être fait en conditions « normales ». Ces modes initiaux sont alors soumis à une analyse de risques menée pour identifier un ou plusieurs « *modes dégradés* » tout en veillant à maintenir la qualité des livrables associés. La Figure 2 illustre ces propos à l'aide d'un exemple concret, l'animation d'un groupe de travail, soumis à un premier risque, la mise en place du télétravail, puis à un second en addition du premier, l'établissement de mesures d'activité partielle.

La réussite d'un projet étant basée sur son acceptation par les collaborateurs, la conduite du changement et plus particulièrement leur implication dans le projet est cruciale [Autissier & Moutot, 2010 ; Eymery et al., 2019 ; Kotter & Rathgeber, 2018]. Ainsi, même en temps de crise, la création d'une équipe projet est indispensable. Dans ce contexte, et pour faire face aux conditions de travail particulières mises en place pendant la crise (télétravail, activité partielle), notre méthodologie doit également **s'adapter à la disponibilité des ressources**. Pour cela, nous avons imaginé un « **mode de travail collaboratif, distanciel et asynchrone** », assurant une flexibilité aux collaborateurs quant à leur implication dans le projet. La communication étant un élément essentiel dans le cadre de projets collaboratifs [Maranzana et al., 2008] nous avons comparé les différents modes de communication

existants afin d'identifier des facteurs clés de succès [Hisarciklilar & Boujut, 2006 ; Hisarciklilar, 2008]. La communication verbale synchrone est bien évidemment la plus appropriée, quand les conditions le permettent. En effet, ses échanges simultanés favorisent un apprentissage collectif nécessaire pour une bonne coopération entre les membres de l'équipe. Son caractère informel et spontané favorise la compréhension partagée et permet de créer une conscience de groupe. Enfin, la communication non verbale associée est très utile pour mieux appréhender certaines paroles, notamment par l'observation des gestes, et d'éviter ainsi des malentendus. En effet, selon Mehrabian & Ferris, 93% de la communication passe par le comportement, le non-verbal [Knapp et al., 2013 ; Mehrabian & Ferris, 1967]. Avec tous ces atouts, la communication verbale synchrone correspond donc à un « *mode initial* ».

A l'opposé, la communication asynchrone, qui représente donc un « *mode dégradé* », est plus complexe à maîtriser. Les échanges se font de manière différée, formalisée et unidirectionnelle empêchant la clarification de l'information et laissant la place à l'interprétation voire à l'incompréhension. Forts de cette analyse, nous avons ainsi déterminé les facteurs clés de succès d'un travail collaboratif asynchrone. En premier lieu, il faut définir un cadre pour faciliter la coordination et le partage d'informations entre les acteurs, et les faire converger vers des objectifs communs. Il faut également garantir un même niveau de compréhension entre les participants en essayant de retrouver les caractéristiques de la communication informelle et non verbale.

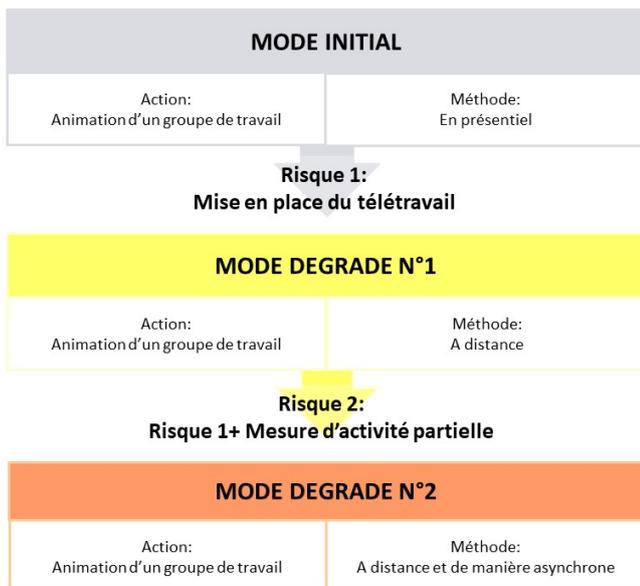


Figure 2. Illustration de la mise en place de « modes dégradés » face aux risques potentiels

Nous avons ensuite décliné ces adaptations, selon le PDCA, dans une méthodologie découpée en 6 lots de travaux (Figure 3). Ce découpage est particulièrement important dans un contexte où le temps et les ressources sont limitées car il permet de mesurer concrètement l'avancement du projet, étape par étape, et de maintenir les équipes motivées [Autissier & Moutot, 2010 ; Eymery et al., 2019 ; Kotter & Rathgeber, 2018].

3.1 Phase ACT

Nous proposons de commencer par la phase *Act*, constituée du **lot 1**, qui est indépendante de la crise. Cette phase de réflexion

consiste à analyser l'existant pour identifier les éventuels problèmes et les améliorations à mettre en place pour les éradiquer. La validation du besoin d'amélioration auprès des parties intéressées (clients finaux et commanditaires) est la seconde sous étape de ce premier lot de travail.

3.2 Phase PLAN

La phase *Plan* est constituée des lots 2, 3 et 4.

Le **lot 2** consiste à préparer le projet :

- Un plan d'action associé à un planning est mis en place, constituant le « *mode initial* ». Une analyse de risques est ensuite réalisée afin de définir, s'il y a lieu, les « *modes dégradés* », pour chacune de ces actions (Figure 2).
- Des indicateurs doivent également être mis en place afin de servir de déclencheur du « *mode dégradé* » dans le cas où une dérive serait observée.

Le premier est un indicateur d'avancement du projet, calculé comme le ratio de l'avancement de l'ensemble des tâches (pondérées en fonction du temps qu'elles nécessitent et de leur avancement respectif) par rapport à la durée du projet (indicateur i1).

Le second est un indicateur d'avancement temporel, calculé comme le ratio du nombre de jours passés sur le nombre de jours total du projet (indicateur i2).

La comparaison de i1 par rapport à i2 permet d'identifier et d'analyser les retards éventuels du projet. Si cela se produit 3 solutions sont envisageables :

(1) Si la tâche en cours de réalisation possède un « *mode dégradé* », ce dernier est activé afin de ne pas accumuler plus de retard. En fonction de la sévérité du retard, un second « *mode dégradé* » peut également être activé.

(2) Si la tâche ne peut se faire selon un « *mode dégradé* » sous peine d'impacter négativement sa qualité, c'est celui d'une ou plusieurs tâche(s) ultérieure(s) qui sera activé.

(3) Si la tâche n'est pas indispensable pour poursuivre le projet, elle peut être mise en pause en attendant des jours plus favorables.

- Enfin, une équipe projet doit être constituée, sous un « *mode initial* » et un « *mode dégradé* » correspondant à la plus petite équipe possible, tout en s'assurant de représenter au mieux l'ensemble des parties intéressées.

Le **lot 3** vise à définir les besoins pour aboutir à un cahier des charges détaillé :

- Pour cela, il faut au préalable déterminer les CTQ (*Critical to Quality*, ou essentiel pour la qualité), en combinant la VOC (*Voice of Customer*, ou voix du client) et la VOB (*Voice of Business*, ou voix du business). Dans un contexte normal, « *mode initial* », cela consiste à interroger les clients finaux et les commanditaires du projet afin de recueillir leur voix. Dans un contexte de crise avec réduction de la disponibilité des ressources le premier « *mode dégradé* » consistera à réaliser un sondage auprès d'un panel représentatif de chacune des parties. En cas d'aggravement de la situation, comme par exemple l'indisponibilité des ressources, un second « *mode dégradé* » pourra être mis en place en étudiant par exemple toutes sources d'information disponibles concernant le sujet (emails, rapports précédents, etc.).
- Une fois les CTQ déterminées, une analyse fonctionnelle doit être réalisée. Idéalement cette étape doit se faire en présence de l'équipe projet. En « *mode dégradé* », elle

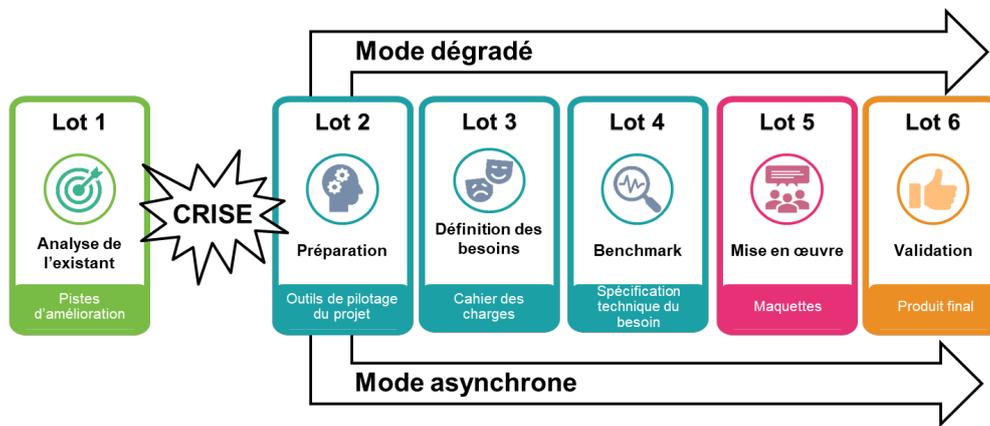


Figure 3. Proposition d'une méthodologie de gestion d'un projet d'amélioration en contexte de crise adaptée du PDCA (P en bleu, D en rose, C en orange, A en vert)

pourra être réalisée avec une équipe projet réduite (un représentant du client et un représentant des commanditaires) voire en autonomie par le chef de projet. Les degrés de flexibilité du livrable de l'analyse fonctionnelle, le cahier des charges, serviront, si nécessaire, de « *mode dégradé* » en cas de retard dans le projet.

Le **lot 4** consiste à réaliser un benchmark, interne et externe, pour identifier des solutions potentielles répondant aux besoins déterminés dans le cahier des charges. Si cette étape nécessiterait des échanges avec des collaborateurs, des consultants, des entreprises extérieures, elle peut également se réaliser, en « *mode dégradé* », de manière autonome grâce aux documents stockés sur les différents réseaux collaboratifs d'entreprises, ou aux ressources disponibles sur internet (littérature scientifique, normes, etc.). Afin d'étendre le champ des recherches sur le *cloud*, la créativité dans la recherche de nouveaux mots clés peut être suscitée par *mind-mapping*. Le livrable attendu à ce stade est donc une spécification technique du besoin.

3.3 Phase DO

La phase *Do*, constituée du **lot 5**, vise à la mise en œuvre de la solution, selon la spécification technique du besoin, à travers différents POC (*Proof Of Concept*) ou maquettes. C'est principalement cette phase qui nécessite l'intervention de l'équipe projet et de la collaboration entre ses membres. Idéalement (« *mode initial* ») l'animation de ce groupe de travail doit se faire en présentiel afin d'utiliser le type de communication le plus simple, la communication verbale et synchrone. Les « *modes dégradés* » présentés dans la Figure 2 font intervenir une communication asynchrone. Pour mettre en place et animer le groupe de travail en « *mode asynchrone* », quatre facteurs nous paraissent, selon notre étude présentée dans l'introduction de ce chapitre, indispensables :

- (1) Etablir un cadre, facilitant la convergence des participants ;
- (2) Susciter, à distance, leur créativité ;
- (3) Maintenir les relations humaines dans ce monde digital ;
- (4) Maintenir la motivation des participants.

3.4 Phase CHECK

La phase *Check*, constituée du **lot 6**, a pour objectif de vérifier l'adéquation entre les attentes des parties intéressées (clients finaux et commanditaires) et le rendu. Cela consiste donc à valider le POC le plus pertinent pour obtenir le résultat

escompté. Le « *mode dégradé* » consiste à réaliser cette étape en parallèle de la précédente, selon un mode agile, permettant de gagner en efficacité.

4 CAS D'APPLICATION

Nous avons appliqué la méthodologie présentée dans la Figure 3, à un projet d'amélioration du pilotage des processus chez *Safran Landing Systems*, spécialisée dans le développement, la fabrication et la maintenance des trains d'atterrissage et des équipements qui leur sont associés (roues, freins, etc.). Cette entreprise, certifiée EN9100 est composée de 14 sites répartis sur trois continents. Le projet a débuté en même temps que l'annonce de la pandémie de Covid-19 en début d'année 2020. Cette crise a fortement impacté le secteur de l'aéronautique en général (-91% de volume de transport aérien mondial entre Mai 2019 et Mai 2020 selon IATA.org) et donc les activités de l'entreprise, qui a choisi, dès le mois de Mars, de prendre des mesures d'activité partielle, transformées par la suite en activité partielle de longue durée, associées à la généralisation du télétravail (100% pour la grande majorité des collaborateurs lors du premier confinement). Dans ce contexte, le projet qui nous avait été confié est le suivant : l'amélioration de la grille de maturité des processus.

4.1 Phase ACT – Lot 1 : Analyse de l'existant

La grille de maturité est utilisée par les pilotes de processus lors de la revue annuelle. Elle permet à chaque pilote de faire un état des lieux du processus dont il a la charge et de définir les actions d'amélioration à mener. Avec une évolution récente du système de management de la qualité de la société, son contenu, composé de 10 critères déclinés en 5 niveaux d'exigence était devenu obsolète. Afin d'avoir une vision exhaustive de la situation, nous avons réalisé une analyse de l'existant, présentée dans la Figure 4.

En ce qui concerne la validation du besoin, nous avons interrogé un panel représentatif des parties intéressées, à savoir des pilotes de processus (les clients) et des responsables du Système de Management de la Qualité (SMQ, les commanditaires). Le besoin d'amélioration a bien été confirmé, les clients trouvant le contenu obsolète et ne permettant plus la prise de décision quant aux opportunités d'amélioration ; et les commanditaires souhaitant encourager le partage de bonnes pratiques entre les différents sites de la société, pour assurer une montée en maturité globale et équilibrée.

QUOI	De quoi s'agit t'il? Grille d'évaluation de la maturité des processus constituée de 10 critères déclinés en 5 niveaux d'exigence.
QUI	Qui remplit cette grille? Les 18 pilotes de processus, localisés au Siège de l'entreprise, et leurs représentants, localisés sur chacun des 14 sites de l'entreprise.
OU	Quel est le support utilisé? Excel.
QUAND	Quand la grille est elle remplie? Annuellement, en amont de la revue de processus.
COMMENT	Comment la grille est-elle remplie? Les pilotes des processus remplissent chacun 9 des 10 critères de leur grille. Les représentants sur les sites remplissent chacun une grille intermédiaire alimentant le dernier critère.
COMBIEN	Combien de fichiers Excel sont remplis? 270 Quelle est la durée totale du traitement? 5h50
POURQUOI	Pourquoi la grille est elle importante? Elle permet d'identifier des actions d'amélioration pour chacun des processus, une exigence de la norme ISO-9001.
	PROBLEME: Le contenu est obsolète. La voix des sites n'est pas suffisamment représentée. La durée de traitement des données est trop chronophage.

Figure 4. Analyse de l'existant réalisée sous forme de « QQQCCP »

4.2 Phase PLAN

4.2.1 Lot 2 : Préparation

Un plan d'action composé des « modes initiaux » et des « modes dégradés » associé à un planning de type Gantt ont été réalisés. La durée totale du projet a été fixée à 6 mois (de Mars à Septembre) afin de pouvoir déployer la nouvelle grille de maturité (Octobre et Novembre) et analyser les retours (Décembre) avant les revues de processus ayant lieu en début d'année. Les 2 indicateurs i1 et i2 ont été incorporés dans le Gantt. Enfin, une équipe projet a été constituée, en « mode dégradé » composée d'un représentant des commanditaires, de 4 représentants des clients (1 pilote de processus localisé au Siège de l'entreprise et 3 représentants localisés sur différents sites) et d'un animateur / chef de projet.

4.2.2 Lot 3 : Définition des besoins

Le recueil de la VOC et de la VOB n'a malheureusement pas pu être réalisé de manière directe en raison des mesures d'activité partielle. Nous avons donc enclenché le « mode dégradé » en étudiant de manière plus détaillée l'analyse de l'existant et la validation du besoin du lot 1, mais également les résultats des grilles de maturité, assorties de commentaires personnalisés, émises par les pilotes de processus et leurs représentants sur les sites au cours des années précédentes. Cela nous a permis d'identifier quatre CTQ :

- CTQ1 : Redonner un rôle de pilotage à la grille de maturité
- CTQ2 : Refléter la maturité de l'ensemble des processus en donnant plus de voix aux sites
- CTQ3 : Favoriser un esprit d'entraide entre les sites
- CTQ4 : Proposer un outil réduisant le temps de traitement des réponses

Quant à l'analyse fonctionnelle, si l'objectif initial (« mode initial ») était de la réaliser avec une équipe pluridisciplinaire, nous avons dû enclenché le « mode dégradé », l'équipe projet ayant à ce moment-là un taux d'activité de 20%. Nous l'avons donc menée en groupe réduit avec uniquement le chef de projet et un commanditaire. Ce dernier étant par ses fonctions le représentant site d'un processus, les clients étaient ainsi représentés.

Le cahier des charges, livrable de ce lot 3, était composé de 3 situations de vie (conception, remplissage, analyse), de 7 fonctions et de 18 critères d'appréciation quantifiés par un niveau, un degré de flexibilité et une limite. Nous avons

accordé beaucoup d'importance aux flexibilités et aux limites de chacun des critères afin de s'assurer d'avoir suffisamment de possibilités de mettre en place des « modes dégradés » le cas échéant.

4.2.3 Lot 4 : Benchmark

Afin de respecter l'esprit de la roue de Deming qui préconise une amélioration pas à pas en repartant de ce qui a d'ores et déjà été construit [Chardonnet et Thibaudon, 2003], nous avons commencé par réaliser un benchmark en interne, au sein du Groupe Safran. Nos collègues étant en activité partielle, nous avons profité de l'existence d'un référentiel documentaire commun à toutes les sociétés du Groupe Safran, pour identifier et analyser les grilles de maturité des processus adoptées par deux autres sociétés. Pour aller plus loin, nous avons également analysé les grilles d'autoévaluation proposées par Safran comme celle concernant le système de management ou celle sur la performance des chantiers d'excellence opérationnelle. L'analyse a consisté à remplir une fiche de synthèse pour chacun de ces documents comprenant les catégories suivantes : outil utilisé, structure de l'évaluation, critères évalués, logique de constitution des niveaux, et démarche d'exploitation des données.

Le benchmark externe, qui devait se faire notamment par la visite d'exposants au sein du Salon de la Digitalisation qui a lieu chaque année au mois de Mars porte de Versailles à Paris, (reporté à l'automne cette année), s'est également réalisé en « mode dégradé ». Nous avons mené à la place des recherches sur internet en les structurant selon la méthode suivante :

- (1) Recherche par brainstorming des mots clés permettant d'identifier la littérature pertinente (ex : « pilotage des processus », « maturité des processus »).
- (2) Utilisation des outils à disposition pour mener ces recherches (Bibliothèque numérique de l'ENSAM, Google Scholar, ResearchGate, ScienceDirect, AFNOR, CMMI, etc.).
- (3) Analyse des documents grâce à la fiche de synthèse (voir plus haut).
- (4) Étendre le champ des recherches grâce à de nouveaux mots clés identifiés par *mind-mapping*.

Ces benchmarks interne et externe nous ont permis de proposer 27 solutions potentielles pour répondre aux 18 critères d'appréciation définis au sein du cahier des charges et d'ainsi de définir la spécification technique du besoin.

4.3 Phase DO– Lot 5 : Mise en œuvre

Cette phase de réalisation s'est déroulée dans un contexte particulier, les membres de l'équipe projet étant localisés sur différents sites à travers le monde, ils n'avaient pas de compatibilité temporelle du fait des mesures d'activité partielle et des différents fuseaux horaires. Leur contribution à ce projet a donc dû se faire selon un « mode asynchrone ». Nous souhaitions les faire intervenir pour les deux premières étapes de cette phase : la définition des critères à partir desquels la maturité des processus sera évaluée et la définition des niveaux d'exigences à associer à chaque critère. Afin de mener à bien ces séances de brainstorming de manière asynchrone, nous avons choisi d'utiliser un logiciel, accessible par tous les acteurs, et nous permettant de respecter les facteurs clés de succès de la communication identifiés dans le chapitre 3.3 : le logiciel Klaxoon®. Ce dernier nous a en effet permis de :

- **Poser un cadre facilitant la convergence des participants.** Nous avons enregistré, sous la forme d'une courte vidéo, une présentation du projet. Pour favoriser la coopération, nous avons rappelé que ce projet

d'amélioration avait pour origine la réponse à leurs besoins et sollicitations exprimés ces 2 dernières années. Nous avons insisté sur notre objectif commun, la réalisation d'un outil permettant d'améliorer le pilotage de leur processus sur l'ensemble des sites de la société. Après avoir stimulé leur envie de s'impliquer, à travers un projet d'envergure, nous avons présenté le mode de travail original que nous allions adopter : un brainstorming asynchrone où chacun pourrait participer indépendamment en fonction de ses disponibilités. Afin de structurer ce travail, des objectifs hebdomadaires ont été fixés et un outil digital, support autoporteur et intuitif, a été créé sur Klaxoon® (Figure 5).

- **Susciter la créativité à distance.** Si les brainstormings stimulent la créativité par l'interaction des idées de chacun, dans notre cas, les acteurs ne pouvaient pas interagir directement et se retrouvaient seuls face à leur ordinateur. Le support proposé devait donc suffire à les stimuler. Mais comment créer cette dynamique créative ? Faut-il générer les premières idées (les premiers post-its), pour pallier le syndrome de la page blanche ? Face au risque d'orienter les idées suivantes et de biaiser les résultats, nous n'avons pas retenu cette option. Nous avons préféré susciter la créativité via un scénario. Nous avons choisi d'adopter l'approche métaphorique suivante : « Imaginez : vous êtes aux commandes d'un avion. Quels sont les paramètres que vous allez surveiller ? De quoi avez-vous besoin pour assurer votre vol ? » Ainsi, Le pilote de l'avion (pilote de processus) est aux commandes de son appareil (son processus) qu'il doit connaître (type, caractéristiques, performance) et a besoin de son équipage (des acteurs du processus), de ses indicateurs de vols (de ses indicateurs), de ses instructions de vols (sa documentation), etc.

- **Maintenir les relations humaines dans un monde digital.** Permettre aux acteurs de se situer par rapport au groupe et susciter leur créativité est un début, toutefois la contrainte de la distance et l'absence de communication informelle impactent la capacité à se comprendre. Pour y pallier, il est important de pouvoir, quand les conditions le permettent, de réaliser des appels vidéo, pour enfin mettre un visage sur un nom et créer une complicité. Quant au support Klaxoon®, il a été personnalisé autant que possible, via des messages informels, des formes et dessins permettant de créer une atmosphère plus humaine et visuelle. A cet effet, nous avons utilisé une option de Klaxoon® permettant aux participants de « liker » le contenu d'un post-it d'un autre membre de l'équipe (via une icône « cœur », matérialisant un « pouce levé » en communication non-verbale), mais également de rebondir et de questionner les notions y figurant (via une icône « ? » assortie de commentaires matérialisant un « haussement de sourcil » en communication non verbale) favorisant ainsi l'apprentissage collectif.
- **Maintenir la motivation.** Nous nous sommes attachés à montrer le caractère vivant et dynamique du projet, de trois façons différentes et complémentaires :
 - (1) La première a consisté à mettre en place un avancement progressif à plusieurs étapes. L'étape 1 a permis de définir les critères que nous avons ensuite synthétisés et fait valider par l'équipe projet. L'étape 2 était consacrée au nombre de niveaux qui, selon l'équipe projet, serait nécessaire et suffisant pour évaluer la maturité. Puis la définition des exigences en différents niveaux a été séquencées en 4 étapes supplémentaires.
 - (2) La seconde astuce pour rendre le projet vivant a consisté à rendre visible son avancement via l'intégration d'indicateurs illustrés par des sabliers, et concernant le pourcentage d'exigences définies et l'avancement temporel.



Figure 5. Création sur Klaxoon® d'un outil digital, support au brainstorming asynchrone. Sur ce template, 3 zones sont représentées : « A propos » pour présenter le projet, « Que dois-je faire ? » expliquant les règles du brainstorming asynchrone et « c'est ici que ça se passe ! » matérialisant la zone de créativité (illustrée en Figure 6).

(3) Enfin, afin de dynamiser le brainstorming, nous avons parfois dû agir en tant que participant pour soutenir le rythme du projet. Afin de rester objectifs dans cette démarche, nous avons uniquement ajouté des éléments identifiés lors de l'analyse du besoin et du benchmark. Ainsi, ces nouvelles idées ont permis non seulement de motiver les participants en voyant le projet progresser mais également de susciter leur créativité.

En termes de résultats, les critères ont été définis par les idées des participants ainsi que par des éléments issus du benchmark (étude des non-conformités émises au cours des audits internes et externes afin de donner une vision concrète des défaillances du pilotage ; analyse des exigences de l'EN9100 [AFNOR, 2018]. Au total 48 idées ont été émises et synthétisées en 9 critères (Figure 6, à but illustratif uniquement : pour des raisons de confidentialité les idées ont été floutées).

Afin de structurer la définition des exigences, nous nous sommes basés sur le modèle CMMI (*Capability Maturity Model Integration*) analysé lors de la phase de benchmark [Trehin, 2011]. Les niveaux ont ensuite été définis par les idées des participants ainsi que par des éléments provenant des retours effectués via les précédentes grilles de maturité, des données issues du benchmark interne ainsi que celles issues des normes EN 9100 [AFNOR, 2018] et ISO 9004 [AFNOR, 2018]. Au total 289 idées ont permis de proposer 5 exigences incrémentales par critère, soit un total de 45 exigences.

Adapter la méthodologie, ce n'est pas toujours la dégrader : il est également possible de l'améliorer ! C'est précisément ce que nous avons fait en décidant de digitaliser la grille de maturité pour la moderniser. Après avoir réalisé un benchmark, nous avons proposé d'utiliser la fonctionnalité « sondage » de l'outil Klaxoon®. En plus de proposer une meilleure expérience utilisateur, la digitalisation permet de simplifier le traitement des données (qui se retrouvent alors toutes dans une même base) et donc de limiter les tâches sans valeur ajoutée.

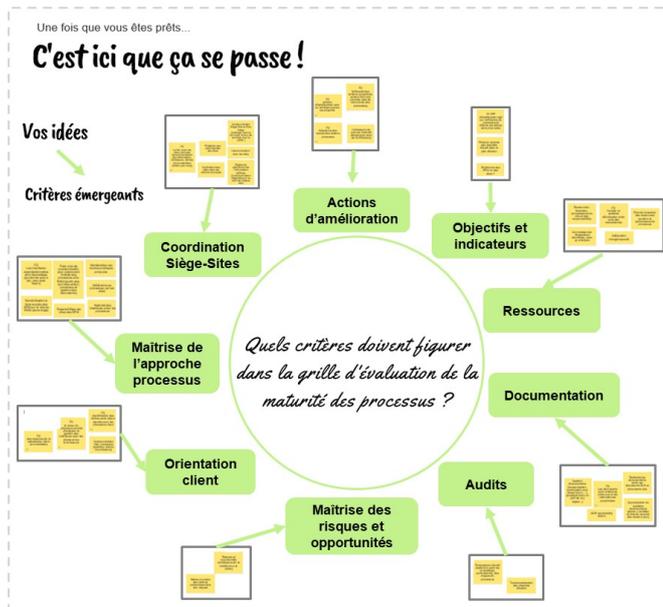


Figure 6. Synthèse des idées émises par les participants en 9 critères d'évaluation pour la grille de maturité. Ce travail a été réalisé dans la zone « C'est ici que ça se passe » de notre outil digital (voir Figure 5). Pour ces raisons de confidentialités, les idées émises (post-its jaune) ont été volontairement rendues illisibles.

Ainsi, à la fin de cette phase *Do*, nous avons pu fournir une maquette digitale de la nouvelle grille de maturité (pour des raisons de confidentialité, celle-ci ne sera pas dévoilée). Un exemple d'un des critères, « documentation », est présenté Figure 7.

9 CRITÈRES	5 NIVEAUX DE MATURITÉ				
	BASIC	BRONZE	ARGENT	OR	PLATINE
Maitrise de l'approche processus					
Orientation client					
Objectifs et indicateurs					
Actions d'amélioration					
Maitrise des risques et opportunités					
Audit					
Ressources					
Documentation					
Coordination siège-sites					

	BASIC	BRONZE	ARGENT	OR	PLATINE
Documentation	La documentation intègre et répond aux exigences internes et externes. Les équipes savent accéder à la dernière version des documents qui les concernent.	La mise à jour de la documentation est planifiée et respecte la procédure de gestion documentaire en vigueur. J'informe les équipes impactées par la création, mise à jour ou suppression de document.	Je participe à la rédaction des documents Safran Landing Systems me concernant. Sur site, je m'assure que les équipes participent à la rédaction des documents site qui les concernent.	Lors de chaque création ou évolution de document site, je veille à ne pas créer de redondance avec les documents Safran Landing Systems et Safran.	Toute la documentation est liée logiquement selon la définition qui lui correspond dans la procédure en vigueur, sans redondance et à jour.



Figure 7. Maquette de la grille de maturité, composée de 9 critères et 5 niveaux (illustrée par le critère « documentation ») et sa digitalisation

4.4 Phase CHECK– Lot 6 : Validation

La phase *Check*, constituée du lot 6, a été réalisée en 4 étapes :

- (1) La validation du contenu de la grille de maturité a été réalisée tout au long de la phase *Do*, en mode agile. Si cela était à l'origine un « mode dégradé » il nous a permis au final de gagner du temps en réajustant au plus tôt les livrables intermédiaires. La validation des critères puis de chacun des niveaux les composant a été réalisée grâce à la fonction sondage de Klaxoon®.
- (2) La vérification de l'adéquation entre les attentes des parties intéressées et le rendu a été réalisée à la fin de la phase *Do*, selon le mode initial. Ainsi, 15 des 18 critères d'appréciation définis dans le cahier des charges étaient bien satisfaits par la maquette digitale (83%). Toutefois, en analysant les indicateurs, cette action a été « mise sur pause », le temps de réaliser l'étape 3.
- (3) Nous avons testé le déploiement de la nouvelle grille de maturité sur un périmètre restreint afin de tester l'acceptabilité et d'évaluer le temps de traitement des réponses. Cette analyse nous a permis de valider les 3

derniers critères d'appréciation, validant ainsi à 100% le respect du cahier des charges.

(4) Enfin, afin d'évaluer la satisfaction de nos clients, nous avons proposé un retour sur les CTQ :

- ✓ CTQ1 : Satisfait par la définition de nouvelles exigences progressives et plus ambitieuses. Une satisfaction de 85% a été obtenue lors d'un sondage préliminaire.
- ✓ CTQ2 : Satisfait par la création d'une grille de maturité unique à destination des sites
- ✓ CTQ3 : Satisfait par le contenu de la grille de maturité, chacun des critères intégrant des exigences de partage de bonnes pratiques.
- ✓ CTQ4 : Satisfait par le nouveau format et la digitalisation de la grille de maturité. La durée de traitement est passée de 5h50 à 55 minutes, associée à une réduction de 92% des tâches sans valeur ajoutée.

Enfin, un dernier fait important, le projet s'est terminé dans les temps avec un ratio i1/i2 de 1 à l'issue des 6 mois.

5 CONCLUSION

Afin de répondre à notre problématique, « Comment mener un projet d'amélioration dans un contexte de crise ? », nous avons proposé une démarche permettant de mener un projet d'amélioration en situation de crise. Cette méthodologie repose sur la démarche de management de projet utilisée de manière classique en amélioration continue, le PDCA, à laquelle ont été incorporées deux adaptations nécessaires pour faire face à la crise : le « *mode dégradé* » et le « *mode asynchrone* ». Nous sommes ensuite passés de la théorie à la pratique en appliquant cette démarche à l'amélioration de la grille de maturité des processus chez *Safran Landing Systems*. Nous avons démontré la fiabilité de cette démarche par la réussite du projet dans les temps impartis et ce malgré le contexte de crise. Afin d'évaluer de manière objective la méthodologie employée lors de ce projet, nous avons réalisé un sondage auprès des membres de l'équipe projet sur plusieurs critères : le « *mode asynchrone* », la communication, le support de brainstorming utilisé, le contenu de la grille de maturité, la satisfaction générale (Figure 8). La satisfaction moyenne ainsi obtenue s'élève à 89% confirmant ainsi le succès de notre méthodologie.

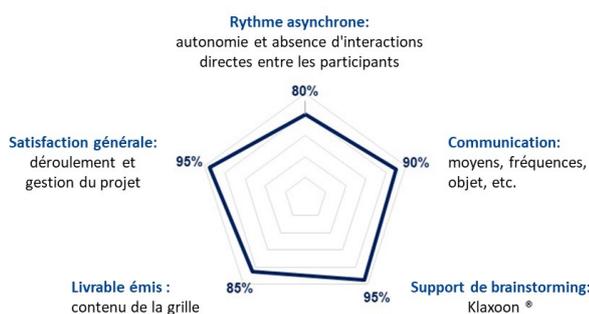


Figure 8. Evaluation de la satisfaction des membres de l'équipe projet

La situation de distanciation physique que nous avons vécue et que nous vivons encore aujourd'hui avec un deuxième confinement, encourageant le télétravail comme nouvelle norme, va certainement être amenée à perdurer, que ce soit chez Safran ou dans le monde de l'entreprise en général. Ainsi, afin de pérenniser les bonnes pratiques issues de ce projet, un bilan s'impose :

- Le « *mode dégradé* » s'est avéré être un vrai système de management des risques projet et un vrai facteur de succès.

L'objectif était d'imaginer différents scénarii possibles à une même tâche afin de ne pas pénaliser l'avancement de projet en cas d'imprévu, ni la qualité du livrable. Cela s'est notamment matérialisé par la réduction des objectifs, la mise en place d'alternatives, le cumul de plusieurs rôles pour le chef de projet, la parallélisation des tâches ou encore leur mise en pause. Alors que certaines phases ont été écourtées, d'autres s'en sont trouvées avantagées, comme le benchmark, réalisé pendant le premier confinement. Cela s'est révélé être une force durant ce projet puisque les résultats de ce dernier ont été utilisés à de nombreuses reprises (lot 3, lot 4). Il est important de noter que la qualité du livrable final n'a pas été dégradée mais est au-delà des attendus grâce à sa digitalisation. Cette tâche non prévue initialement a pu être réalisée grâce à une excellente gestion du temps.

- Le « *mode asynchrone* » a été un réel défi que nous relevé grâce à la mise en place d'un support de travail autoporteur hébergé sur le logiciel Klaxoon® et une gestion de projet cadrée mais souple, s'adaptant pas à pas aux besoins et à l'avancement du projet. Cela a permis de faire participer les acteurs à leur rythme, dans les meilleures conditions possibles. Ils ont pu passer outre le mode informel grâce aux fonctions « *like* » et aux « ? » des post-its virtuels. Les points forts identifiés par les participants souligne cette réussite. Ils ont en effet particulièrement apprécié la flexibilité qui leur était offerte, le côté ludique de Klaxoon® et la parfaite adéquation de la méthodologie face à la situation de crise que nous traversions. La clarté des consignes et les encouragements tout au long du projet ont également été mis en avant, démontrant ainsi que le « *mode asynchrone* » n'est pas un frein à la communication, ni au travail collaboratif.

6 PERSPECTIVES

Cependant, des pistes d'amélioration sont à envisager afin d'améliorer notre méthodologie. En effet, le « *mode asynchrone* » et l'incapacité de faire des réunions avec l'ensemble des participants est un handicap à la création du sentiment d'appartenance à un groupe, le seul point faible relevé par les participants. Il faudrait pour l'avenir envisager un moyen de personnaliser les actions menées sur le support. Dans un monde de plus en plus digital, pourquoi ne pas imaginer des avatars dans un monde virtuel qui présenteraient leurs idées ? Si cela a déjà fait ses preuves, notamment dans le domaine de conception innovante [Guegan et al., 2016], où la personnalisation des avatars et de leur environnement de travail ont un vrai impact sur la créativité des participants, il suffirait d'y rajouter un « *mode asynchrone* ». On peut ainsi envisager des avatars présentant leurs idées, selon un discours préenregistré dans un monde tel que celui proposé par Virbela®, à condition d'y mettre les moyens.

Pour conclure, si cette méthodologie est perfectible, elle peut parfaitement, en l'état et à moindre coût, être transposée à toute entreprise ou secteur d'activité soumis des circonstances et contraintes particulières comme le besoin de maintenir un haut niveau de qualité pour envisager une vision à long terme malgré des ressources réduites et dispersées. Cette méthodologie peut également être mise en place pour gérer des projets dans des entreprises où les ressources nécessaires à sa mise en œuvre sont réparties sur différents sites et fuseaux horaires, une situation de plus en plus courante pour faire face à la mondialisation.

7 REFERENCES

- AFNOR (2015). Systèmes de management de la qualité - Principes essentiels et vocabulaire. NF EN ISO 9000:2015.
- AFNOR (2015). Systèmes de management de la qualité- Exigences. NF EN ISO 9001:2015.
- AFNOR (2018). Management de la qualité – Qualité d’un organisme – Lignes directrices pour obtenir des performances durables. NF EN ISO 9004:2018.
- AFNOR (2018). Systèmes de management de la qualité – Exigences pour les organismes de l’aéronautique, l’espace et la défense. NF EN 9100:2018.
- Attussier, D. & Moutot J. M. (2010). Méthode de conduite du changement. Dunod.
- Chardonnet, A. & Thibaudon, D. (2003). Le guide du PDCA de Deming. Editions d’Organisation.
- Deming, W. E. (1982). Quality Productivity and Competitive Position. Conroy Edition.
- Eymery, A., N. Maranzana, A. Aoussat and J.P. Mathevet (2019). Proposition d’une démarche de digitalisation du système documentaire : application au manuel qualité Safran. 13^{ème} Congrès International de Génie Industriel (CIGI-QUALITA’19), Montréal, Canada.
- Guegan, J., S. Buisine, F. Mantelet, N. Maranzana and F. Segonds (2016). Avatar-mediated creativity: When embodying inventors makes engineers more creative. *Computers in Human Behavior*, Vol. 61 (August 2016), pp. 165-175
- Hisarciklilar, O., & Boujut, J.-F. (2006). Reducing the “information gap” between synchronous and asynchronous co-operative design phases. *Virtual Concept*, Cancun, Mexico.
- Hisarciklilar, O. (2008). Formes et structures des annotations sémantiques pour supporter la communication en conception collaborative asynchrone. Thèse de doctorat. Institut National Polytechnique de Grenoble.
- Knapp, M. L., Hall, J. A. and T. G. Horgan (2013). *Nonverbal Communication in Human Interaction*. Wadsworth Publishing.
- Kotter, J. & H. Rathgeber (2018) *Alerte sur la banquise !*, Pearson.
- Maranzana, N., N. Gartiser and E. Caillaud (2008). From concurrent engineering to collaborative learning of design. *International Journal of Design and Innovation Research*. Volume 4, Issue 1, pp 39-51.
- Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31(3), 248–252.
- Shankar, R. (2009). *Process improvement using Six Sigma: A DMAIC guide*. ASQ Quality Press.
- Trehin, Y. (2011). *Modèles CMMI v. 1.3 – Manuel d’utilisation à l’usage des professionnels des systèmes d’information*. Editions ENI.

Risk Analysis: Moroccan Drug Approval Process

OUMAIMA OULLADA¹, AHMED ADRI², SAID RIFAI³

¹ Ecole Nationale Supérieure d'Electricité et de Mécanique, Laboratoire de Mécanique Productique et Génie Industriel, Ecole Supérieure de Technologie

UNIVERSITE HASSAN II DE CASABLANCA, B.P.8012, OASIS, CASABLANCA, MAROC.

oumaimaoullada@gmail.com

^{2,3} Ecole Supérieure de Technologie

UNIVERSITE HASSAN II DE CASABLANCA, B.P.8012, OASIS, CASABLANCA, MAROC

ahmedadri@gmail.com

dptgmp@gmail.com

Résumé – Introduction: Pour passer d'une entité chimique ou biologique à un médicament. Le développement d'un médicament nécessite dix à quinze ans de recherche. Notamment les tests précliniques, les essais cliniques et le développement industriel, qui sont strictement encadrés par la loi. Au Maroc la direction de médicament et de la pharmacie (DMP) se charge de l'octroi des autorisations de mise sur le marché des médicaments à usage humain (AMM). L'objectif de cette étude est d'évaluer les risques au cours du processus d'octroi des AMM des médicaments à usage humain au sein de la DMP.

Méthodes: il s'agit d'une étude des risques a priori par la méthode de l'analyse des modes de défaillance, de leurs criticité (AMDEC), afin de cerner mieux les risques liés au processus d'octroi des AMM des médicaments à usage humain.

Résultats: En 2018, 11 risques ont été identifiés comme prioritaires et l'indice de criticité total de l'analyse s'élevait à 451. Après mise en place d'actions d'amélioration, une nouvelle AMDEC a été réalisée un an plus tard montrant une baisse de l'indice de criticité totale de 72 % (126) avec 6 risques restant prioritaires. Ainsi, grâce à l'outil AMDEC, des risques ont pu être identifiés à l'année « n », corrigés par des actions puis réévalués efficacement à l'année « n + 1 »

Discussion: l'AMDEC nous a permis d'identifier les risques possibles, de reconsidérer certaines procédures et de proposer des plans d'actions pour la prise en charge des risques les plus critiques.

Conclusion: cette analyse permet, par des évaluations périodiques, d'entrer dans une véritable démarche qualité, ce qui renforce la satisfaction des patients, les industriels ainsi que l'ensemble des acteurs de la direction de médicament et de la pharmacie.

Mots clés: DMP, AMM, AMDEC, Indice de Criticité.

Abstract – Introduction: To switch from a chemical or biological entity to a drug. The development of a drug requires ten to fifteen years of research. This includes preclinical tests, clinical trials and industrial development, which strictly regulated by law. In Morocco, the Directorate of Medicine and Pharmacy (DMP) is responsible for granting marketing authorizations for drugs for human use (MAA). The objective of this study is to assess the risks during the process of granting MAA for drugs for human use within the DMP.

Methods: it is a study of the risks a priori by the method of Failure Mode and Criticality Analysis (FMECA), in order to better identify the risks related to the process of granting MAA for drugs for human use.

Results: In 2018, 11 risks were identifying as priorities and the total criticality index of the analysis was 451. After implementing improvement actions, a new FMECA has carried out one year later showing a decrease in the total Reliability & Risk Priority Number (RPN) of 72% (126) with 6 risks remaining priorities. Thus, thanks to the FMECA tool, risks were identified in year "n", corrected by actions and then effectively reassessed in year "n + 1".

Discussion: the FMECA allowed us to identify possible risks, to review certain procedures and to propose action plans to deal with the most critical risks.

Conclusion: this analysis allows, through periodic evaluations, to enter into a real quality approach, which reinforces the satisfaction of patients, industrialists as well as all the actors of the drug and pharmacy management.

Key words: DMP, MAA, FMECA, RPN.

1 INTRODUCTION:

Before a medicine or drug can be sold or prescribe to citizens across the Moroccan country, a marketing authorization of the drugs must be obtained. The Moroccan regulation offers one options for the authorization of medicinal products: the

national procedure from the submission of the Common Technical Document (CTD) to the granting MAA of drugs for human use. To be marketed in the Moroccan country, with the objective to ensure their efficacy, safety and quality. Confirmatory clinical trials had usually performed to inform a

benefit–risk decision, the results of which will be the basis for a treatment recommendation. Marketing Authorization Application: Across all European markets, plus Australia, New Zealand, South Africa and Israel (exceptions among major markets include USA, Canada, China and Japan), the MAA is a common document used as the basis for a marketing application (an application for approval to market the product based on a full review of all quality, safety and efficacy data, including clinical study reports). In the USA, the New Drug Application (NDA) is the MAA equivalent. In Canada, the New Drug Submission (NDS) is the MAA equivalent (Regnstrom et al., 2010). Also, in Moroccan country the MAA the terms using. The process of registration or gaining marketing authorization in different countries called by different terms that mean essentially the same thing the product may be marketed and sold there (Sheets, 2017). However, in Morocco for obtaining the granting MAA of drugs for human use, there is a pharmaceutical regulatory authority who has called Directorate of drugs and Pharmacy is one of the Directorates of the Ministry of Health that works in the deployment of the national policy on drugs and health products at the technical and regulatory levels. It ensures that these various interlocutors assume their responsibilities, on the basis of the expertise it conducts and the controls it carries out in order to guarantee the quality, effectiveness and safety of use of the products marketed (DMP, n.d.). Hence the importance of the application of a quality assurance system in the line of the strategy of the DMP that guarantees the safety of patients first and foremost, to meet the needs of clients including industrial pharmaceutical establishments and the regulatory sector (Maroc, n.d.). In this regard, several methods are available, and among the most effective: Failure Mode, Effects and Criticality Analysis (FMECA), which is an a priori risk analysis method. Properly used, this approach would allow an optimal management of the MAA circuit and an improvement in the quality of MAA delivery services. Among the existing risk management methods and the criteria for choosing between these methods, taking into account our problem of providing companies with a decision support tool and a global vision that helps to guarantee their sustainability. We have proceeded to the use of Failure Mode Analysis of their effect and criticality (FMECA). This method is also proposed by the Institute for Safe Medication Practices (ISMP) in Canada (Darcissac et al., 2020); (Poulin & Poirier, 2014). Our objective of this study is to assess the risks during the process of granting MAA for drugs for human use within the DMP.

2 MATERIALS AND METHODS

2.1 Definition of FMECA

Failure mode effect and critically analysis (FMECA) is a key analytical tool that aims to understand the root causes of failures associated with the operation of critical components in crucial systems. Its history is back to 1950's (Catelani et al., 2015). NASA has the credit of onset of FMEA for failure prevention, prediction and reliability requirements in 1963 (Bhangu & Grover, 2019). Failure Mode and Effects Analysis (FMEA) can be considered a systematic procedure that can be used in order to analyze a system with the aim to identify potential failure modes, their causes and effects on performance and, eventually, their effects on the safety of the personnel, on the environment as well as on the system (Standard 11266, 2014). Rank the recognized failure modes by their risk priority numbers (RPNs) and take remedial actions

for the higher risky failures. Although FMEA is a useful and popular tool for safety and reliability analysis (Huang et al., 2020). Over the past few years, several literature reviews in relation to FMEA have accomplished under different outlooks. For instance, the implementation of FMEA in service operation. Bouti and Kadi (ไทรทับทิม, 2554), carried out a literature review of FMEA/FMECA in 1994 mainly including its basic principles, main types, technique enhancement, automation, combination with other methods, and specific applications. By extending the literature review of Bouti and Kadi (ไทรทับทิม, 2554), Spreafico et al (Spreafico et al., 2017), surveyed the FMEA/FMECA papers and patents from 1978 to 2016 with the aim to identify the evolution of the method and highlight the technical contributions from industry. This is a study conducted within the Drug Directorate and Pharmacy unit over a period of one year. To analyze the risks, we opted for the FMECA method, a preventive, qualitative and quantitative process analysis technique.

2.2 Progress

As related in the theoretical part of our work, the FMECA-Process applies to manufacturing processes; it used to analyze and evaluate the criticality of all potential failures of a product caused by its process. This analysis requires the application of the Ishikawa Diagram, a brainstorming method used to search for and represent the different causes of a problem, an effect in the case of FMECA. Based on the identified effect, the brainstorming conducted around five categories of causes (Manpower, Machines, Materials, Methods, and Management). Also called the 5 M rule. The failure mode defined by a faulty or malfunctioning system. The criticality of the failure quantified by triple rating. Our work consisted in carrying out a quality risk analysis on a very important process at the level of the drug and pharmacy management, with a view to develop risk management within the DMP. The teams: quality unit, the head of department and us, had then contacted to make us aware of the importance of the approach and to explain its different stages. The work was planned on the following four processes like showing the figure: 1, it described the process of granting MAA in the form of a flowchart.

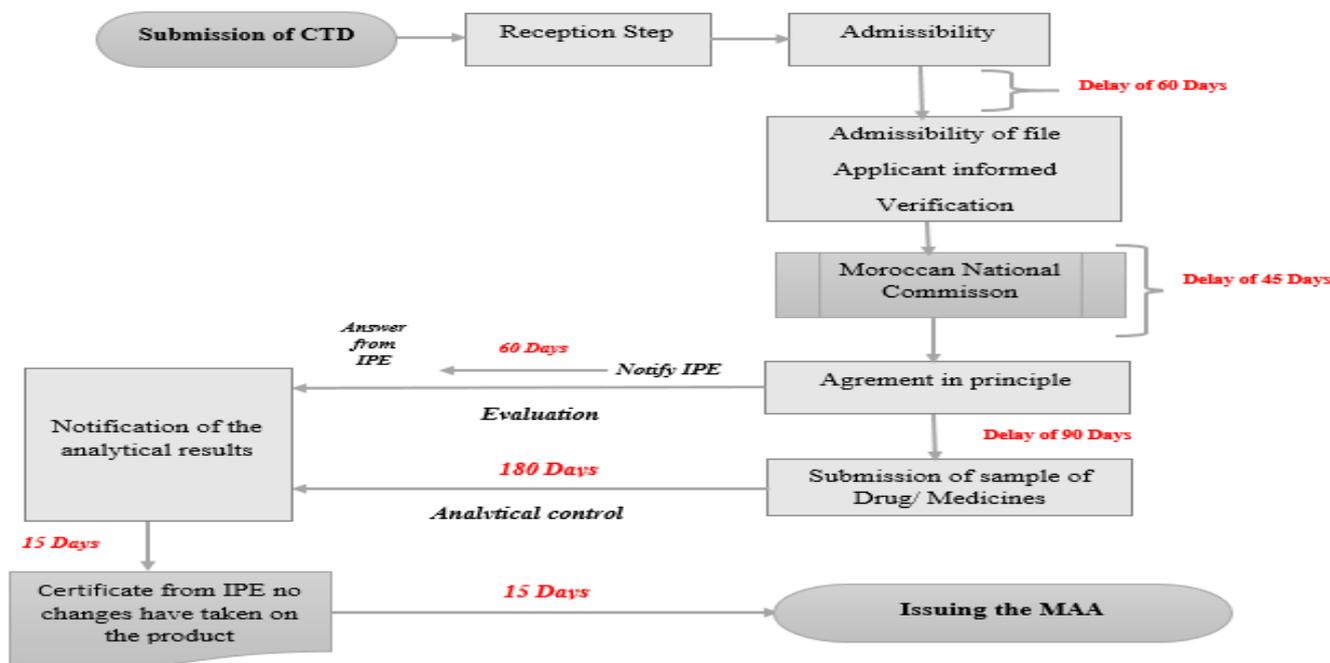
2.3 Process :

In order to obtain the granting of the MAA by the Ministry of Health, there is a procedure, the national procedure, where the responsible pharmacist must submit the file in CTD format. The file deemed admissible within 60 days after verification of the authenticity of the document. When the file is admissible, the industrial pharmaceutical establishment informed in writing. The DMP has a period of 45 days from the date of admissibility to rule on the efficacy and safety of medicines after receiving the opinion of the National Marketing Authorization Commission. By indicating whether the drug is a generic drug, the period shortened by 15 days. Once the efficacy and safety have defined, the applicant shall submit the sample of the medicinal product concerned, the documents, the reagent and the means necessary for the control. If the presentation has not made within 90 days from the date of the agreement in principle, the agreement in principle becomes invalid. The article of the decree mentioned when the results of the analytical control and the evaluation of the file are conclusive, the applicant is invited to provide within 15 days a certification attesting that any modification has taken place in the elements of the products. The Ministry of Health issues the MAA within 15 days from receipt of the certification. The period of the validity of the MAA is 5 years, renewable. The

delay to obtain the granting MAA for a new drug (Original Drugs) is 10 months, and for the generic drug is 9 months (DMP, n.d.), (Decree 2-14-841 n.d.). The DMP can withdraw the MAA depending on the gravity of the finding pursuant to the Article 31. In the event of suspension or withdrawal of the marketing authorization, the holder of the authorization shall have a maximum period of 6 months to rectify the deficiency noted.

2.4 Quantification of risks to prioritize them:

The failure mode defined by a faulty or malfunctioning system. The criticality of the failure quantified by triple rating (Meyrieux et al., 2012), (Table 1).



IPE: Industrial Pharmaceutical Establishment ; CTD: Common Technical Document

Figure 1: Process of Issuing MAA in the Directorate of Drug and Pharmacy

Table 1: The Quotation Grid:

Severity Quotation		
1	Minor	Without consequence MAA route complied with and delivered on time
2	Significant	Without probable consequences? incomplete receipt of files
3	Dramatic	Prolonged duration Difficulty in evaluating the administrative file
4	Serious	Risk of output data error
5	Fatal	Delay in issuing MAA
Occurrence Quotation		
1	Very rare or unlikely	The event has never happened or the event could happen only once
2	Rare	Event has already occurred or may occur: at most once a year
3	Occasional	The event has already occurred or could occur: more than once a month (1 to 3 times a month)
4	Frequent	The event has already happened or could happen (1 to 2 times a week)
5	Very fréquente	The event has already occurred or could occur: more than twice a week (10 or more times a month), at each file of advertising to healthcare professionals
Detectability Quotation		
1	Automatic	Observable defect: it cannot escape visual detection (at the time of the evaluation).
2	Quasi-certain	Detection is easy to implement
3	Difficult	Sampling is mandatory to detect the failure
4	Not detectable	There are no warning signs of the failure (External claims/customers)
Reliability & Risk Priority Quotation		Code
RPN > 33		High Reliability Priority risk reduction actions
16 < RPN < 32		Average Reliability Risk reduction actions to be taken after treatment of priority actions
RPN < 15		Low Reliability No Risk Reduction Action

- «S» grade: Severity of the failure, it is an evaluation on a multi-level scale;
- «O» grade: occurrence of the failure, this is an evaluation of the number of cases per unit of time;
- «D» grade: detection of the failure, represented by a coefficient on a multi-level scale, reflecting the attenuation of the seriousness of the consequences in case of detection weighted by the probability of detection. The product of the three scores obtains the criticality index or the Reliability & Risk Priority Number: RPN: $RPN = S * O * D$. In addition, Reliability is important, the more the mode of failure considered is of concern.

In practice, we applied the method according to the following steps:

- Definition of the objective of the perimeter to be study: visa approval and authorization service the drug and pharmacy department
- Constitution of a multidisciplinary working group: the service team it should be note that one of the members of the group is a qualitian, the head of the service and us.
- Determination and analysis of the steps in the process for monitoring the marketing authorization of a drug (originator and generic) for human use within the department.
- Highlighting of potential failures, causes and effects, according to the Ishikawa diagram (fish bone). The Ishikawa diagram helps to identify causes. Causes identifying and classified on the figure 2:

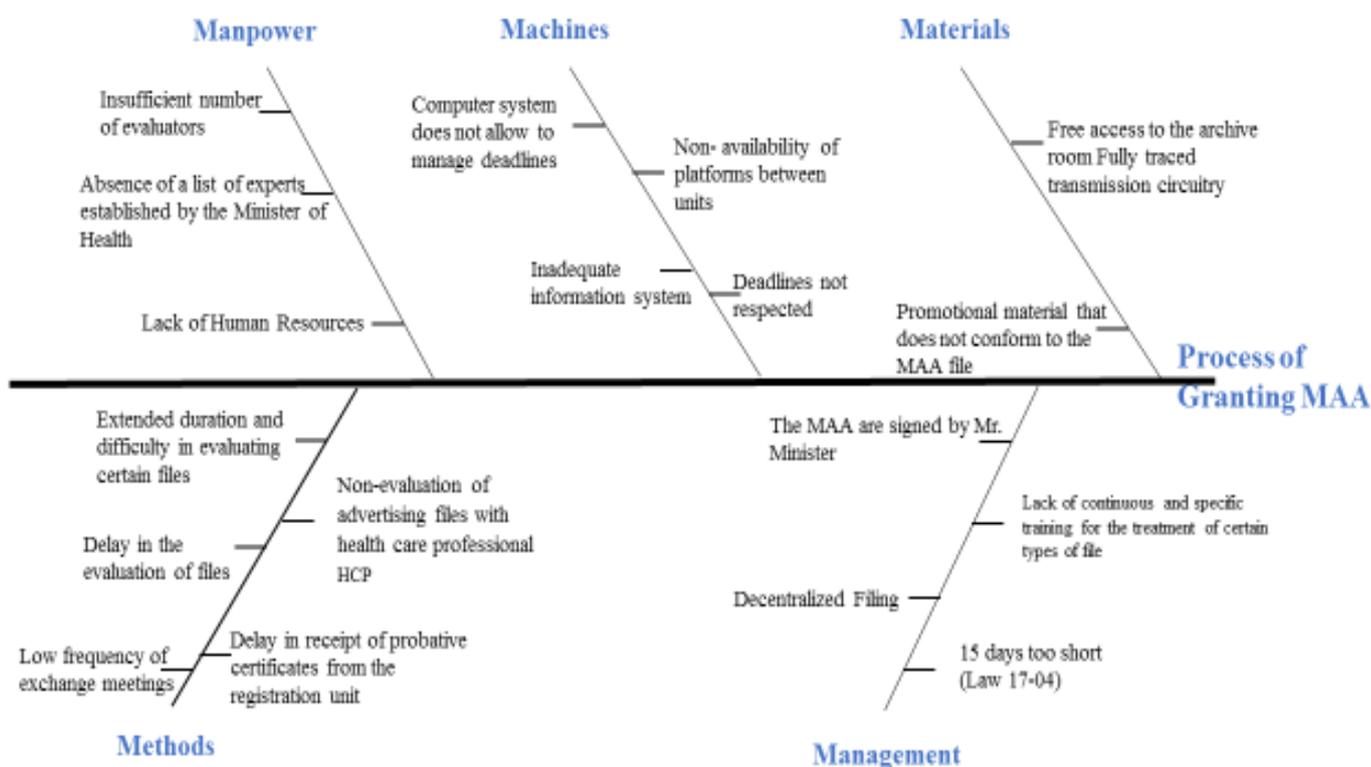


Figure 2. Ishikawa's Diagram process of MAA of Drugs for Human Use.

3 FMECA WORKSHEET:

3.1 FMECA 1 WORKSHEET

The following (Table 2) presents the FMECA 1 worksheet for the administrative process of the visa homologation and authorization service SVHA of the drug and pharmacy directorate. This table gives a detailed description of the components of the process, the failure modes are the ways in which the component could fail and the date of failure identification. The next step is to specify all kinds of possible and reasonable causes that cause a failure to occur. Finally, the effect of the failure defined as the result of the failure mode on the process function. It describes what happens when each failure mode occurs. For each risk obtained, the criticality (or

Reliability & Risk Priority Number= RPN) was calculated taking into account the Occurrence (O), severity (S) and detectability (D) of the risk (Table 1) according to:

$$RPN = O \times S \times D$$

3.2 FMECA 2 WORKSHEET

One year after the first risk analysis (FMECA 1), a second analysis was carrying out (FMECA 2). The working group met again for four brainstorming sessions, which allowed the identification, and quantification of risks, according to the same modalities as described for FMECA 1. The objectives of FMECA 2 were to evaluate the effectiveness of the actions implemented following FMECA 1, and to take into account the emergence of new risks (Table 3).

Table 2: FMECA 1 WORKSHEETS of failure mode of the Process of Granting MAA

Potential failure mode	Effects of Failure	Possible Causes of Failure	Date of Risk identification	S	O	D	RPN
Extended duration and difficulty in evaluating certain file	Delay in the evaluation of files Customer dissatisfaction	-Insufficient number of evaluators - Multiplicity of service assignments in relation to the number of evaluators	At the time of the evaluation	4	5	3	60
Inadmissibility of file	Delays in processing files	-Lack of Human Resources	Internal claim at the time of the evaluation	4	5	3	60
Absence of experts of the same specialty at National the MAA commission	Failure to keep the commission /failure to meet deadlines	-Absence of a list of experts established by the Minister of Health,	At the time of sending the output data/customer complaints	4	4	2	32
Deadlines not respected	-Non-compliance with regulatory deadlines. -Impact on the quality of the evaluation; delay of the evaluation.	-Computer system does not allow to manage deadlines -Inadequate information system -Lack of continuous and specific training for the treatment of certain types of file	At the time of sending the output data/customer complaints	5	2	3	30
Input data error between units	Transmission of erroneous data/image of the DMP	-Non- availability of platforms between units - Low frequency of exchange meetings	Customer Complaints	5	3	3	45
Case of file disaster/ Insecurity	- Delay in the review of files, -Image damage of the DMP	-Free access to the archive room Fully traced transmission circuitry	Internal and external Complaints	5	4	3	60
Non-evaluation of advertising files with health care professional HCP	Uncontrolled advertising, / Promotional material that does not conform to the MAA file	-15 days too short (Law 17-04) - Promotional material that does not conform to the MAA file	Permanent	2	5	3	30
Difficulty in evaluating some kinds of files	Delay in the evaluation of files	-Lack of continuous and specific training for the handling of certain types of files.	Internal Complaints	2	5	3	30
Risk of communication between the division SVH and the LNCM	Delay of the decision on the notification of IPE	-Low frequency of exchange meetings -Delay in receipt of probative certificates from the registration unit	Internal and external Complaints	4	2	3	24
Variation request filed only at the LNCM	Transmit the MAA to the client with erroneous data	Decentralized filing	Customer Complaints	5	2	2	20
Delay in the issuance of MAA, letters	Customer non-satisfaction	The MAA are signed by Mr. Minister	Customer Complaints	5	4	3	60

Table 3: FMECA 2 WORKSHEETS of actions implemented on the process of granting MAA

Potential failure mode	Actions	S	O	D	RPN
Extended duration and difficulty in evaluating certain file	- To provide for agreements with international drug agencies, particularly the Benchmark countries; -Improvement of the computer system and establishment of the handling guide	4	3	2	24
Inadmissibility of file	-Plan to Strengthen the department's workforce,	4	3	2	24
Absence of experts of the same specialty at National the MAA commission	- Validation of the list of experts by the Minister of Health and commitment of experts -Confirmation of experts -Validation of the internal regulations by the experts and the Minister of Health	4	2	2	16
Deadlines not respected	- Improve the information system -Strengthening the SVHA membership - Staff awareness	5	2	2	20
Input data error between units	-Non- availability of platforms between units - Low frequency of exchange meetings	2	2	1	2
Case of file disaster/ Insecurity	- Securing the archive room - Better traceability of the file reception circuit, - Provide files in e-CTD format - Immediate verification of any document received	2	2	1	2
Non-evaluation of advertising files with health care professional HCP	-Control the 48-hour time limit for the transmission of advertising files from the recording unit, -Amendment of Law 17-04 to review the deadline, -Assignment of staff to the SVHA for the creation of the advertising unit	2	5	3	30
Difficulty in evaluating some kinds of files	- Engage the evaluators in different training courses, even international training courses for the DMP's counterparts	2	2	1	2
Risk of communication between the division SVH and the LNCM	- Promote communication between divisions (Pharmacy and LNCM) and registration unit.	2	2	1	2
Variation request filed only at the LNCM	-The submission of the variation file to the LNCM must be accompanying by an acknowledgement of receipt at the pharmacy division level. - Centralize the deposit of the pharmacy division and the LNCM division,	2	2	1	2
Delay in the issuance of MAA, letters	-MAA are signing at the DMP level.	2	2	1	2

4 DISCUSSION :

In a first FMECA after having identified 11 failure modes, according to the MAA granting process, we were able to detect the root causes of each risk by indicating the severity, occurrence and detectability for a well-defined of RPN calculation, which helped us to detect criticality first. In addition, the calculation of RPN guided the work team to propose preventive measures and corrective actions. The objectives of these proposals are significantly to reduce the criticality rating or the RPN, and continuously to improve the quality management system in place. They are adaptable for the DMP. They enable a renewal of practices and a certain regularity in the process of granting MAA. Then a second FMECA 2 performed to see the effectiveness of the actions implemented while recalculating the RPN.

The figure: 3 shows the result, a comparison of the RPN total before and after FMECA of the process of granting MAA on the SVHA in the DMP. Evidently, the total number of RPNs has increased from 451 to 126. The reduction in the RPN clearly indicates a decrease in the possibility of failures.

For the first failure: Extended duration and difficulty in evaluating certain file. The calculated RPN goes back to 60, following the importance of this critical stage of the process of granting AMM, after the implementation of the corrective actions the risk is persisting with a RPN of 24. An unmanageable risk, while still waiting to improve the computer system at the heart of the DMP. As well as to increase the number of correspondences between the counterparts of the DMP to benefit the maximum of their own experiences. Even for the Inadmissibility of file, still an unmanageable risk,

pending the strengthening of the DMP's services. For the Absence of experts of the same specialty at National the MAA. Commission is the third step of the process of granting MAA, where the medical experts meet according to the specialty of the candidate drug where its therapeutic interest is pronounced. This step often indicates the absence of experts of the same specialty at the MAA commission, FMECA 1 demonstrated an RPN=32. The proposed actions (Table 3), which had partially implemented, have played a very important role in the development of the RPN and certainly in the long term. The failure Deadlines not respected, always persists with a new RPN= 20, a risk that has an impact on customer satisfaction, for which the DMP will have to think about a new recruitment strategy, in particular to strengthen its services and improve its system. For the Input data error between units and Case of file,

Disaster/ Insecurity the application of corrective actions have been effective, and then the RPN has been reducing. Also this mode of failure: the Non-evaluation of advertising files with health care professionals this is permanent due to the short period of time enacted by the Law 17-04 (Decree2-14-841 n.d.). The value of the RPN with the two FMECA studies has stagnated at 30 in anticipation of validating the amendment of the Law 17-04 to review the deadline also, allocation of staff to the SVHA for the creation of the advertising unit. The other risks mentioned on the table their RPN index perfectly decreased to 2, due to the actions carried out on the spot. Risks are perfectly controlled and monitored so that they do not fall back on them.

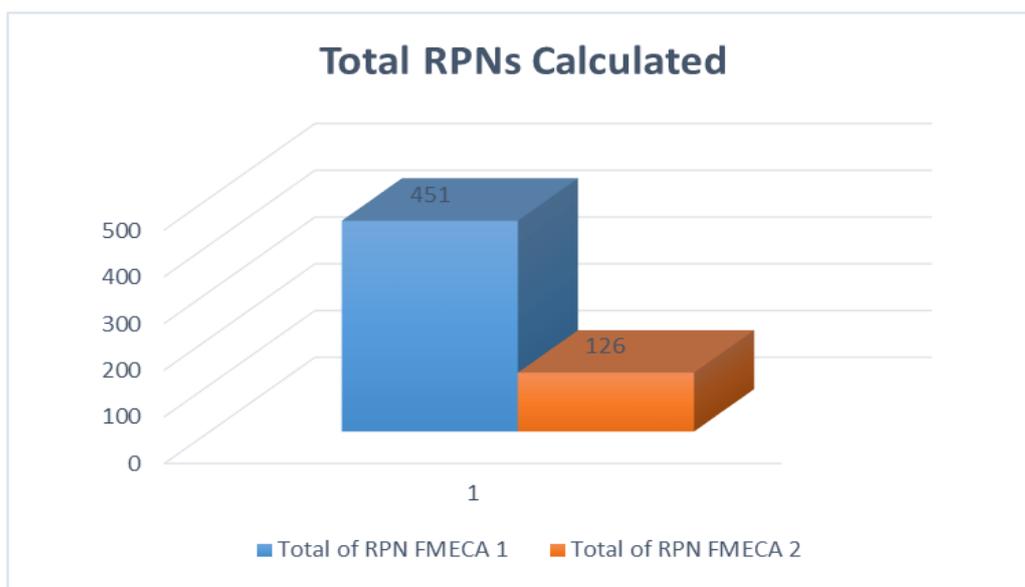


Figure 3: Comparison of RPN Pre and Post FMEAC for process of Granting MAA on SVHA

5 CONCLUSION :

The drug regulatory system is subject to a single issue: user satisfaction and the personal and professional development of its players. However, the equation of effective management and the stakes reached requires a quality system combining vigilance and safety. The tool proposed in this work is basing on the risk analysis using the FMECA method, which also presents itself as a powerful means of aiding in the design of action plans. Moreover, the approach followed with the FMECA method is now more in line with a regulatory framework and mandatory for health institutions, rather than a purely scientific approach. The FMECA is an a priori risk analysis method that is in line with the DMP's policy of improving its services and maintaining its sustainability. Our FMECA is a first learning experience on this particular process. This study has assessed the feasibility and interest of using the FMECA on the process of granting MAA for drugs for human use. It highlighted the complexity of the process and the associated risks 11 failure modes with a total RPN of 451. Following the risk analysis, corrective actions were taking showing a decrease in the total RPN of 72% (126). Monitoring and maintaining these changes in practices over time is often complex. Renewing the FMECA at one year has made it

possible to follow up on the actions taken and monitor the emergence of new risks. It therefore appears to be an effective way to pilot the process of granting MAA for drugs for human use.

6 REFERENCES:

- 2-14-841.pdf. (n.d.) Decree No. 2-14-841 of 19 Chaoual 1436 (August 5, 2015) on the marketing authorization of medicines for human use Ministry of Health Morocco.
- Bhangu, N. S., & Grover, S. (2019). Risk assessment of gear box of wind turbine using FMEA approach. *International Journal of Mechanical and Production Engineering Research and Development*, 9(6), 725–734. <https://doi.org/10.24247/ijmperdddec201961>
- Catelani, M., Ciani, L., Cristaldi, L., Faifer, M., Lazzaroni, M., & Khalil, M. (2015). Toward a new definition of FMECA approach. *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, 2015-July, 981–986. <https://doi.org/10.1109/I2MTC.2015.7151403>
- Darcissac, C., Duvert, L., Hoegy, D., Chappuy, M., Pivot, C., & Janoly-Dumenil, A. (2020). Prospective risk

- analysis in a retrocession unit: Focus on drug dispensation process. *Annales Pharmaceutiques Francaises*, 78(1), 12–20. <https://doi.org/10.1016/j.pharma.2019.06.006>
- DMP. (n.d.). Retrieved November 26, 2020, from <http://dmp.sante.gov.ma/>
- Huang, J., You, J. X., Liu, H. C., & Song, M. S. (2020). Failure mode and effect analysis improvement: A systematic literature review and future research agenda. *Reliability Engineering and System Safety*, 199(November 2019), 106885. <https://doi.org/10.1016/j.ress.2020.106885>
- Maroc, R. (n.d.). *POlitique PHarmaceutique National This document was produced on the basis of the recommendations of the Consultative Commission for Medicines and Health Products (decision of the Minister of Health n°111 of July 9, 2012), on the analytical studies of 2004 - 2012 mentioned in reference and with the technical support of the World Health Organization (WHO)..*
- Meyrieux, C., Garcia, R., Pourel, N., Mège, A., & Bodez, V. (2012). Analyse des risques a priori du processus de prise en charge des patients en radiothérapie : Exemple d'utilisation de la méthode Amdec. *Cancer/Radiotherapie*, 16(7), 613–618. <https://doi.org/10.1016/j.canrad.2012.07.188>
- Poulin, M.-C., & Poirier, L. (2014). Risques, qualité, sécurité : L'analyse des modes de défaillance et de leurs effets : une méthode d'analyse proactive des risques en établissement de santé. *Pharmactuel*, 41, 241–244.
- Regnstrom, J., Koenig, F., Aronsson, B., Reimer, T., Svendsen, K., Tsigkos, S., Flamion, B., Eichler, H. G., & Vamvakas, S. (2010). Factors associated with success of market authorisation applications for pharmaceutical drugs submitted to the European Medicines Agency. *European Journal of Clinical Pharmacology*, 66(1), 39–48. <https://doi.org/10.1007/s00228-009-0756-y>
- Sheets, R. (2017). Marketing authorization. *Fundamentals of Biologicals Regulation: Vaccines and Biotechnology Medicines*, 79–98. <https://doi.org/10.1016/B978-0-12-809290-3.00006-7>
- Spreafico, C., Russo, D., & Rizzi, C. (2017). A state-of-the-art review of FMEA/FMECA including patents. *Computer Science Review*, 25, 19–28. <https://doi.org/10.1016/j.cosrev.2017.05.002>
- Standard 11266, I. (2014). International Standard International Standard. 61010-1 © Iec:2001, 2014, 13.
- ABDELKADER BOUTI and DAOUD AIT KADI . (2554). International Journal of Reliability, Quality and Safety EngineeringVol. 01, No. 04, pp. 515-543 (1994)No Access
- A STATE-OF-THE-ART REVIEW OF FMEA/FMECA <HTTP://LIBRARY1.NIDA.AC.TH/TERMPAPER6/SD/2554/19755>. PDF

CIGI 2021

Continuous Improvement process model: A Knowledge Management approach

DIANA CARDENAS¹, DAVY MONTICOLO¹, Laurent MULLER¹, Pascal LHOSTE¹

¹ ERPI – Equipe de Recherche sur les Processus Innovatifs
Université de Lorraine

8 rue Bastien Lepage, 54000, Nancy - France

cardenas2@univ-lorraine.fr, davy.monticolo@univ-lorraine.fr, l.muller@univ-lorraine.fr, pascal.lhoste@univ-lorraine.fr

Résumé – Le contexte en constante évolution et les défis qu'il représente, amènent les entreprises à la nécessité d'améliorer leur performance de façon permanente. Depuis ses origines et à un rythme d'adoption croissant, l'amélioration continue (vue sous les différentes méthodologies comme le Lean Manufacturing, le Lean Six Sigma, etc.) a joué un rôle important, car sa mise en œuvre est considérée comme un moyen d'améliorer le niveau de performance et les processus. Néanmoins, et malgré les nombreuses études développées, il y a un grand nombre d'échecs sur ces mises en œuvre. Les études actuelles se sont surtout concentrées sur l'identification des facteurs d'échec ou de succès, mais la clarification du processus organisationnel reste inexplorée et mal comprise. En conséquence, suivant la méthode MASK et le méta-modèle KROM, cette étude propose un modèle de processus de l'amélioration continue en mettant en évidence les objectifs, les missions et les cinq processus clés. Ceci est réalisé en capitalisant les connaissances issues des entretiens avec 5 experts sur l'amélioration continue provenant d'entreprises multinationales.

Abstract – The current changing context and challenges it poses, bring companies to the need to improve their performance in a continuous way. For this means, since its origins and at an increasing adoption rate, the Continuous Improvement (seen under the different methodologies like, Lean Manufacturing, Lean Six Sigma, etc.) has played an important role, as its implementation is understood to lead to a better performance level and improved processes. Nevertheless, and despite the numerous developed studies, there is being seen a high number of failures on these implementations. Current studies have been focused mostly on the identification of failure or success factors, but the clarification of the organizational process remains unexplored and misunderstood. Accordingly, following MASK method and KROM meta-modelling, this study proposes a continuous improvement process model evidencing the goals, missions and the five key processes. This is done by capitalizing the knowledge from the interviews with 5 experts on Continuous Improvement from multinational companies.

Mots clés – Amélioration Continue, Amélioration de la performance, Gestion des connaissances, Modélisation des processus.
Keywords – Continuous Improvement, Performance Improvement, Knowledge Management, Process modelling, Knowledge Capitalization.

1 INTRODUCTION

Nowadays, companies are seeking to constantly improve their level of performance. Thus, the Continuous Improvement of the performance in an organization plays a fundamental role. Continuous Improvement has been known over the years as an approach with the main purpose of enhancing the performance in an organization or their processes [Bond, 1999; Gonzalez Aleu & Van Aken, 2016]. It has been widely popularized and implemented through various methodologies, among those, the most important and well-known are Lean Manufacturing, Lean Six Sigma and Six Sigma [Gutierrez-Gutierrez and Antony, 2019; McLean et al., 2017; Message Costa et al., 2018].

It is found in the literature that the widespread adoption of the Continuous Improvement approach not only leads to a better performance level, but it also contributes to the innovation of the organization. And that it is in fact through its implementation that is possible for a firm to promote a culture that responds to the challenges and changes in the context for

enduring its sustainability [Hyland et al., 2007]. This extensive interest, its well-known popularity and widespread implementation is said to be highly motivated by the changes in the business environment, which constantly forces companies to keep on improving their performance in order to remain competitive and be able to fulfill customer demands [Khan et al., 2019; Singh & Singh, 2013].

With no doubt, the Continuous Improvement plays a critical and strategic role for the organizations. Nonetheless, there is not much clarity in the literature on the process. Its management is frequently misunderstood, which mainly starts with the blurry use of the term, because it is both used for the complete process, but also refers to just the outcomes of that process [Bessant et al., 2001]. Also, even though the great interest in research on continuous improvement and the widely disseminated promised benefits, what has been appreciated lately is that there is still a lot of gaps to be filled, gaps that are leading to a high percentage of companies that fail in the implementation of continuous

improvement, and thus a great amount of resources are being wasted [McLean et al., 2017; Sanchez-Ruiz et al., 2020].

In the literature we observe that the reasons for failure are the lack of knowledge on Continuous Improvement, lack of an appropriate measurement and follow up system [Sanchez-Ruiz et al., 2020], and lack of alignment with the strategy of the companies [Middel et al., 2007]. Very valuable studies as the two previously mentioned, have been focused on evidencing or identifying the failure or success factor on continuous improvement implementation, but leaving aside the clarification of its whole process. Which is indeed, of mayor importance, as having a formal and well-established process is critical for successful implementation of improvements [Middel et al., 2007]. Thus, it becomes highly relevant to analyze the process of continuous improvement, but not only from the perspective of the execution of projects or outcomes of the process, as it has been mainly analyzed in studies as the ones from Tennant et al. [2002] and Unzueta et al. [2020], but rather to consider the whole complete process in organizations, which has been poorly researched.

Therefore, in order to have a better management of the Continuous Improvement and consequently, helping in reducing the failure rates of implementations, there is still the lack of an overarching vision of this process in the industry. One that is not only focused on isolated solutions or actions, but that allows to understand the process in a comprehensive manner. Thus, the scientific question that arises is: how to model the process of Continuous Improvement of the performance in the organizations? as this is not clear in the literature. Consequently, the aim of this study is to propose a unified model of this process, through MASK, a Knowledge Management approach based on the capitalization of experts' knowledge [Ermine, 2013]. From which the identification of five principal processes of the continuous improvement is done. For doing so, the next section reviews the concepts of Continuous Improvement and Knowledge Management. Then, the methodology for this study is explained in the third section. Followed by the results and discussion in the fourth section. And finally, in the fifth section the conclusions are shown.

2 BACKGROUND

2.1 Continuous Improvement

From its beginnings, Continuous Improvement has been seen as a key element in the performance improvement, constantly increasing its adoption rate [Hyland et al., 2007]. According to the literature, Continuous Improvement was originated in Japan after the Second World War, was then when it was started to be used as an approach and philosophy for performance improvement and not only as isolated initiatives. Becoming then, one of the foundations of the success of the Japanese manufacturing [Singh and Singh, 2015]. As highlighted by [Sanchez and Blanco, 2014], some of the first definitions of Continuous improvement date from 1982 and 1989 by authors Deming and Imai, respectively. From which, it is highlighted its iterative characteristic and the involvement of the different levels at the organization.

The concept has been evolving along with the changes in the environment, it has been constantly ratifying its great importance for organizations. Being seen as an umbrella term and as a necessity for organizations in order to have a better organizational performance, to remain competitive in the

market and to adapt to the new context [Khan et al., 2019; Singh and Singh, 2013]. It is certainly a process in which a key factor is the knowledge and involvement of the entire organization [McLean et al., 2017; Singh and Singh, 2015].

Despite this, continuous improvement is a subject that still needs further research and to be analyzed in a more complete way. In this sense, there are studies, like the one from [Almaiman and McLaughlin, 2018], who identifies the cultural aspects that facilitate Continuous Improvement, while stating its importance for achieving the desired benefits. Also, the one from [Heavey et al., 2014] that proposes a framework for continuous improvement by determining its key forces for improving customer value. And finally, the one from [Berger, 1997] who evidences the core principles that are the basis of the Continuous Improvement. Even with the importance and relevance of those studies, there is still lacking a clarification on what is the process of Continuous Improvement.

2.2 Knowledge Management

Knowledge Management has been largely studied, having over the last two decades a noticeable increase of interest from academic researchers as well as from practitioners [Lee and Chen, 2012; Sarka et al., 2019]. It is being understood as mechanisms and processes through which is enabled the creation, sharing and re-utilization of knowledge [Amaral et al., 2016; Bolis et al., 2012; Mårtensson, 2000; Poage, 2003]. Being thus, strategic in the era of the knowledge economy, and aiming at creating value, developing intangible resources, and improving the competitiveness in organizations [Patil and Kant, 2014].

The theory states that knowledge is created through the social interaction of individuals and organizations, and this creation process involves the interaction of two types of knowledge, the tacit and the explicit one, hence enabling a continuous process of knowledge creation [Nonaka et al., 2000]. This process involves four modes of knowledge creation, which are, socialization, externalization, combination, and internalization. It expresses the knowledge creation or crystallization as a never-ending spiral that starts with the individual tacit knowledge level that is subsequently being externalized, transformed into explicit knowledge and then shared among the organization, for ultimately being internalized by individuals, providing a greater understanding and then generating more tacit knowledge [Nonaka and von Krogh, 2009].

Knowledge management builds up on the theory stated by Nonaka and Takeuchi [1995]. It is in fact a strategic approach that supports the knowledge creation process [Matta et al., 2001], which is focused on analyzing the organization knowledge as a resource, and with the main objectives of knowledge capitalization (valorize the knowledge and preserve it), sharing (not only the knowledge circulation, but to reach a collective intelligence) and creation (linked to the permanent innovation of organizations) [Ermine and La, 2003].

This process for knowledge capitalization, the externalization from tacit knowledge of the experts to explicit one, can be either done by an approach of only knowledge transcriptions, or under a knowledge engineering approach, which in the literature is considered to be even more complete and robust [Ermine and La, 2003]. Among the different methods of knowledge engineering, MASK (Method for Analysis and Structuring

Knowledge) stands out [Matta et al., 2001]. MASK is an evolution of knowledge management of nearly 30 years, it is a well-structured method which allows the knowledge capitalization from experts in a specified subject, it is based on the interviews to experts, the modeling of the knowledge and the model validation from the experts [Aries et al., 2008].

3 METHODOLOGY

With the purpose of providing a clearer understanding and modelling the process of the Continuous Improvement, the following methodology is proposed (Figure 1). This is based on a knowledge management approach, more specifically MASK method and KROM (Knowledge Reuse Organizational Meta-Model) [Girodon et al., 2012]. This methodology, under the approach of Knowledge Management aims at capitalizing the knowledge from experts in the industry, to transform their tacit knowledge on explicit [Matta et al., 2001], for consequently sharing it and make it available to the use of the scientific community, and thus continue to contribute to the cycle of knowledge in the subject.

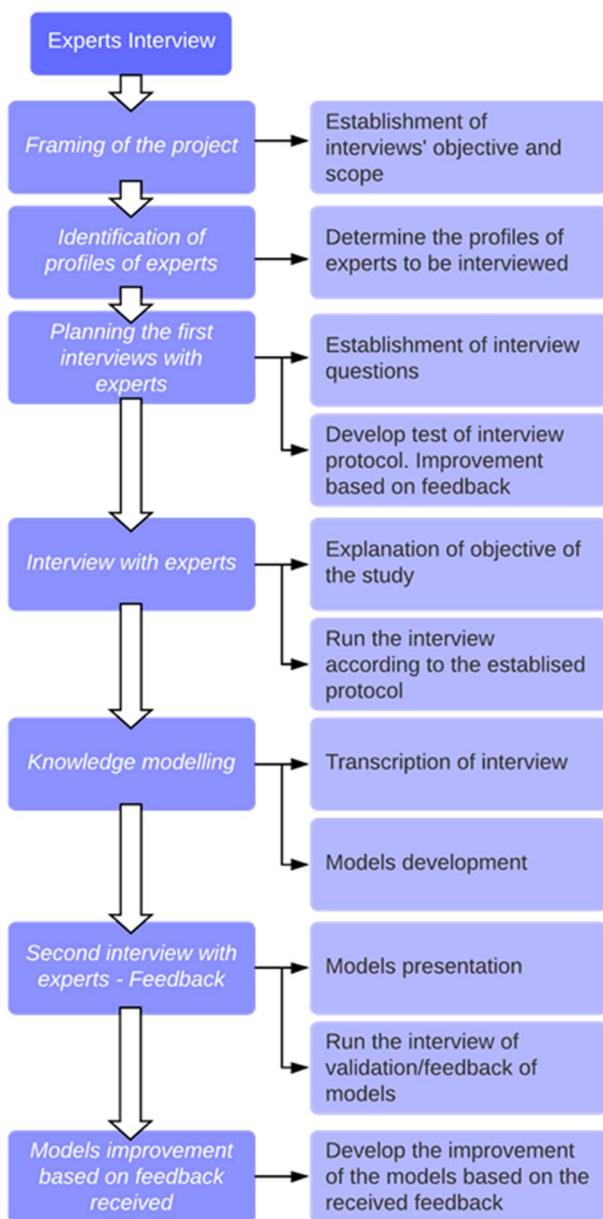


Figure 1. Methodology

For this study, the steps of MASK method as described in [Ermine, 2013; Ermine and La, 2003] were followed: starting with the framing of the knowledge capitalization project by establishing the objectives and scope. Followed by the identification of the profiles of the experts to be interviewed, for this, different organizations and locations were privileged, as experts' diversity is often seen as an important factor for taking into account different social practices [Nonaka and von Krogh, 2009]. Then, the planning of the interviews was done, for which, a series of questions and a protocol of interview were established and then tested to further improve them. After that, it is the actual interviews to the selected experts.

For the next step, based on the interviews, the knowledge modelling is done. For this study we opted to develop the knowledge modelling with KROM meta-model, as it is centered on organizational modelling based on knowledge [Girodon et al., 2012; Monticolo, 2015]. Through it, the organizational and process models are developed, for making it possible to clarify the Continuous Improvement process, by showing its goals, missions, subprocesses and the activities within them. After developing the models, an important step to follow is to develop a second interview with experts in order to have their validation on the gathered knowledge and feedback on the models for further improving them and confirm its accuracy.

4 RESULTS AND DISCUSSION

By following the MASK method, the first thing that was done was the framing of the project of knowledge capitalization, which is shown in (Table 1) and contains the objective and five w's and how. This is a very important step to do, as this gives us clarity regarding the scope of the knowledge field to clarify.

Table 1. Framing of Knowledge Capitalization project

Framing of the project	
Objective	To access and capitalize the knowledge of experts in Continuous Improvement, with the objective to understand and clarify its processes, activities, characteristics.
Five W's and How	
What	The domain of study is the process of Continuous Improvement in the industry.
Who	The group of experts to interview are industry professionals with vast experience on Continuous Improvement.
Where	The interviews are developed by videoconference due to the actual situation and for having access to a diverse group of experts.
When	Each interview is designed to last 1 hour maximum. The total length of the process, including modeling and second round of interviews (experts' feedback on built models) is about 3 to 4 months.
How	A first interview with a defined protocol is done with the experts for gathering information. Following, after analyzing and structuring the information, a second round of interviews with available experts is developed to validate the structured models and receive feedback for further improvement.
Why	Because there is a need to give clarity on how the actual organizational process of the

Continuous Improvement in the industry is, looking towards a lower failure ratio on implementations and a future transformation of the process.

Based on the described framing, the profile of the experts to be interviewed was determined, also the protocol of interview with the proposed questions was defined, taking into account the models to be used. In order to validate and improve this protocol prior to interviewing the experts, a mock interview was developed with one of the authors, who has more than 20 years of experience in the subject to study. With the intention to get a diversity of context on the experts for proposing more generalized models, a total of 5 experts from 4 different companies (located in France, Switzerland, Mexico, United States) were interviewed. All experts have a vast experience in Continuous Improvement, with positions such as Managers, Directors and Expert specialists in Process and Operational Excellence, Lean, Production and Continuous Improvement.

After gathering and analyzing the information from each of the interviews, the authors proceeded to capitalize the experts' knowledge by means of the models proposed by KROM. Since the interview to the first expert, the models were developed and consequently improved with the knowledge from the subsequent experts. By this, it was possible to observe that besides the particularities of the context of each expert, their knowledge on the subject was consistent. It was observed that the experts shared the same vision on Continuous Improvement and through the process even some of their answers were very similar, meaning in this way that a good quality and pertinence of the knowledge that captured under this Knowledge Management approach.

To confirm the relevance of the models developed, a second round of interviews with two of the experts was conducted. In these, as suggested in the MASK method [Aries et al., 2008], the aim is to obtain feedback and validation of the models by the experts, which process is considered sufficient when there is nothing more to add to the models. One of these second interviews was at the beginning of the process, in order to validate the early stages of the models; and the other one at the end of the process, after capitalizing the knowledge from all of the experts and consolidating them into one set of unified and comprehensive models.

As results of the capitalization of the knowledge, Figure 2 and Figure 3 show respectively the organizational structure model and the process model for Continuous Improvement. The model in Figure 2 is structured for representing the three principal missions of Continuous Improvement: to give support to the improvement of the processes in the company, to guide the establishment of the performance indicators, and to maintain a high level of performance. At the same time, it can be seen that these missions contribute to the achievement of the following goals: to excel from the competition, to quickly adapt to customer's (either external or internal) expectations, and to allow the company to sustain itself over time. Representing in this way the identity of the Continuous Improvement, which is in fact very important to clarify, as we have seen that one of the failure reasons in implementations is actually the lack of alignment with the company strategy [Middel et al., 2007], as it is often not seen as the whole organizational process.

Moreover, for fulfilling the identified missions, they rely on five processes, namely: to give support to the problem resolution, support the performance strategy, to train and coach the

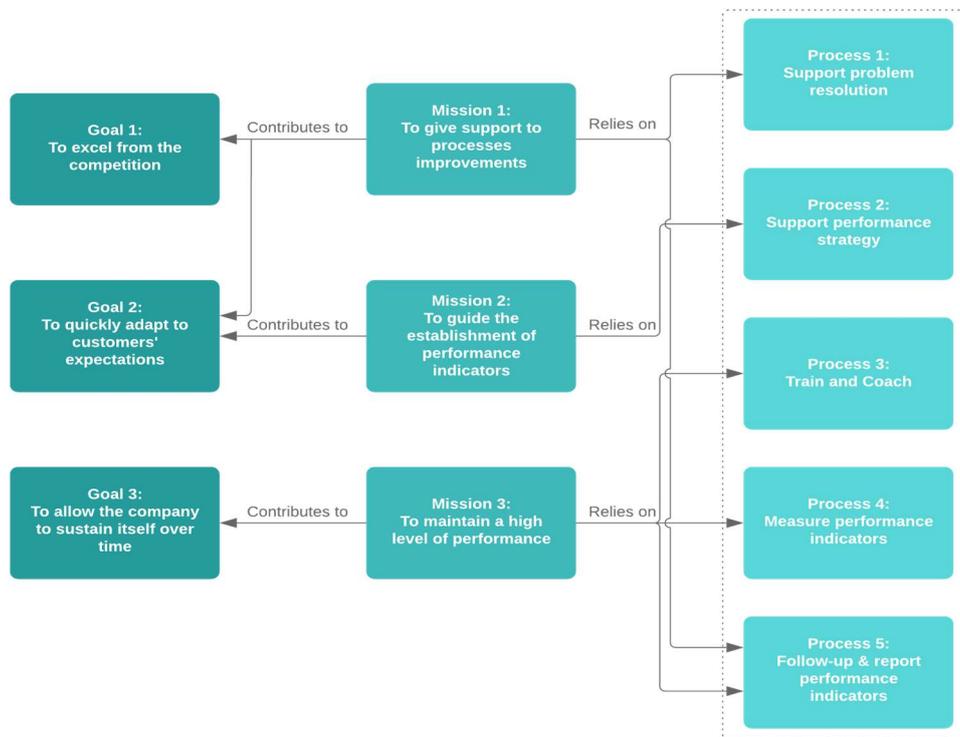


Figure 2. Continuous Improvement Organizational structure model

to map the whole organizational process and to identify the activities within them, its entries (e), exits (s), controls (c) and mechanisms (m). From the five identified processes, as an example, in Figure 4 and Figure 5 we represent “support problem resolution” and “support performance strategy” respectively.

In the process model for “support problem resolution” (Figure 4) are represented all the activities that it is constituted by. In this, it was evidenced the critical importance of well-defined performance indicators, as well as the data reliability of its measures, in which the role of data manager is lately becoming more and more valuable. This process receives as controls and entries, among others, the performance indicators measure, their target levels, and their level of conformity.

These come from other of the defined processes (follow-up and report performance indicators, support performance strategy, measure performance indicators), and become crucial for the right development of this process in order to obtain the desired outputs (lessons learned, action plan performance indicators, improved processes). Which highlights once again the need to see the Continuous Improvement as a whole organizational process and not as isolated bricks.

Likewise, in Figure 5 is shown the process model for “support performance strategy”. This represents a very crucial process, as in its activities we find: to support the establishment of the mission and vision of the organization, from which the performance indicators are identified and defined with its target levels (generally done under a top-down approach, but taking into account that the reporting usually follows a bottom-up one), then the dashboard in which they are going to be reported is

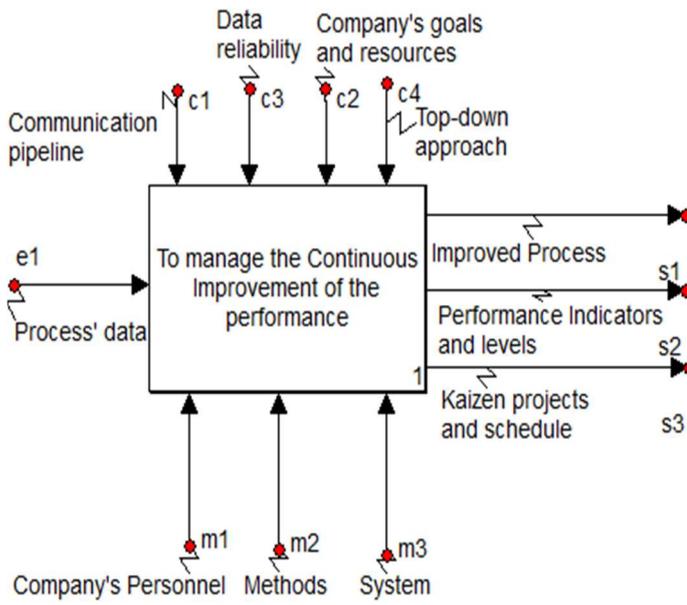


Figure 3. Continuous Improvement diagram A-0

personnel, measure the performance indicators, and follow up and report the established performance indicators. These processes can be consequently represented by the activities within them. This is done by the process model proposed by KROM, which is based on SADT/IDEF0. Thus, following in Figure 3 is represented the model of the whole process based on SADT, in which each of the five sub processes are contained. From the interviews it was possible to not only identify for each mission which are the processes that they rely on and how these missions are contributing to the goals. But also, it was possible

To support problem resolution

A1

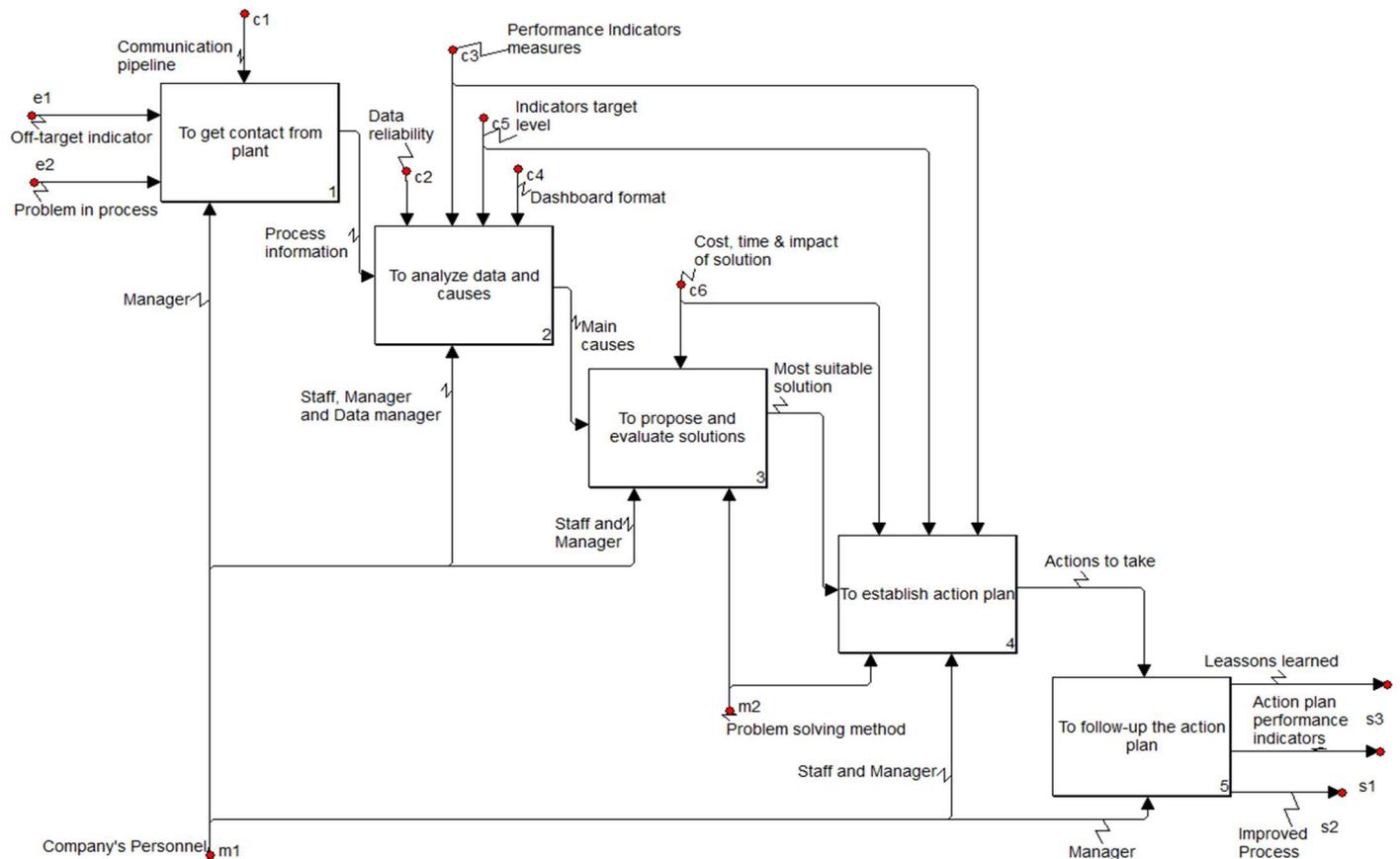


Figure 4. Process model: Support problem resolution

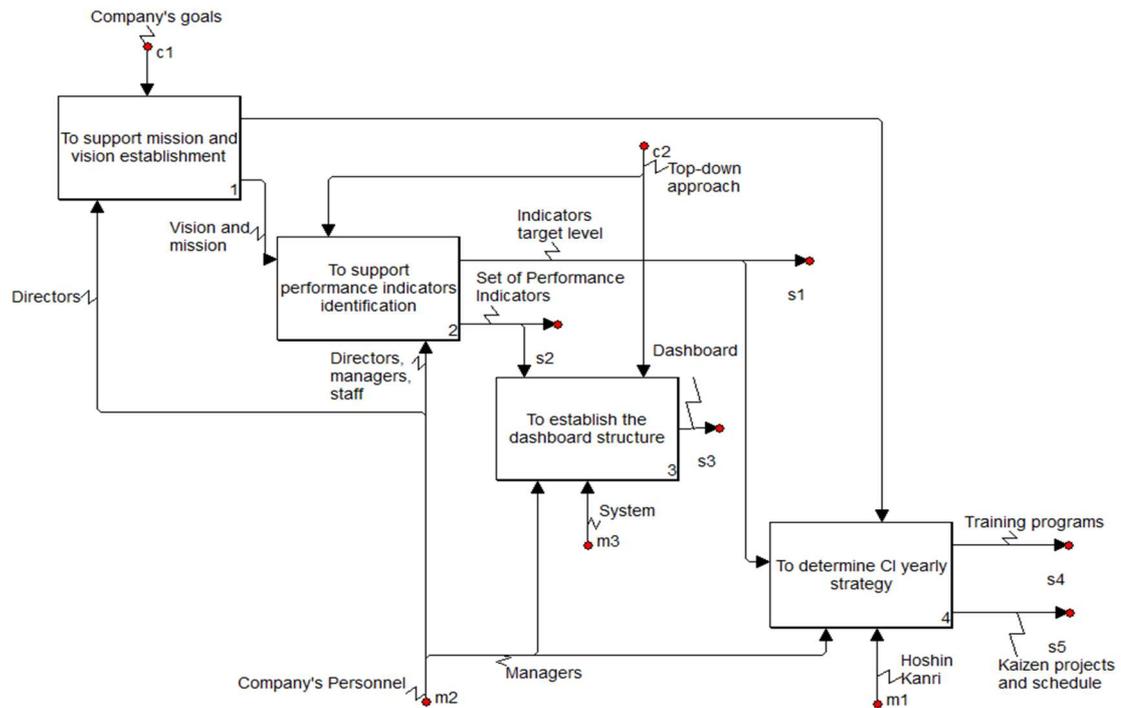


Figure 5. Process model: Support performance strategy

structured, and also it is determined the continuous improvement yearly strategy, with the training programs and kaizen projects and their schedule. Throughout the interviews, it was evidenced that even if the identification of the performance indicators plays a highly important role, it is mainly done by instinct. Its identification and dependability connection between the different levels of indicators (low, medium, high) represents without doubt an important room for further improvement and systematization.

Consequently, with the development of these models under a Knowledge Management approach, we seek to give more clarity on the organizational process of the Continuous Improvement, which is important and vital for the survival of the companies, but so far poorly studied as a complete process. A general and mostly agreed opinion among the experts is that Continuous Improvement is fundamental for companies seeking to be adaptable and competitive. It must be considered as an essential part of its strategy, no matter under which specific methodology it is implemented (Lean, Six Sigma, Lean Six Sigma, TPS...) or the way it is represented in the companies (a whole department, a group of experts, a shared philosophy with each process owner empowered). Thus, a clearer way to understand it, is shown in this paper, which is expected to be seen as a reference for the as-is process and later analyze the to-be one.

Moreover, the developed models can be confronted with another reference describing performance improvement requirements: ISO 9000 standards. In ISO 9001 standard (Fundamentals and vocabulary) [International Organization for Standardization, 2015], one of the seven principles of quality management is improvement. This principle reminds us that the success of an organization is based on a constant desire for improvement and proposes to bring the organization's processes under control. In ISO 9001 standard (Requirements), chapter 9 (Performance

assessment) and chapter 10 (Improvement) are directly concerned with performance improvement, they broadly take up the elements presented in our model. Finally, in the ISO 9004 standard (Quality of an organization - Guidance to achieve sustained success) [International Organization for Standardization, 2018], the term improvement (in different forms) appears 146 times, thus, reflecting the great importance of this notion in this quality standard. By analyzing the standard, we believe that the models developed are in line with the standards, they are complementary. In our models we can find the processes and activities in detail, but the standard, in contrast with our models, opens the scope to the outside of the organization (in chapter 5.2 involves the interested parties for detecting opportunities for sustainable performance improvement, and chapter 10.4.3 talks about comparing the organization's performance with agreed benchmarks) and considers the practice of internal audits and self-assessments with the staff. The standards give a much more general vision, nevertheless with our models obtained from the capitalization of expert knowledge, we propose a more detailed vision that can become a great insight for practitioners.

5 CONCLUSIONS

This study developed the organizational and process models of the Continuous Improvement, intending to give a better understanding, highly motivated by its fundamental role for organizations. This was developed under a Knowledge Management approach, following the MASK method, and capitalizing the knowledge of the experts through the models proposed by KROM. This study is providing in the results, a model depicting the three goals, three missions and five processes of the organizational process of Continuous improvement, as well as the activities within two of these processes. With this, we intend to give more clarity and to help towards the creation and sharing of knowledge in order to help

to reduce the failure rate in the implementation of Continuous Improvement initiatives.

Furthermore, these models represent how the Continuous Improvement is being managed and developed in the industry nowadays, however the authors believe it is important to be conscious that it should be evolving along with the changes in the environment, like the fourth industrial revolution, which should take us to further adapt and improve it. In this regard some key elements were identified in the interviews, like the importance of continuously tracking the performance and the analysis and reliability of the data, for improving the decisions management and base it on data instead of just intuition.

The results from this study could lead to various perspectives and applications, with the developed models, a more understandable view of the Continuous Improvement is given, which could become the basis for guiding practitioners through its implementation and for developing a diagnosis of the Continuous Improvement process in the companies. It could allow based on this process, to propose a decision-making system for helping engineers and managers to better analyze and improve the performance. It could as well be the starting point towards an analysis for the improvement of the process, as for this, having a clear understanding of the process is a must. And finally, for developing the digitalization of the workflow of the process in order to have a better visibility and control.

6 REFERENCES

- Almaitan, S. and McLaughlin, P. (2018), "Facilitating a continuous improvement culture: A literature review", *Advances in Transdisciplinary Engineering*, Vol. 8, IOS Press BV, pp. 493–498. DOI:10.3233/978-1-61499-902-7-493.
- Amaral, A., Araújo, M.M. and Rodrigues, C.S. (2016), "Organizational learning and knowledge management—insights from industrial managers", *Advances in Intelligent Systems and Computing*, Vol. 490, Springer Verlag, pp. 403–415. DOI:10.1007/978-3-319-41697-7_36.
- Aries, S., Le Blanc, B. and Ermine, J.-L. (2008), *MASK : Une Méthode d'ingénierie Des Connaissances Pour l'analyse et La Structuration Des Connaissances. Management et Ingénierie Des Connaissances: Modèles et Méthodes*, Hermes Science Publications-Lavoisier.
- Berger, A. (1997), "Continuous improvement and kaizen: Standardization and organizational designs", *Integrated Manufacturing Systems*, Vol. 8 No. 2, pp. 110–117. DOI:10.1108/09576069710165792.
- Bessant, J., Caffyn, S. and Gallagher, M. (2001), "An evolutionary model of continuous improvement behaviour", *Technovation*, Vol. 21 No. 2, pp. 67–77. DOI:10.1016/S0166-4972(00)00023-7.
- Bolis, I., Brunoro, C. and Sznalwar, L.I. (2012), "The workers role in knowledge management and sustainability policies", *Work*, Vol. 41, pp. 2713–2720. DOI:10.3233/WOR-2012-0515-2713.
- Bond, T. (1999), "The role of performance measurement in continuous improvement", *International Journal of Operations and Production Management*, Vol. 19 No. 12, pp. 1318–1334.
- Ermine, J.-L. (2013), "Knowledge Management with the MASK method", *Knowledge Management for Sustainable Development*, available at: <https://hal.archives-ouvertes.fr/hal-0208044>.
- Ermine, J.-L. and La, J.-L.E. (2003), *La Gestion Des Connaissances*, Vol. 2, Hermes Sciences Publications, available at: <https://hal.archives-ouvertes.fr/hal-00997696>.
- Girodon, J., Monticolo, D., Bonjour, E. and Perrier, M. (2012), "KROM: An organizational Meta-Model oriented to knowledge: A case from Ophthalmic industry", 8th International Conference on Signal Image Technology and Internet Based Systems, SITIS 2012r, Vol. 2, pp. 845–851. DOI:10.1109/SITIS.2012.127.
- Gonzalez Aleu, F. and Van Aken, E.M. (2016), "Systematic literature review of critical success factors for continuous improvement projects", *International Journal of Lean Six Sigma*, Emerald Group Publishing Ltd., Vol. 7 No. 3, pp. 214–232. DOI:10.1108/IJLSS-06-2015-0025.
- Gutierrez-Gutierrez, L. and Antony, J. (2019), "Continuous improvement initiatives for dynamic capabilities development: A systematic literature review", *International Journal of Lean Six Sigma*, Vol. 11 No. 1, pp. 125–149. DOI:10.1108/IJLSS-07-2018-0071.
- Heavey, C., Ledwith, A. and Murphy, E. (2014), "Introducing a new continuous improvement framework for increased organisational return on investment", *TQM Journal*, Emerald Group Publishing Ltd., Vol. 26 No. 6, pp. 594–609. DOI:10.1108/TQM-06-2013-0065.
- Hyland, P.W., Mellor, R. and Sloan, T. (2007), "Performance measurement and continuous improvement: Are they linked to manufacturing strategy?", *International Journal of Technology Management*, Vol. 37 No. 3–4, pp. 237–246. DOI:10.1504/IJTM.2007.012260.
- International Organization for Standardization. (2015), *Quality Management Systems - Requirements (ISO Standard No. 9001:2015)*.
- International Organization for Standardization. (2018), *Quality Management - Quality of an Organization - Guidance to Achieve Sustained Success (ISO Standard No. 9004:2018)*.
- Khan, S.A., Kaviani, M.A., Galli, B.J. and Pharma, A. (2019), "Application of continuous improvement techniques to improve organization performance A case study", *International Journal of Lean Six Sigma*, Vol. 10 No. 2, pp. 542–565. DOI:10.1108/IJLSS-05-2017-0048.
- Lee, M.R. and Chen, T.T. (2012), "Revealing research themes and trends in knowledge management: From 1995 to 2010", *Knowledge-Based Systems*, Elsevier, Vol. 28, pp. 47–58. DOI:10.1016/J.KNOSYS.2011.11.016.
- Mårtensson, M. (2000), "A critical review of knowledge management as a management tool", *Journal of Knowledge Management*, Vol. 4 No. 3, pp. 204–216. DOI:10.1108/13673270010350002.
- Matta, N., Ermine, J.-L., Aubertin, G. and Trivin, J.-Y. (2001), "How to capitalize knowledge with the MASK method?", *IJCAI 2001 Workshop on Knowledge Management and Organizational Memories*, Seattle, Washington USA, August, Vol. 6, pp. 1–13.
- McLean, R.S., Antony, J. and Dahlgard, J.J. (2017), "Failure of Continuous Improvement initiatives in manufacturing environments: a systematic review of the evidence", *Total Quality Management & Business Excellence*, Vol. 28 No. 3–4, pp. 219–237. DOI:10.1080/14783363.2015.1063414.
- Message Costa, L.B., Filho, M.G., Fredendall, L.D., José, F. and Paredes, G. (2018), "Lean, six sigma and lean six sigma in the food industry: A systematic literature review", *Trends in Food Science and Technology*, Vol. 82, pp. 122–133. DOI:10.1016/J.TIFS.2018.10.002.
- Middel, R., Op De Weegh, S. and Gieskes, J. (2007), "Continuous improvement in the Netherlands: A survey-based study into current practices", *International Journal of*

- Technology Management, Inderscience Publishers, Vol. 37 No. 3–4, pp. 259–271. DOI:10.1504/IJTM.2007.012262.
- Monticolo, D. (2015), *Approches Organisationnelles Pour La Conception de Systèmes Multi-Agents Dédiés à La Gestion Des Connaissances ; Application Aux Projets d ' Ingénierie et d'Innovation*, Université de Lorraine.
- Nonaka, I. and von Krogh, G. (2009), "Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory", *Organization Science*, Vol. 20 No. 3, pp. 635–652. DOI:10.1287/ORSC.1080.0412.
- Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford university press.
- Nonaka, I., Toyama, R. and Konno, N. (2000), "SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation", *Long Range Planning*, Vol. 33 No. 1, pp. 5–34. DOI:10.1016/S0024-6301(99)00115-6.
- Patil, S.K. and Kant, R. (2014), "Methodological literature review of knowledge management research", *Tékhné, Elsevier BV*, Vol. 12 No. 1–2, pp. 3–14. DOI:10.1016/J.TEKHNE.2014.07.001.
- Poage, J.L. (2003), "Covering the Intangibles in a KM Initiative", *IT Professional*, Vol. 5 No. 6, pp. 17–23. DOI:10.1109/MITP.2003.1254964.
- Sanchez-Ruiz, L., Gomez-Lopez, R. and Blanco, B. (2020), "Barriers to effectively implementing continuous improvement in Spanish firms", *Total Quality Management and Business Excellence*, Routledge, Vol. 31 No. 13–14, pp. 1409–1426. DOI:10.1080/14783363.2019.1699783.
- Sanchez, L. and Blanco, B. (2014), "Three decades of continuous improvement", *Total Quality Management & Business Excellence*, Vol. 25 No. 9–10, pp. 986–1001. DOI:10.1080/14783363.2013.856547.
- Sarka, P., Heisig, P., Caldwell, N.H.M., Maier, A.M. and Ipsen, C. (2019), "Future research on information technology in knowledge management", *Knowledge and Process Management*, John Wiley and Sons Ltd, Vol. 26 No. 3, pp. 277–296. DOI:10.1002/KPM.1601.
- Singh, J. and Singh, H. (2013), "Continuous Improvement Strategies: An Overview", *IUP Journal of Operations Management*, Vol. 12 No. 1, pp. 32–57. DOI:10.4324/9781315154237.
- Singh, J. and Singh, H. (2015), "Continuous improvement philosophy – literature review and directions", *Benchmarking*, Emerald Group Publishing Ltd., Vol. 22 No. 1, pp. 75–119. DOI:10.1108/BIJ-06-2012-0038.
- Tennant, C., Warwood, S.J. and Chiang, M.M.P. (2002), "A continuous improvement process at Severn Trent Water", *TQM Magazine*, MCB UP Ltd, Vol. 14 No. 5, pp. 284–292. DOI:10.1108/09544780210439716.
- Unzueta, G., Esnaola, A. and Eguren, J.A. (2020), "Continuous improvement framework to develop cultural change: case study, capital goods company", *TQM Journal*, Emerald Group Holdings Ltd., Vol. 32 No. 6, pp. 1327–1348. DOI:10.1108/TQM-02-2019-0051.

**Session 2.3 - Session spéciale 3 et 7 :
Surveillance intelligente des processus
et maintenance prédictive pour
l'industrie du futur**

Chairman : Nassim Boudaoud - Zineb Simeu

CIGI 2021

Tolérancement GPS et Pilotage

THOMAS MULLER^{1,2}, MAURICE PILLET¹, JEAN LUC MAIRE¹, DAVY PILLET²

¹ Université Savoie Mont Blanc
Annecy, France
thomas.muller@univ-smb.fr
maurice.pillet@univ-smb.fr
jean-luc.maire@univ-smb.fr
² Ellistat
Annecy, France
Davy.pillet@ellistat.com

Résumé – Dans la construction d'un plan de surveillance, on cherche à garantir la conformité du produit pour le client. Si la relation entre cette conformité et le pilotage de la production est relativement simple à obtenir dans le cas d'un intervalle de tolérance dimensionnel (10 ± 0.1), cette relation est plus complexe dans le cas d'un tolérancement géométrique (GPS).

Nous proposons dans cette communication une démarche structurée d'analyse des phases de fabrication dans le but de générer un plan de surveillance garantissant la fonctionnalité de la pièce, et cela tout en limitant au strict nécessaire les contrôles sur poste.

La situation de capacité et les dérives potentielles sont analysées en fonction du processus de fabrication et selon les six composantes de la spécification géométrique.

Cette analyse conduit à la construction d'un plan de surveillance qui sépare deux types de contrôle :

- les contrôles hors ligne qui peuvent être réalisés sur une machine à mesurer tridimensionnelle déportée.
- les contrôles en ligne sur le poste qui nécessitent un retour d'information immédiat pour permettre le pilotage du processus.

Dans la démarche proposée, nous dissocions clairement les mesures destinées à la vérification de la conformité de celles destinées au pilotage des processus. Quelques exemples illustrent la démarche proposée.

Abstract – In the construction of a monitoring plan, the aim is to guarantee the conformity of the product for the customer. If the relation between this conformity and the production control is relatively simple to obtain in the case of a dimensional tolerance interval (10 ± 0.1), this relation is more complex to obtain with a geometric tolerance (GPS).

In this paper, we propose a structured approach for analyzing the manufacturing phases in order to generate a monitoring plan that guarantees the functionality of the part, while limiting on-site controls to the strict necessary.

The capability situation and potential drifts are analyzed according to the manufacturing process and the six components of the geometrical specification.

This analysis leads to the construction of a monitoring plan that separates two types of control:

- off-line controls which can be carried out on a remote coordinate measuring machine.
- on-line controls on the workstation which require immediate feedback to allow process control.

In the proposed approach, we clearly dissociate the measurements intended for checking compliance from those intended for process control. A few examples illustrate the proposed approach.

Mots clés – GPS, Tolérancement, Pilotage, Plan de surveillance, Décomposition modale

Keywords – GPS, Tolerancing, Control, Monitoring Plan, Modal decomposition

1 INTRODUCTION

Le tolérancement géométrique GPS [ISO5458, 1998] [Anselmetti, 2011] (Geometrical Product Specification) est de plus en plus utilisé dans le monde de l'automobile, mais également dans l'ensemble de l'industrie mécanique. Incontestablement, le tolérancement géométrique permet de mieux exprimer la conformité du produit répondant au besoin fonctionnel.

Cependant, si la fonctionnalité est mieux décrite par ce type de tolérancement, celui-ci introduit plusieurs problèmes au niveau des processus de production habitués à des tolérances dimensionnelles de type bipoint " 10 ± 0.2 ".

Une dérive couramment constatée avec le tolérancement GPS consiste à transformer le tolérancement géométrique en dimensions locales et à donner aux opérateurs un plan de fabrication dans lequel le tolérancement a été simplement « traduit ». Cette façon de concevoir un plan de contrôle présente trois inconvénients majeurs :

1. On ne dissocie pas la vérification de la conformité et les actions de pilotage du processus
2. La traduction amène souvent à resserrer inutilement les tolérances ce qui entraîne une augmentation des coûts de production

- On ne garantit pas toujours la conformité des spécifications géométriques qui sont pourtant essentielles pour assurer le besoin fonctionnel

Une autre dérive consiste à fournir en production directement le plan de définition du produit exprimé en tolérancement géométrique, à charge pour le régleur de garantir les spécifications à partir de ce plan. Cette façon de concevoir le pilotage d'une production a trois inconvénients majeurs :

- On ne réalise pas une analyse du processus qui permet pourtant de garantir la conformité des spécifications géométriques
- On confond conformité et pilotage de la production
- On surcharge la production en contrôles inutiles qui réduisent la productivité

Nous proposons dans ce papier une méthodologie permettant à la fois de donner des règles de pilotages non ambiguës et de garantir la conformité des spécifications à moindre coût de contrôle.

2 PILOTAGE ET CONFORMITE

2.1 Le but de l'industrialisation

Une industrialisation idéalement réussie d'un produit serait une industrialisation qui permettrait d'obtenir la garantie de la conformité du produit sans qu'il y ait besoin de le mesurer ou sans qu'une intervention humaine soit nécessaire. Evidemment, nous sommes souvent loin de cet idéal et nous devons, tout au long des processus de production, contrôler les produits pour s'assurer de leur conformité. Cependant, au-delà de la simple conformité, l'objectif du pilotage d'un processus de production est de placer le plus de caractéristiques du produit sur leurs cibles afin de limiter le coût de non-qualité en accord avec la fonction perte de Taguchi [Taguchi, 1987].

La conduite d'un processus doit satisfaire deux objectifs qui sont malheureusement souvent confondus :

- La vérification de la conformité du produit doit s'appuyer sur le dessin de définition du produit
- La limitation de la variabilité autour de la cible doit s'appuyer sur le respect de règles de pilotage

Le second point doit prendre en compte le souci de l'automatisation des ajustements pour aller vers une plus grande autonomie des machines [Dittrich et al., 2019a] [Möhring 2020].

2.2 Les différents types de dérives

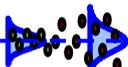
L'observation des différentes caractéristiques rencontrées en mécanique nous amène à établir le tableau des différentes dérives ou variations qui peuvent être constatées sur une machine (Tableau 1). Nous pensons que ce tableau est quasi exhaustif (nous sommes évidemment preneurs de toute autre situation ne correspondant à aucune situation listée dans ce tableau).

A chacune de ces situations correspond un type de contrôle parfaitement adapté (voir Tableau 1).

Si on élimine le cas de l'évènements erratique qui peut subvenir durant la production et qui contraint à un contrôle à 100%, on s'aperçoit que seules les dérives permanentes ou non prévisibles en dispersion et en position nécessitent un suivi périodique à fréquence élevée.

Cette situation n'est pourtant pas celle que nous rencontrons le plus souvent dans les entreprises dans lesquelles une grande proportion des caractéristiques fait l'objet d'un suivi périodique, voire même d'un contrôle à 100% !

Tableau 1 – Récapitulatif des différentes dérives

	Schéma	Exemples	
		Ponctuel ou Arrivée prévisible	Permanent ou Arrivée non prévisible
Processus stable		Géométrie copie de la géométrie de la machine. Entraxe dans outillage de découpe.	Validation à la conception du processus
Décentrage		Changement outils. Remontage machine.	Casse, collision... Problème sur un lot de matière première
		Contrôle systématique après événement causal	Contrôle libératoire fréquentiel
Dérive en position		Machine froide Premières minutes de fonctionnement	Usure d'outils. Variations thermiques
		Contrôle systématique après événement causal	Dérives d'usure, dérives thermiques
Dérive en dispersion		Défauts de localisation (outil mal centré dans la broche)	Usure d'un foret Jeux processus
		Contrôle systématique après événement causal	Dérives d'usure, dérives thermiques
Evénement			Copeaux sur butée. Erreur opérateur.
			Poka-Yoke, Contrôle à 100%

2.3 Les contrôles "off-line" et "on-line"

S'il est clair que le contrôle périodique à fréquence élevée doit se faire au plus proche du moyen de production pour pouvoir réagir au plus vite en cas de besoin, cela n'est pas forcément nécessaire pour les autres types contrôles qui requièrent moins de proximité et une fréquence moins élevée. Lorsque l'industrialisation a été réalisée correctement, tous les autres contrôles n'ont pas vocation à induire un pilotage du processus mais permettent simplement de vérifier la conformité des caractéristiques. Il est donc tout à fait envisageable de réaliser ces contrôles en off-line et avec une fréquence plus faible.

Nous dissociions ainsi deux types de contrôles :

- Les contrôles on-line : ils ont pour vocation d'induire un pilotage du processus et doivent être fait au plus proche de la machine de façon synchrone à celle-ci.
- Les contrôles off-line : ils ont pour objectif de surveiller la conformité des produits qui découlent d'une bonne industrialisation. Ils surveillent les évènements improbables et ne visent à induire des corrections sur le processus. Ils peuvent être réalisés de façon asynchrone au processus.

D'autres types de pilotage sont à considérer mais sortent du périmètre de cet article, comme par exemple :

- Le contrôle des déplacements de la machine, avec notamment le comportement dynamique [Semm et al., 2019], qui dépend en effet du constructeur de la machine et non pas de son utilisateur.
- La correction des trajectoires 5 axes sous l'effet de la déformation [Dittrich et al., 2019b].

Ce dernier point peut être réalisé en on line [Lyu 2020] ou en off line. Duong [Duong et al., 2019] propose une approche intéressante : La formulation de l'erreur de contour proposée est basée sur l'estimation des points de contact de l'outil et l'OGA (offline gain adjustment) s'inspire de l'idée du contrôle prédictif du modèle (MPC).

3 TRADUCTION D'UN TOLERANCEMENT GPS EN PLAN DE SURVEILLANCE

Le tolérancement GPS n'est pas un tolérancement qui permet de piloter directement un processus de production. Le passage d'une spécification GPS vers un plan de surveillance demande une analyse de la part de l'industrialisation afin de le décomposer en spécifications pilotables et intrinsèques. Il convient alors d'identifier les différentes sources de dispersion d'une surface tolérancée en GPS et de les restituer dans le contexte des référentiels locaux déterminés par le système de référence GPS [Morse et al., 2018].

Analysons par exemple deux surfaces tolérancées par une localisation (Figure 1).

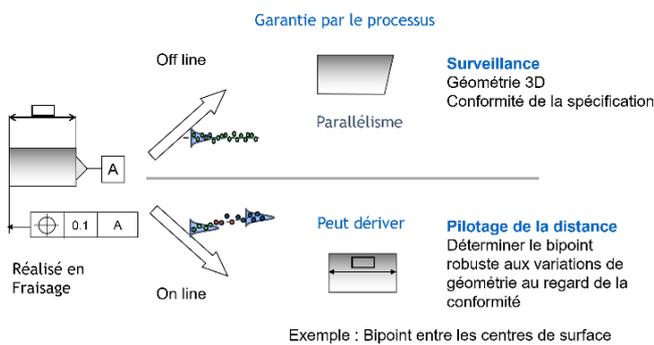


Figure 1 - Surveillance off-line et surveillance on-line

Les sources de dispersion sur cette surface sont, d'une part des variations de géométrie (parallélisme, planéité) et, d'autre part, des variations dimensionnelles pouvant provenir d'un mauvais réglage ou d'une usure d'outil. Les variations géométriques n'évolueront sans doute pas lors de la production, car elles sont données par la précision de la machine et les conditions de la gamme. On peut se contenter d'une validation à la conception/réception du processus et éventuellement d'un contrôle libératoire fréquentiel qui peuvent tout à fait être réalisés en off-line.

Par contre, les dimensions peuvent subir des dérives en cours de production et doivent donc être suivies par un contrôle fréquentiel réalisé en on-line pour garantir la conformité.

Dès lors plusieurs problèmes se posent :

- Comment identifier la parties pilotable et la partie intrinsèque d'une spécification ?
- Comment décomposer davantage ces parties en spécifications élémentaires ?
- Si on décompose la spécification, faut-il alors spécifier de nouvelles tolérances sur chacune de ces spécifications élémentaires ?

Dans l'exemple simple que nous avons présenté précédemment, nous avons pu identifier les caractéristiques géométriques comme étant la partie intrinsèque de la spécification, et les caractéristiques dimensionnelles comme étant la partie pilotable de la spécification.

Concernant la décomposition en spécifications élémentaires, une machine-outil utilise des correcteurs unidimensionnels. Il

convient donc de trouver des critères unidimensionnels à piloter. Dans notre premier exemple, le bipoint entre les centres des surfaces est un bon choix. Si ce bipoint est centré sur la longueur cible, alors il sera robuste à des variations géométriques comme le parallélisme.

Enfin, pour la question du tolérancement de ce bipoint, il est possible de fixer une tolérance nécessairement plus serrée que la localisation. Cependant, si on utilise une approche de pilotage de type APC [Muller et al., 2019], ou inertiel [Pillet et al. 2007] il n'est pas essentiel de fixer des tolérances. On cherchera à piloter cette longueur directement sur la cible.

Pour pouvoir garantir la conformité et, même davantage, la qualité d'une production, il faut donc pouvoir décomposer le processus en deux types d'éléments à contrôler :

- Ceux ne nécessitant pas de pilotage en temps réel mais une validation au démarrage et un contrôle de libération. Ils peuvent être surveillés en off-line.
- Ceux nécessitant une surveillance en temps réel pour un pilotage du processus. Ils seront surveillés en on-line.

Le contrôle de conformité qui concerne l'ensemble des spécifications géométriques ne doit pas être confondu avec le pilotage du processus.

4 DECOMPOSITION D'UNE SPECIFICATION GEOMETRIQUE

L'approche que nous avons présentée précédemment et consistant à distinguer ce qui relève d'un contrôle off-line et d'un contrôle on-line pilotable, est empirique. La décomposition de la spécification pour la partie pilotable consiste à définir les variations de la surface réelle selon les six degrés de libertés dans le repère de la machine.

Un tolérancement GPS localise l'intervalle de validité d'un élément dans l'espace par rapport à un repère. Illustrons cette décomposition sur un autre exemple (Figure 2).

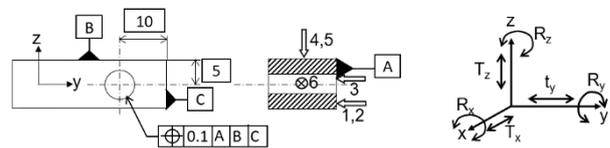


Figure 2 – Décomposition d'une spécification GPS

On décompose la tolérance de localisation selon les six degrés de liberté sur les axes X, Y et Z (Tableau 2). Pour chaque degré de liberté, on identifie la dérive potentielle et les possibilités de correction sur le moyen de production.

Tableau 2 – Décomposition de la tolérance géométrique

Degré de liberté	Dérive potentielle	Type	Occurrence	Action de correction en pilotage
Tx	Non concerné			
Rx	Non concerné			
Ty	Mauvais réglage	Décentrage	A chaque début	Correcteur TY, TZ
	Mauvais montage du foret	Dispersion	A chaque changement de foret	Changement Foret
Tz	Usure foret	Dispersion	Durée de vie > 150 pièces	Changement Foret

Ry	Posage non conforme	Décentrage	A chaque début	Néant
Rz	Pièce mal posée	Erratique	Jamais	Néant

On voit à l'aide de cette approche empirique que le pilotage continu de la qualité peut se faire en intervenant sur les deux correcteurs Tx et Ty.

Il faut donc identifier deux caractéristiques unidimensionnelles qui formeront la partie pilotable.

Dans ce cas, les deux bipoints candidats sont moins évidents à identifier que dans l'exemple précédent.

On pourra prendre par exemple pour ces deux bipoints :

- le point au centre du cylindre (pour être robuste aux rotation Ry et Rz) et le point projeté sur la surface de référence B, de longueur L1
- le point au centre du cylindre et le point projeté sur la surface de référence C, de longueur L2

Le plan de surveillance se déduit alors naturellement (Tableau 3). Ce plan est évidemment très différent d'une simple vérification au poste de la conformité de la spécification géométrique.

Tableau T3 – Plan de surveillance

Spec	Valeur	Moyen	Type de Surveillance	
			On-line	Off-line
Loc	0.1	Machine à Mesurer tridimensionnelle		Démarrage série Changement foret Fréquentiel off-line 3/Jour
L1	5			
L2	10		Départ série Freq 2/Heure	

Le contrôle de la localisation, qui est un contrôle de conformité, peut se faire en off-line. Les deux longueurs L1 et L2 étant impliquées dans le pilotage court terme de la machine nécessitent une surveillance accrue on-line.

5 UTILISATION DE LA DECOMPOSITION DANS LA BASE MODALE DES CORRECTEURS

Piloter une machine, c'est réduire la variabilité en agissant sur des correcteurs en nombre nécessairement limité. Comme il n'est pas possible d'annuler tous les écarts, on doit rechercher un critère d'optimisation des actions simultanées de tous les correcteurs. Il convient pour cela de déterminer les parties des déviations observées qui peuvent être compensées.

L'approche proposée est ici décrite à l'aide de deux cas : un cas simple de deux plans dont l'un est en référence et un cas plus complexe de surfaces gauches.

5.1 Cas 1 : localisation d'une surface

Dans le cas d'une localisation simple d'une surface (Figure 3), on a un vecteur d'écart de dimension égal au nombre de points palpés. Deux actions de pilotage sont possibles :

1. Une translation suivant l'axe Y de la machine
2. Une rotation autour de l'axe Z de la trajectoire de l'outil

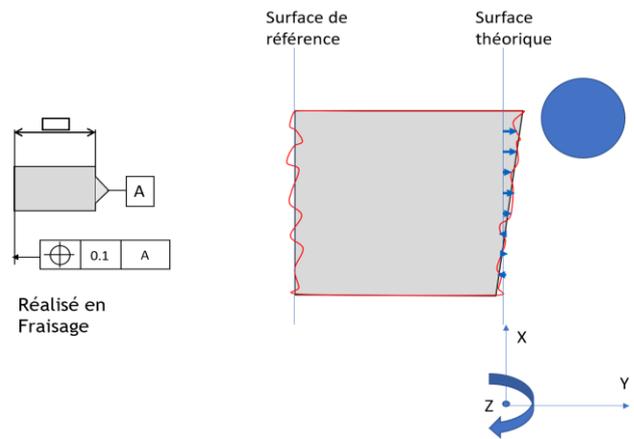


Figure 3 – Cas 1 : Spécification GPS et pilotage

L'objectif du pilotage est de trouver les deux actions simultanées qui réduisent le plus possible les écarts.

Le vecteur des écarts peut se décomposer dans une base modale à deux dimensions, chaque dimension étant représentée par l'action théorique d'une variation unitaire de la valeur d'un correcteur. Dans l'exemple présenté, le premier mode correspond à l'action du correcteur suivant l'axe Y et il est calculé comme le produit vectoriel entre cette action et la normale au point de mesure. Le second mode correspond à l'action de rotation de la trajectoire de l'outil et il est calculé par une approximation des petits déplacements.

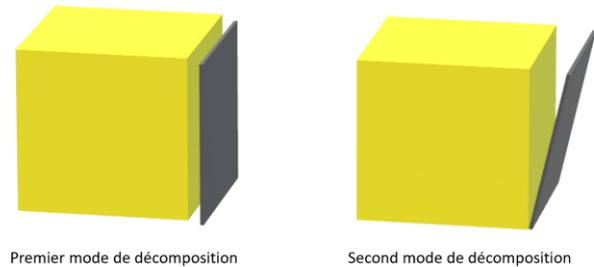


Figure 4 – Décomposition modale

Le vecteur d'écart peut alors être décomposé dans la base modale des correcteurs de la façon suivante :

$$V_{Mesure} = \sum_{i=1}^N \lambda_i Q_i + \varepsilon$$

Avec

V : le vecteur des écarts

N : le nombre de correcteurs disponibles

λ_i : Les coordonnées du vecteur V dans la base des modes Q_i

ε : le résidu de la projection

L'ensemble des coordonnées modales peuvent être regroupées sous forme d'un vecteur λ , obtenu via l'équation :

$$\lambda = Q^* \cdot V_{Mesure} = (Q^t \cdot Q)^{-1} \cdot Q^t \cdot V_{Mesure}$$

Dans l'exemple simple d'une seule surface, les valeurs du vecteur λ sont simplement égales aux corrections à appliquer pour minimiser les écarts résiduels.

5.2 Cas 2 : Localisation de plusieurs surfaces dépendantes

Dans ce cas (Figure 5), le vecteur d'écart V_{mesure} se décompose en deux sous-vecteurs d'écart V_B et V_C :

$$V_{mesure} = \begin{bmatrix} V_B \\ V_C \end{bmatrix}$$

Avec

V_B pour les écarts de la surface B en prenant pour référence la surface A

V_C pour les écarts de la surface C en prenant pour référence la surface B

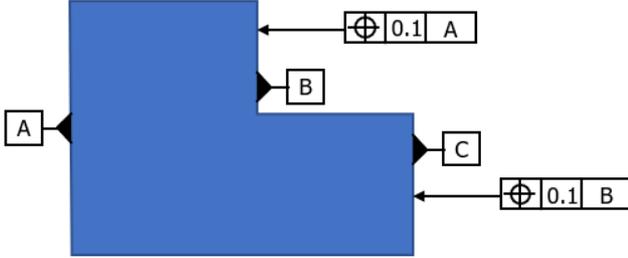


Figure 5 – Cas 2 : Plusieurs localisations dépendantes

A l'aide de la décomposition modale,

$$V_{mesure} = \sum_{i=1}^N \lambda_i Q_i + \varepsilon$$

On recherche les valeurs du vecteur λ qui minimiseront l'ensemble des écarts. Dans ce cas, la conformité en localisation sera optimale.

$$\lambda = Q^* \cdot V_{mesure} = (Q^t \cdot Q)^{-1} \cdot Q^t \cdot V_{mesure}$$

Pour cela, on doit identifier la matrice Q qui contient autant de colonnes que de modes élémentaires, chacun d'eux correspondant à une action de pilotage. Si pour chaque surface on dispose d'une action de translation et d'une action de rotation alors la matrice Q pourra s'exprimer sous la forme suivante :

$$Q = \begin{bmatrix} Q_{BB} & Q_{BC} \\ Q_{CB} & Q_{CC} \end{bmatrix}$$

avec

Q_{BB} qui représente l'effet sur la surface B des correcteurs agissant sur la surface B

Q_{CB} qui représente l'effet sur la surface C des correcteurs agissant sur la surface B

Q_{BC} qui est une matrice nulle dans ce cas

Q_{CC} qui représente l'effet sur la surface C des correcteurs agissant sur la surface C

5.3 Cas général 3D

La Figure 6 donne un exemple de spécifications 3D réalisées à partir de localisations dépendantes.

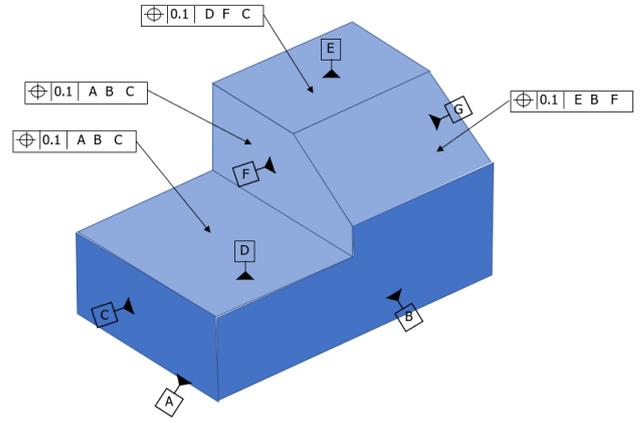


Figure 6 – Cas général : Spécification GPS 3D

Dans cet exemple, la matrice Q se décompose de la façon suivante :

$$Q = \begin{bmatrix} Q_{DD} & \cdots & Q_{DG} \\ \vdots & \ddots & \vdots \\ Q_{GD} & \cdots & Q_{GG} \end{bmatrix}$$

Avec

Q_{ij} qui représente l'effet sur la surface i des correcteurs agissant sur la surface j

La matrice Q représente la décomposition modale [Formosa et al., 2007] dans l'espace des correcteurs qui permet de trouver les réglages minimisant les écarts résiduels.

5.4 Hiérarchie des dépendances

Dans le tolérancement GPS, la décomposition modale implique le choix d'un ordre dans le référencement des surfaces. Par exemple, dans la Figure 7, la décomposition modale pour la surface C sera différente entre l'ordre de référencement AB et l'ordre de référencement BA.

Pour la première surface de référence choisie, il n'y a pas de précautions particulières à prendre, et on peut directement calculer l'effet des correcteurs sur les autres surfaces. En revanche, pour la surface suivante, il ne faut prendre que la partie de l'effet du correcteur sur les degrés de liberté bloqués par la seconde surface.

Pour cela, on calcule l'incidence projetée sur la direction du degré de liberté bloqué par la surface, que l'on projette à nouveau sur la normale à la surface.

On a donc :

Soit \vec{e} le déplacement du point p sous l'effet du correcteur

Soit \vec{n} la normale à la surface au point considéré

Soit \vec{u} la direction du degré de liberté bloqué par la surface

Soit \vec{v} la projection orthogonale de \vec{e} sur \vec{u}

$$\vec{v} = \frac{\vec{e} \cdot \vec{u}}{\|\vec{u}\|^2} \cdot \vec{u}$$

L'effet de la correction à prendre en compte est alors :

$$c = -\vec{v} \cdot \vec{n}$$

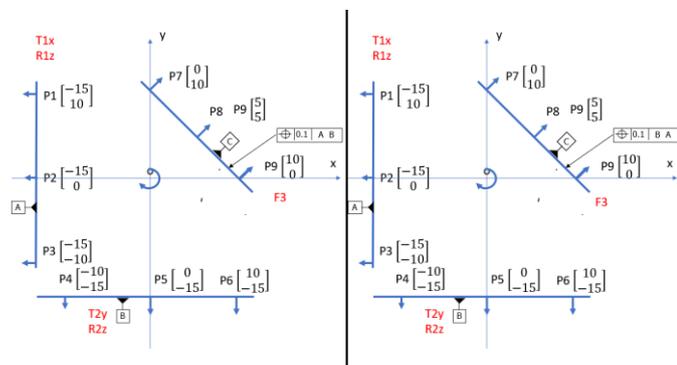


Figure 7 – Influence de l'ordre des surfaces

6 CONCLUSION

La normalisation GPS a incontestablement apporté une amélioration importante dans le tolérancement fonctionnel. Cependant, une dérive souvent observée dans l'entreprise est d'envoyer directement les plans en production sans effectuer une analyse au préalable de leurs implications en termes de pilotage des moyens de production.

Ce papier a décrit la démarche que nous proposons pour réaliser cette analyse, et cela de manière à aboutir à un plan de surveillance garantissant la conformité de la pièce. Une première approche consiste à séparer, de manière intuitive, ce qui relève d'un contrôle off-line et ce qui relève d'un contrôle on-line. Une seconde approche, plus générale, permet d'obtenir directement les réglages à partir d'un tolérancement GPS. Les exemples décrits ont illustré l'application de ces deux approches dans le cas d'une localisation, mais elles peuvent également être généralisées à d'autres spécifications.

7 REFERENCES

- Anselmetti, B. (2011). Manuel de tolérancement. *Lavoisier*.
- Dittrich, M. A., Denkena, B., Boujnah, H., & Uhlich, F. (2019a). Autonomous machining—recent advances in process planning and control. *Journal of Machine Engineering*, 19.
- Dittrich, M. A., Uhlich, F., & Denkena, B. (2019b). Self-optimizing tool path generation for 5-axis machining processes. *CIRP Journal of Manufacturing Science and Technology*, 24, 49-54.
- Duong, T. Q., Rodriguez-Ayerbe, P., Lavernhe, S., Tournier, C., & Dumur, D. (2019). Contour error pre-compensation for five-axis high speed machining: offline gain adjustment approach. *The International Journal of Advanced Manufacturing Technology*, 100(9), 3113-3125.
- Formosa F., Samper S., Perpoli I., *Modal expression of form defects*, In Models for Compute Aided Tolerancing in Design and Manufacturing, Springer Series, 2007, pp. 13-22
- Hong, Y., S., Chang, T., C. *A comprehensive review of tolerancing research*, International Journal of Production Research, Vol. 40, No. 11, pp 2425-2459, 2002
- ISO - ISO5458:1998 Geometrical Product Specifications (GPS) – 1998
- Lyu, D., Liu, Q., Liu, H., & Zhao, W. (2020). Dynamic error of CNC machine tools: a state-of-the-art review. *The International Journal of Advanced Manufacturing Technology*, 106(5), 1869-1891.

Möhring, H. C., Wiederkehr, P., Erkorkmaz, K., & Kakinuma, Y. (2020). Self-optimizing machining systems. *CIRP Annals*, 69(2), 740-763.

Morse, E. P., Shakarji, C. M., & Srinivasan, V. (2018). A brief analysis of recent ISO tolerancing standards and their potential impact on digitization of manufacturing. *Procedia CIRP*, 75, 11-18.

Muller, T., Pillet, M., Maire, J. L., & Pillet, D. (2019, March). Why and How to Move from SPC (Statistical Process Control) to APC (Automated Process Control). In *International Conference Design and Modeling of Mechanical Systems* (pp. 33-40). Springer, Cham.

Pillet, M., Adragna, P. A., Ozouf, V., & Guerra, A. S. (2007). Pilotage par carte de contrôle dans le cas d'un tolérancement inertiel. *Qualita 2007*.

Semm, T., Rebelein, C., & Zaeh, M. F. (2019). Prediction of the position dependent dynamic behavior of a machine tool considering local damping effects. *CIRP Journal of Manufacturing Science and Technology*, 27, 68-77.

Taguchi, G. (1987). *System of experimental design; engineering methods to optimize quality and minimize costs*. Vol 1 et 2, American Supplier Institute, INC

State Estimation considering the Sensor's Degradation

HASSAN HACHEM¹, HAI CANH VU², MITRA FOULARIRAD¹

¹ Université de technologie de Troyes
10300, France
hassan.hachem@utt.fr

² Université de technologie de Compiègne
60200, France
hai-canh.vu@utc.fr

¹ Université de technologie de Troyes
10300, France
mitra.fouladirad@utt.fr

Résumé – Dans cet article, nous nous intéressons à estimer la dégradation d'un système en tenant compte de la dégradation des capteurs. Sur cette base, nous avons besoin de deux processus stochastiques pour modéliser notre dégradation, un pour le système et l'autre pour le capteur. Concernant l'implémentation, nous avons utilisé deux algorithmes différents pour estimer la dégradation: le filtre particulaire (PF) et les méthodes de Monte-Carlo par chaînes de Markov (plus précisément Metropolis-Hastings (MH) est utilisé pour cette implémentation). Les performances de notre approche sont illustrées à travers des exemples numériques.

Abstract – In this paper, we are interested in the estimation of the degradation of a system taking into consideration the degradation of the sensors. Based on that, two stochastic processes are required to model the degradation, one for the system and the other for the sensors. Regarding the implementation, we used two different algorithms to estimate the degradation: Particle filter (PF) and Markov chain Monte Carlo (MCMC) (more precisely Metropolis-Hastings (MH) is used for this implementation). The performance of our approach is presented in simulated examples.

Mots clés - Dégradation du capteur, processus Gamma, processus Wiener, filtrage bayésien.

Keywords – Sensor degradation, Gamma process, Wiener process, Bayesian filtering.

1 INTRODUCTION

Modern products evolve from one generation to another. Automatically, this evolution generates a reduction in time-to-market, and a rise of products newly introduced to the market. In complex systems (modern products), we are not only projecting our focus on the system's failure but rather on information about how does the system deteriorates. Therefore, we can have more subtle information about degradation by gathering monitoring data from sensors. However, we will be facing a new problem, which is the possibility of having drifted data caused by the sensor's degradation and not only due to a random measurement noise. Meanwhile, product reliability has become an indispensable component of the consumer preferences. Higher standards of reliability coupled with lower costs have become a critical concern for companies to deliver competitive products. Methods such as condition monitoring and degradation analysis are developed to examine the reliability of modern products. Degradation analysis is shown to be a significant toolkit, particularly for those subject to restricted test time and sample size [Bae, 2007].

Predictive maintenance (PdM) has been recently developed and largely applied during the last decade. If the PdM is correctly implemented, it can help to avoid the potential risks of system failures and their consequences, to ensure the proper functioning of the system at the lowest cost possible. According to the PdM, maintenance decisions are made based

on the system Remaining Useful Life (RUL), regarding the estimation of the RUL one should propose a model to study the degradation. The RUL is usually estimated from the degradation data (vibration, temperature, crack size, etc.) collected from different sensors by using different prognostic approaches such as stochastic filters, Markov processes, machine learning. Hence, degradation processes have received considerable attention over the last decades. It has become essential to bring out the monitoring and prognosis of the degradation process to ensure the proper functioning of a system; this will avoid the potential risks of breakdowns and their consequences.

In the context of degradation modeling, stochastic process-based models are generally used [Guida, 2013]. Gamma and Wiener's processes are two of the most common classes of stochastic processes for degradation modeling. These two classes were well studied in the literature. The gamma process and its extensions in degradation modeling have been investigated in the works [Lawless, 2004]. The applications of the Wiener process and its extensions to the degradation modeling process have also been investigated in the works [Tang, 2008].

The consideration of degradation as a stochastic process is not recent. Several research studies have been carried out in this field, [Abdel-Hameed, 1975] was the first to implement the Gamma process as a model of random deterioration

phenomena. [Abdel-Hameed, 2010] and [Frangopol, 2004] present an overview of the most popular probabilistic models used to describe the system's degradation. [Noortwijk, 2007] presents the advantages of the gamma process, according to him this process is an effective tool if one wants to model a system's degradation with independent and positive increments; it has a simple, explicit mathematics calculation. [Noortwijk, 2009] presents a detailed study on the history, the definition, the calibration, and the different applications of the homogeneous and not homogeneous Gamma process. Unlike the Gamma process, which offer a monotonous and increasing path of evolution, the Wiener process can give a non-monotonic model of a random system. [Platen, 2006] indicates that this process is widely used as a mathematical object for modeling different random and strongly fluctuating random phenomena. Therefore, these two processes will be used to describe the system and the sensor degradation.

1.1 Related Work

In the literature, most of the models used to describe the deterioration of the system with the consideration of sensor degradation are linear additive models. The models can be divided into two categories: deterministic models and the probabilistic models.

In the deterministic models, there is three main articles. In [Ma, 2018]'s paper three sensor performance degradation models are considered in their study, namely, bias, drifting, and gain. All these models are described and studied in their article. According to [Wang, 2014] either the sensor works perfectly, or it fails. There is no other alternative considered in their work. In [Liu, 2016]'s paper, they suppose a sensor network consisting of n sensor nodes is exploited to measure the output of the target plant.

In the probabilistic models, one main article focuses on the study of the system with the consideration of sensor degradation which is [Liu, 2019]'s paper. They suppose a linear additive model with white noise. The Wiener process is employed to describe the system's degradation and the sensor's degradation. According to [Dorr, 1997] the impact of sensor degradation lies in the distortion of measurements, which makes the observations biased and severely deviate from the true value, which justifies the addition of white noise.

In this paper, we aim at estimating system and sensor degradation using Particle filter and Metropolis Hastings from a noisy measurement data for health monitoring and maintenance planning. Unlike the other system degradation studies, our proposed model considers the sensor degradation.

The paper is organized as follows. Section 2 describes the deterioration model, and introduces the stochastic processes used in this paper. The algorithms used to estimate the system and sensor degradation are given in Section 3. In Section 4, simulated examples are provided to show the performance and the flexibility of the method, and the final section presents our conclusion.

2 DETERIORATION MODELLING

2.1 Deterioration Model

Based on what was mentioned above, a probabilistic linear additive model is proposed, to study the system with the consideration of sensor degradation. Let $\{Z(t), t \geq 0\}$ indicate

the measurement process, which refers to the uncertain observation of the corresponding system and the deteriorating state of the sensor at time $t \geq 0$. Once the effect of the sensor and the degradation of the system are combined, the measurement at time t is given as:

$$Z(t) = f(X(t), Y(t)) + \epsilon = X(t) + Y(t) + \epsilon,$$

where $X(t)$ and $Y(t)$ are two random variables resulting from stochastic processes, denote the system's degradation level and the sensor's degradation level respectively, at time t . And ϵ is the measurement error, $\epsilon \sim N(\mu, \sigma^2)$. In the coming sections, several stochastic processes that describe the degradation of the system and the sensor (mainly Wiener process, Gamma process) will be taken into consideration. As mentioned in [Liu, 2019]'s paper, sensors can exhibit different degradation processes in different industries. The sensor can be subjected to various degradation processes, depending on the application and the environmental effects. Which justifies the choice of several stochastic processes.

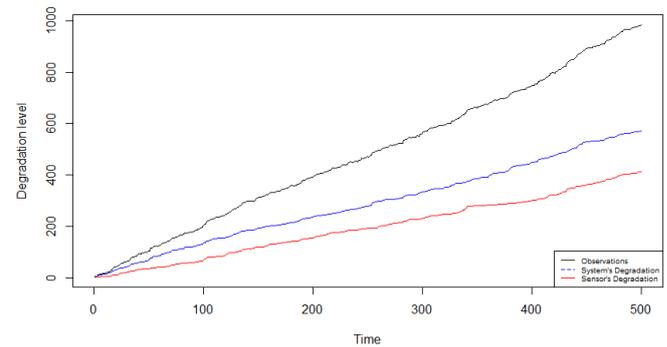


Figure 1. Representation of observations, system and sensor degradation

Figure 1 describes the measurements, system, and sensor degradation at the examination. If the sensor's degradation is considered, then the measurements at inspection differ from the system degradation as shown in the Figure above.

2.1.1 Gamma Process

In its simplest form a homogeneous Gamma process of a parameter (α, β) is defined as a stochastic process $\{G(t), t \geq 0\}$, having the following properties:

1. $G(0) = 0$.
2. The increments of this process are independent;
3. For every $0 < s < t$, the random variable of increment $G(t) - G(s)$ follows a Gamma distribution $Ga(\alpha(t-s), \beta)$.

2.1.2 Wiener Process

A Wiener process is a stochastic process $\{W(t), t \geq 0\}$ indexed by non-negative real numbers t with the following properties:

1. $W(0) = 0$.
2. Has independent increments: for every $t > 0$, the future increments $W(t+u) - W(t)$, $u \geq 0$, are independent of the past values $W(s)$, $t \geq s$.
3. W has Gaussian increments: $W(t+u) - W(t)$ is normally distributed with mean a and variance b^*u , $W(t+u) - W(t) \sim N(a, b^*u)$.
4. W has continuous paths: $W(t)$ is continuous in t .

For an efficient study of the degradation, it is necessary to estimate the non-observable state of the system. In this aim, the hidden degradation states vectors $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ should be evaluated from the observations $\mathbf{Z} = (Z_1, \dots, Z_n)$.

Due to the uncertainties associated to the vector \mathbf{X} and \mathbf{Y} to the observations \mathbf{Z} , proposing a relevant approximation of \mathbf{X} and \mathbf{Y} and being able to evaluate the estimation results, we should first measure the conditional density of the non-observable states X_1, \dots, X_n and Y_1, \dots, Y_n with respect to the observations Z_1, \dots, Z_n .

Due to the existence of a large number of integrals in the equation of the conditional density, the estimation of the conditional density is tricky and often impossible. To prevent this complicated estimation, the Particle Filter (PF) and the Metropolis-Hastings (MH) algorithms are proposed to estimate conditional densities.

3 ALGORITHMS

3.1 Particle Filter

Particle filter - also called bootstrap filter [Gordon, 1993], Monte Carlo filter [Kitagawa, 1996], or Condensation filter [MacCormick, 2000] - calculates sequentially $p(X_k|Z_{1:k})$ and $p(Y_k|Z_{1:k})$ from a given initial density. The aim of the particle filter is to approximate $p(X_k|Z_{1:k})$ and $p(Y_k|Z_{1:k})$ by a finite sets of particle $\{X_k^i\}_{i=1,\dots,N}$ and $\{Y_k^i\}_{i=1,\dots,N}$ with weights. The weight of each particle reflects the probability that the particle is at a relevant position of the state space. Sequential Monte Carlo methods allow the implementation of a Bayesian filter by numerical approximations.

To use the Particle filter first one needs to initialize the filter, our initialization is supposed as a random vector of size N , which follows a Gaussian distribution. After that, the known SIR algorithm will be used.

The SIR algorithms works as follows:

- Generate N particles using the prior distribution (i.e distribute them and give them a weight).
- Acquire the observations.
- Update the weight according to the likelihood.
- Normalize the weights.
- Resample particles set (use N_{eff} if necessary).
- Predict.

3.2 Metropolis Hastings

Metropolis Hastings [Hastings, 1970] starts by proposing a move from current state, with a conditional probability density, then compute the Hastings ratio is computed, after that the candidate is accepted with probability $a = \min(1, r)$, where r is the Hastings ratio. That is, the post-update state is the proposed state with probability a , and the post-update state is the current state with probability $1-a$. The final step, which is called «Metropolis rejection», the candidate, is accepted or rejected according to a uniform random number on $[0, 1]$.

4 SIMULATION RESULTS

As mentioned in section 2.1, different cases will be proposed for our degradation model:

1. Two Wiener processes that describe the system's degradation and the sensor's degradation plus additive white noise.
2. Two Gamma processes, which describes the system's degradation and the sensor's degradation plus additive white noise.

3. Finally, the system's degradation is supposed to be a Gamma process and the sensor's degradation is a Wiener process plus an additive white noise. Then we will compare the performance of our two algorithms.

4.1 First Case

We suppose that X and Y are two Wiener processes defined by these parameters: $(\mu = 4, \sigma = 1)$ and $(\mu = 6, \sigma = 1)$ respectively, and $\epsilon \sim N(0, 1)$.

For MH, we start with a random starting value, and we run our algorithm for 10000 iterations.

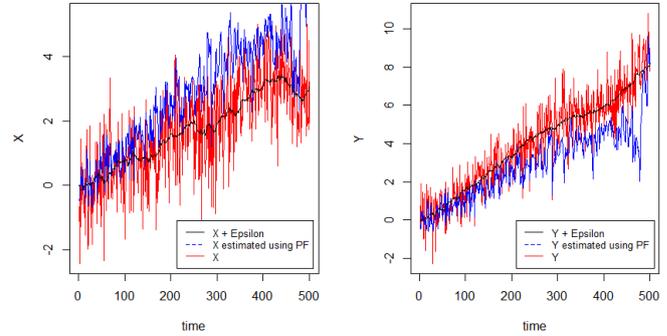


Figure 2. Estimation of X and Y using PF.

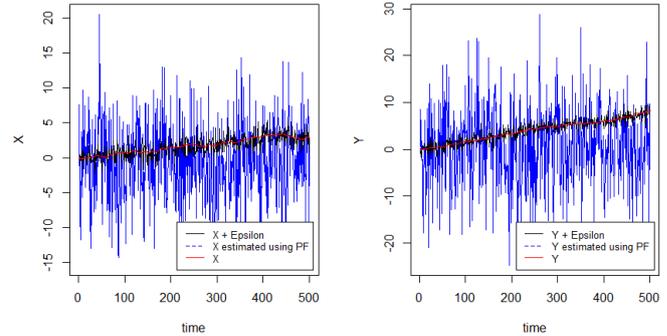


Figure 3. Estimation of X and Y using MH.

From Figures 2 and 3, we can see that the estimation of X and Y are close to the true state, which means that our algorithms work well for this case.

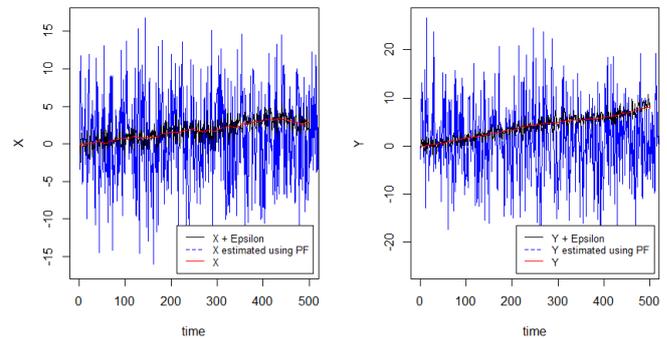


Figure 4. X and Y estimation using MH with a different starting value.

We tried the algorithm with different starting values; none of the starting values affected the results (see Figures 3 and 4). To compare the methods, we make use of the MSE and RMSE.

Table 1. MSE and RMSE of the PF and MH prediction

Vector	PF		MH	
	MSE	RMSE	MSE	RMSE
X	0.509	0.714	0.9071	3.406
Y	0.665	0.815	0.952	1.846

By comparing the estimated values (X estimated and Y estimated) and the real values (X and Y) from Table 1. We conclude that in both cases, the estimation of X (system's degradation) is better than the estimation of Y (sensor's degradation). In this case, PF works better than MH.

4.2 Second Case

We suppose that X and Y are two Gamma processes defined by these parameters: $(\alpha = 2, \beta = 1)$ and $(\alpha = 1, \beta = 1)$ respectively, and $\epsilon \sim N(0, 1)$.

For MH, we start with a random starting value, and we run our algorithm for 10000 iterations.

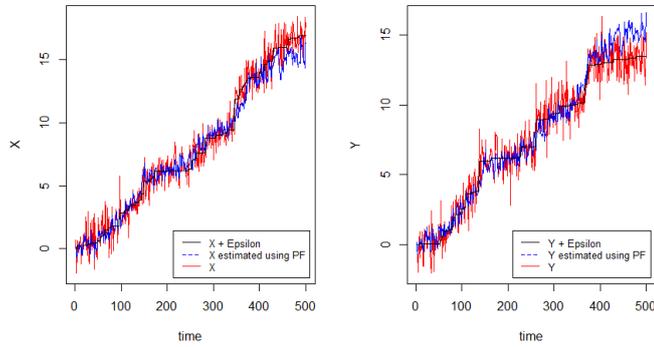


Figure 5. Estimation of X and Y using PF.

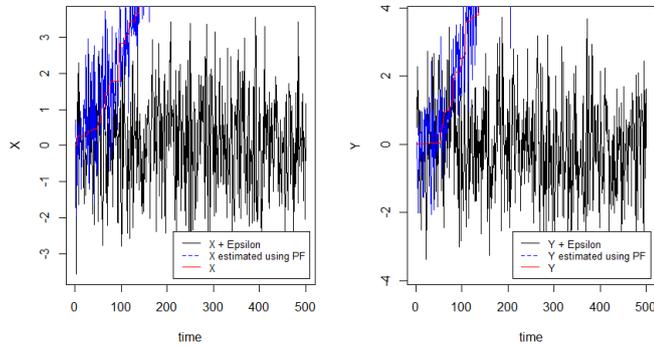


Figure 6. Estimation of X and Y using MH.

As shown in Figure 5, the estimation of X and Y works well using PF. However, it does not work well with MH (see Figure 6) with a random starting value. As we can see, the first estimated values are close to the true states, which means that our algorithm works well but it needs more time to achieve the convergence. Hence, to gain time, we will suppose that we know our starting values, which are the true X and the true Y, we obtain the following result:

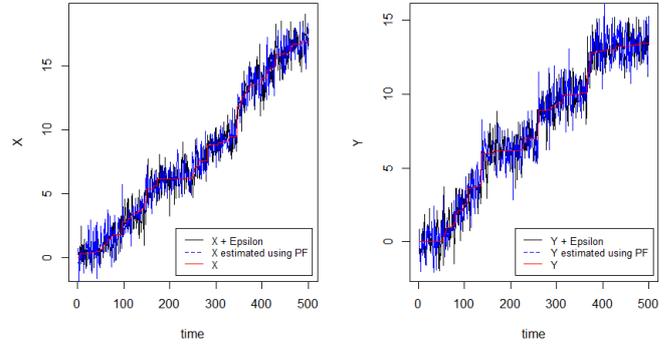


Figure 7. X and Y estimation using MH with X and Y as starting values.

Just by changing our starting values, we obtain good results (see Figure 7). This is because, we supposed our proposal follows a Gaussian distribution, which means our proposal vectors will always be around the mean of our proposal, while Gamma process is a monotonous process, hence we can suppose our starting value as a regression linear (such that: $X = ax + b$).

Table 2. MSE and RMSE of the PF and MH prediction

Vector	PF		MH			
	MSE	RMSE	MSE		RMSE	
Vector	MSE	RMSE	First case	Second case	First case	Second case
X	0.691	0.831	0.099	91.98	0.101	9.59
Y	0.951	0.975	0.314	77.16	0.318	8.78

By comparing the estimated values (X estimated and Y estimated) and the real values (X and Y) in Table 2. We observe that the estimation of X is better than the estimation of Y, except for the case where MH did not achieve the convergence. From our point of view, we think PF works better than MH because it is not affected by a starting value, hence we would recommend using PF in this case to estimate the state of X and Y.

4.3 Third Case:

We suppose that X is a Wiener process defined by these parameters $(\mu = 6, \sigma = 1)$, and Y is Gamma process defined by these parameters $(\alpha = 2, \beta = 1)$.

For MH, we start with a random starting value for X. To avoid the problem of non-convergence that we have seen in Section 4.2, since Y is a Gamma process, we suppose our starting value is the true state Y otherwise the result will not be accurate (see figure 8), and we run our algorithm for 10000 iterations.

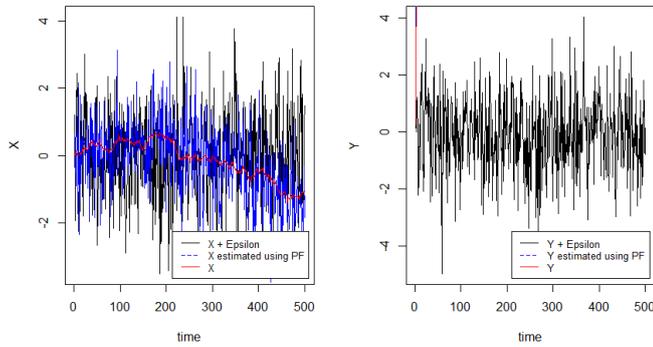


Figure 8. Estimation of X and Y using MH.

Figure 8 shows that MH does not work well with a random starting value as mentioned in the previous case.

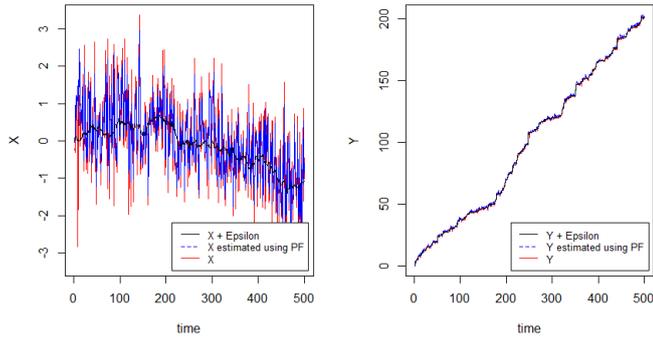


Figure 9. X and Y estimation using PF.

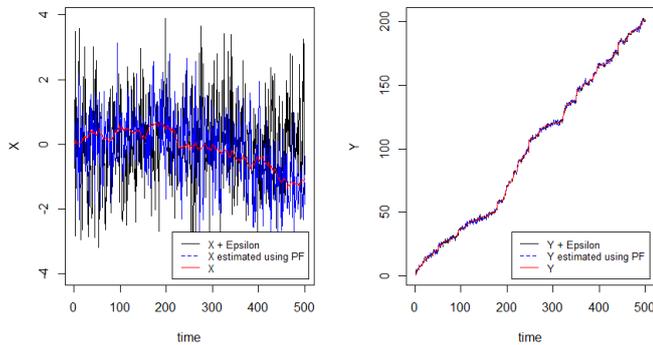


Figure 10. X and Y estimation using MH with Y as a starting value.

Figures 9 and 10 show that both of the considered algorithms work well in this case (if we suppose Y as a starting value for MH).

Table 3. MSE and RMSE of the PF and MH prediction

Vector	PF		MH	
	MSE	RMSE	MSE	RMSE
X	0.565	0.752	2.107	1.451
Y	43.58	6.6	0.848	0.92

Table 3 gives us the MSE and the RMSE of the estimated values (X estimated and Y estimated) and the true values (X and Y) of the two considered algorithms. Therefore, if one wants to estimate the system's degradation only, we can use

both algorithms. Nevertheless, if we want to estimate both the system and the sensor degradation, we recommend using the PF, because it does not require any special condition.

5 CONCLUSION

The two considered algorithms, PF and MH, for system and sensor degradation, estimates the states of both degradations. The PF and MH are used to deal with the degradation and measurement uncertainty. Our simulation studies show that PF works well for all our proposed cases, because it does not require any special condition. However, MH works well for Wiener processes. While it requires a special starting value to work well for Gamma processes.

From future perspectives, we envisage to extending this work by implementing the two considered algorithms while using different parameters. We also envisage, extending this work to predict the system and sensor degradation. And to estimate the RUL using PF and MH, for one of these cases and propose a maintenance policy.

6 REFERENCES

- Abdel-Hameed, M. (1975). A gamma wear process. *IEEE transactions on Reliability*, 152-153.
- Abdel-Hameed, M. (2010). Degradation processes: An overview. *Advances in Degradation Modeling*, 17-25.
- Dorr, R. a.-L. (1997). Detection, isolation, and identification of sensor faults in nuclear power plants. *IEEE Transactions on Control Systems Technology*, 42-60.
- Frangopol, D. M.-J. (2004). Probabilistic models for life-cycle performance of deteriorating structures: review and future directions. *Progress in structural engineering and Materials*, 197-212.
- Gordon, N. J. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE proceedings F (radar and signal processing)*, 107-113.
- Guida, M. a. (2013). The inverse Gamma process: A family of continuous stochastic models for describing state-dependent deterioration phenomena. *Reliability Engineering & System Safety*, 72-79.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Oxford University Press*.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of computational and graphical statistics*, 1-25.
- Lawless, J. a. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime data analysis*, 213-227.
- Liu, B. a. (2019). Stochastic filtering approach for condition-based maintenance considering sensor degradation. *IEEE Transactions on Automation Science and Engineering*, 177-190.
- Liu, Q. a. (2016). A resilient approach to distributed filter design for time-varying systems under stochastic nonlinearities and sensor degradation. *IEEE Transactions on signal processing*, 1300-1309.
- Ma, S. a. (2018). Structural response recovery based on improved multi-scale principal component analysis considering sensor performance degradation. *Advances in Structural Engineering*, 241-255.
- MacCormick, J. a. (2000). A probabilistic exclusion principle for tracking multiple objects. *International Journal of Computer Vision*, 57-71.

- Platen, E. (2006). A benchmark approach to finance. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 131-151.
- Suk Joo Bae, W. K. (2007). Degradation models and implied lifetime distributions. *Reliability Engineering & System Safety*, 601-608.
- Tang, J. a.-S. (2008). Estimating failure time distribution and its parameters based on intermediate data from a Wiener degradation model. *Naval Research Logistics (NRL)*, 265-276.
- van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 2-21.
- van Noortwijk, J. M.-J. (2007). Gamma processes and peaks-over-threshold distributions for time-dependent reliability. *Reliability Engineering & System Safety*, 1651-1658.
- Wang, X. a. (2014). Stochastically resilient extended Kalman filtering for discrete-time nonlinear systems with sensor failures. *International Journal of Systems Science*, 1393-1401.

Combining First Principles and Machine Learning for optimal maintenance of electrical assets

MARCEL CHEVALIER^{1*}, SYLVAIN MARIÉ^{1*}, BARTOSZ BOGUSLAWSKI¹, MICHEL CERCUEIL²,
FRANÇOIS CHUPOT², ANTOINE VIGNON¹, WEDIAN YOUSSEF²

¹ Schneider Electric Industries,
160, avenue des Martyrs
38000 Grenoble, France

² Schneider Electric Industries, France

*Corresponding authors : marcel.chevalier@se.com, sylvain.marie@se.com

Résumé – Les acteurs industriels sont confrontés à un besoin accru de réduire le risque d’incidents imprévus et d’optimiser leur maintenance. Ces besoins appellent à une réponse double à la fois sur les détections court-terme et la surveillance de vieillissement long terme. Chez Schneider Electric nous répondons à ces deux problématiques en combinant l’apprentissage artificiel et les modèles physiques et statistiques, afin de fournir des services avancés pour les équipements de distribution électrique. L’apprentissage artificiel est utilisé pour modéliser les paramètres clés des équipements électriques en fonctionnement normal, afin de prédire le comportement normal attendu en toute circonstance. Un système de surveillance peut se servir de cette référence pour détecter des déviations statistiquement significatives, révélant un défaut potentiel. En ce qui concerne les modèles de vieillissement, ils sont basés sur des modèles physiques et statistiques, dans la continuité de l’état de l’art du Prognostic Health Management. Il s’agit ici de décomposer les équipements électriques en sous-systèmes, et d’identifier pour chacun d’eux l’ensemble des modes de dégradation concurrents et leurs facteurs d’influence (environnement et usage). En combinant cela avec les cinétiques appropriées nous pouvons estimer la fin de vie. Ces deux approches, court-terme et long-terme, sont combinées sur le terrain pour fournir des services de maintenance prédictive. Ils s’inscrivent dans une initiative plus large visant à partager des algorithmes avec l’écosystème de partenaires de Schneider Electric [SE, 2019-1].

Abstract – Industrial end users face an increasing need to reduce the risk of unexpected failures and optimize their maintenance. This calls for both short-term analysis and long-term ageing anticipation. At Schneider Electric we tackle those two issues using both Machine Learning and First Principles models, to provide advanced services for electrical distribution assets. Machine Learning algorithms are used to model key parameters of electrical assets in normal operation. They predict the expected normal behavior, which serves as a reference. The monitoring system can therefore detect statistically significant deviations from this reference, indicating potential failures. Concerning ageing models, they are based on First Principles, similar to Prognostic Health Management techniques. It consists in breaking down electrical assets into sub-assemblies, identifying for each of them all possible competing degradation modes and their influencing factors (environment and usage). Combining this with relevant kinetics laws we can estimate the expected end of life. These two approaches, short-term and long-term, are combined in operation to provide predictive maintenance services. They belong to a broader initiative aimed at sharing analytics beyond the Schneider Electric scope [SE, 2019a].

Mots clés – Pronostic, Apprentissage Artificiel, Modèles de dégradation, Vieillissement, Détection d’anomalies.

Keywords – Prognostic Health Management, Machine Learning, Degradation models, Ageing, Anomaly detection.

1 INTRODUCTION

Nowadays digitization and Industrial Internet of Things (IIoT) make it extremely easy to collect a vast amount of data concerning electrical assets, during their operational life in real time conditions on customer plants. It allows users to get information on all environmental and usage conditions for products in real situation, together with data on observed failures.

Collected data can be used to learn normal behavior models of assets. Such models are relevant to detect anomalies characterized by a statistically significant deviation from the normal situation. While this kind of anomaly detection method is suited for short term asset monitoring, it is not appropriate

for long term degradation trends. To capture such trends, we can use First Principles degradation models (i.e. physical degradation models such as Arrhenius law); they complete the monitoring system to get a global view on asset health.

In the following sections we will first detail the short-term strategies (section 2) then the long-term ones (section 3) and conclude (section 4).

2 SHORT-TERM ANOMALY DETECTION

2.1 Context and motivation

Electrical assets may experience issues under some conditions. One of the most dramatic consequence to avoid is fire, which may impact both humans and equipment, in addition to

stopping customer business / process. Electrical fires cause an estimated \$3 billion-plus of damage to non-residential structures in the United States and Europe every year according to the National Fire Protection Association (NFPA) and the European Fire Safety Alliance (EFSA).

Designing equipment and switchboards in accordance with IEC 61439 standards ensures a proper IP level and significantly reduces the risk of fire due to insulation failure or excessive temperature rise. Then, one of the leading causes of electrical fires in medium- and low-voltage installations is faulty power connections of cables, busbars, busbars, and circuit breakers, particularly when the connections are completed onsite [SE, 2019b].

Cable, busbar, and withdrawable circuit-breaker connections can start to deteriorate due to loose connections caused by improper tightening torque or constant vibrations over time. Deterioration can also occur because of damaged surfaces due to corrosion, excessive pressure, or excessive friction. [Hirschbold and Chabert, 2018] provide a review of these cases and their causes.

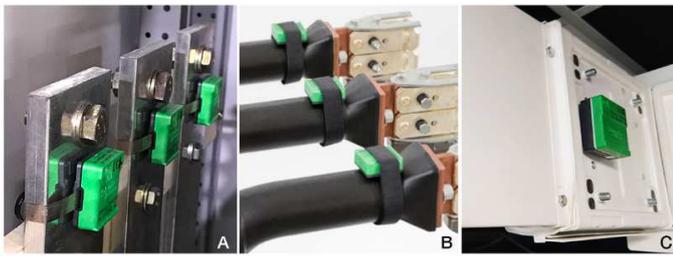


Figure 1. Wireless temperature sensors (Easergy TH110)

To prevent dramatic events to happen and detect such issues early, thermal surveys are typically performed using infrared (IR) thermography technology. However, these are done on a scheduled basis and may miss critical conditions occurring between scheduled scans. Continuous thermal monitoring overcomes these limitations by connecting wireless thermal sensors (Figure 1) to a data collection and monitoring infrastructure (Figure 2) [Hirschbold and Chabert, 2018]. [Hoffman et al., 2020] provide a review of alternate sensing technologies.

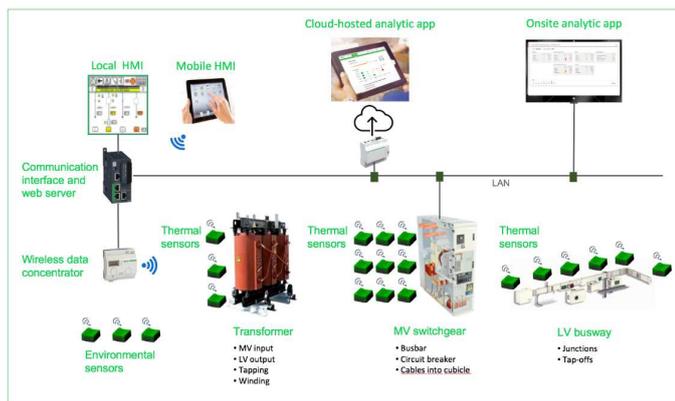


Figure 2. Continuous thermal monitoring architecture

Although such continuously collected temperature data provides a way to detect critically high temperature levels, typically using a “high temperature” threshold, resulting alerts might happen too late (reactive maintenance) as the threshold has to be set high enough to not trigger annoying false positives. This calls for a more adaptive temperature threshold,

based on context information: when the asset is not much loaded and external temperature is low, the threshold should ideally be lower than when the loads draw a large amount of current and external weather is already very warm. In the following section we show how such context-adaptive thresholds can be created with virtual sensors, so as to detect deviations early, before they transform into an actual fault (predictive maintenance).

2.2 Methodology

2.2.1 Soft / virtual sensors

Since 2012, Schneider Electric has developed a program to predict various industrial data based on usage and environmental conditions. Such idea of “Soft sensor”, also known as “Virtual sensor”, is not new and has been described in many ways in the literature and implemented in industrial products. For example, in the control theory domain, so-called “observer models” are fit to predict the discrete or continuous value of a variable. Popular models include white box detailed physical models, gray box simplified physical models, and black box data-driven models. The last two categories have their equivalent in, or reuse techniques from, the statistics and machine learning domains, learning models from appropriately pre-processed data. See [Curreri, 2020] for a review.

Gray-box models can be derived from the simplified physics of the asset or system at hand, for example with reduced order models and thermal-electrical equivalent models. Thermal radiation, conduction and convection are the main phenomenon at play when it comes to Temperature. Heat is mostly generated from drawn current through the Joule effect $P \sim I^2R$, as well as from magnetic induction.

For Machine Learning-based virtual sensors, popular pre-processing methods include *sliding window* to allow the model to pick up numerical delays and n-th order numerical derivative ; *(temporal) feature extraction* including rolling averages and other statistics, Fast Fourier Transform (FFT), Wavelets, etc. ; principal component analysis (PCA) and singular value decomposition (SVD) to compress and optimize input data so that models learn in a more robust way, etc. See [Curreri, 2020] and [Hoffman et al., 2020] for a review.

A virtual sensor can predict a category (classification) or a number (regression). Popular models include Linear and Logistic Regression, Kalman Filters (KF), Particle filters, Hidden Markov Models (HMM) and more generally Dynamic Bayesian networks (DBN), multi-layer Neural networks (NN), Recurrent Neural Networks (RNN), Support Vector Machines (SVM), Relevance Vector Machines (RVM), Decision Trees and Regression Trees (DT, M5’), Random Forest and Xtreme Gradient Boosted Trees (XGBoost). See [Souza, 2016] for a review of the most popular techniques.

Generalization performance of the models can be estimated using split- or k-fold cross-validation, as per the Machine Learning state of the art. Usual performance metrics include the Root Mean Square Error (RMSE) and the R-squared (R^2).

2.2.2 Adaptive threshold: significant drift detection

Besides their use for better process control, virtual sensors are used to tackle many other problems, such as back-up of a real sensor, what-if analysis, sensor validation and fault diagnosis. The latter consists typically in monitoring a significant

deviation between the actual monitored data and the reference provided by the soft sensor.

A usual way to perform this step is to model the prediction error (so-called residuals) and use this model to detect a statistical outlier. Such models can be simple and straightforward (e.g. normal distribution, possibly leveraging generalization error statistics, see previous section); they can otherwise be integrated in the virtual sensor model itself (e.g. quantile regression forests [Meinshausen, 2006]). Fault detection *per se* can be as simple as bayesian inference or rely on the temporal nature of the data at hand to tackle both single-point deviations and consistent multi-sample deviations. Techniques from the field of Statistical Process Control (SPC) such as Control Charts, Cusum or EWMA can be relevant for this aim.

Figure 3 illustrates how such technology can be assembled to create an adaptive temperature monitoring solution and solve the problem described previously in Section 2.1.

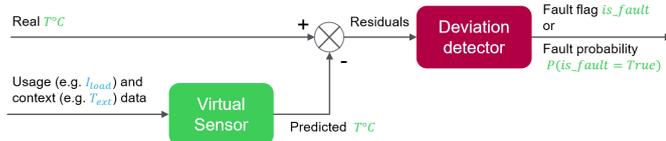


Figure 3. Virtual Sensor Fault Detection principle

2.2.3 Industrialization and deployment on the field

Our virtual sensor and virtual sensor fault detection analytics components have been industrialized at Schneider Electric in 2013 and 2016 respectively in the form of web service APIs on the cloud, so as to be reused in the various offers in the Schneider Electric catalog. Both are complementing the existing Avantis PRISM offer that was already present in the company's portfolio fault detection

The virtual sensor component was tested successfully onsite at a food and beverage industrial customer site as well as on several customer datasets, for example water pump networks. The virtual sensor fault detection approach was applied on a range of applications since 2016 including dry transformer overheating, solar panels soiling detection, LV or MV switchboards thermal monitoring, Air Handling Units (AHU) fault detection [Gao et al., 2019], refrigeration cabinet temperature sensor backup, and more. It was integrated within the Conext Advisor 2 and Asset Advisor software offers as part of the Schneider Ecostruxure suite.

Since 2019 this technology has been transferred into the Schneider Exchange platform [SE, 2019a] in the form of pay-per-use Analytics APIs, for all actors in the energy and automation ecosystem to leverage into their own business.

3 LONG-TERM AGEING MODELS

Since 2015, Schneider Electric has developed a program to develop algorithms to compute the ageing of assets with respect to time, environmental conditions, and usage conditions. The idea here is to offer new services to customers, based on advanced preventive maintenance and Reliability-Centered Maintenance features; these new algorithms may help developing time-based maintenance, spare part

management, condition-based maintenance, and predictive maintenance when used in simulation on what-if scenarios.

To do this, we must ensure that we can quantify usage and environmental conditions, and that we master the ageing laws with respect to these conditions. This is now possible thanks to digitization: we can rely on lots of data coming from assets in operation, providing information not only on their operational state, but also on their environment and usage, which can help define and tune ageing models, when compiled.

Rather than trying to define an approximative mission profile for our assets on customer plants, we can now have a clear view of the real environmental conditions (temperature, hygrometry, degree of salty atmosphere, concentration of corrosive gases, etc.) and of the real usage conditions (intensity, voltage, power, reactive power, harmonics, number of trips, etc.).

In addition, we have built a methodology that enables us to enrich the traditional physical ageing laws with a statistical approach that will be fed by data coming from digitization.

In the following sections, we will develop and explain this approach to determine a realistic value for the age of assets, leading to new services offers.

3.1 Ageing models history and limits

Since the early 1900's, many ageing models have been developed in the specialized literature [Pagès]:

- endurance models for mechanical devices;
- thermal ageing for electronic devices;
- resistance/constraint models for architecture;
- corrosion models for different metals;
- etc.

One common limitation of these models is that they rarely consider multiple environmental conditions, but more frequently they concentrate on a limited number of influencing factors (only the temperature for Arrhenius models, as an example).

Another limitation of these models is that they are rarely considered together, but are more often dedicated to a given situation, such as the corrosive atmosphere for metallic stick models for example.

What we are currently trying to do in Schneider Electric can be illustrated by what happens on our own car:

- it includes many different parts, with different ageing models (brakes do not age the same way as tires);
- it is sensitive to environmental conditions: the ageing of your car will not be the same, whether you live in Paris, Stockholm or Riyadh;
- it is sensitive to usage conditions: the ageing of your car will not evolve the same way if you drive very slowly and carefully and park your car in your garage each evening, or if you drive very nervously, with frequent harsh engine solicitations.

Anyway, when you drive your car, it may send you some visual warnings, suggesting you have a look on the motor, the oil level, or the air pressure in the tires. These warnings happen before the failure occurs, since they are based on the evolution of characteristics that reflect the health status of certain parts of the car. And all this information will be computed considering the environmental and usage conditions, then combined to provide valuable information to the driver.

For electrical devices such as a Low Voltage Circuit Breaker (LVCB), things are more complicated, because if one can rather easily measure a pression in the tires, the level of gasoline in a tank, or a rotation speed on a motor, it is much

more complicated to measure something on a device whose main mission is:

- a) to stay motionless 99.99% of the time;
- b) to react imperatively when solicited (since circuit breakers are safety devices).

Since we can't measure internal physical modifications in the breaker to infer a degradation, we must rely on models based on external environmental and usage conditions.

3.2 New possibilities induced by digitization

As explained in previous section, digitization allows now to know more about the way customers use their assets: with adequate sensors, we can monitor the *environmental conditions* (temperature, hygrometry, corrosive gases level, salt concentration in atmosphere, etc.) and the *usage conditions* (number of operations or trips, variation of intensity, harmonics and voltage impacting the asset, etc.). The sampling rate should be selected appropriately to provide a precise view of the asset's real life.

In parallel, thanks to internal failure event loggers we can also monitor the different failures the asset experiments, for each part of the asset: mechanics, electronics or electrotechnics.

Putting together this information will allow:

- to feed the ageing models with relevant information concerning the asset environment and the constraints it would be submitted to;
- to refine the ageing models when realizing that the predictions are either too optimistic or too pessimistic, considering precisely the given type of asset, rather than a wide family in an asset range. This improvement will be an evolutive process enabled by a methodology mixing physics-based and statistical approaches: we start from existing knowledge delivered by our experts and improve this knowledge by leveraging data coming from customers, from maintenance services, and from labs. In addition, Machine Learning capabilities may be used to process the data from multiple sources.

Relying on ageing and probability of failure information, we can then provide customers new services related to their connected assets, which will free them from performing tasks they are usually not that fond of managing: this may be connected services that deliver information related to risk of failure and need for maintenance, or this may be maintenance intervention initiatives on the assets, even before the customer has been warned that something went wrong.

Obviously, these services will be available only for the customers who will accept to share the connected data of their assets. In a broader way, one could think about extending the cooperation between end users in a dedicated segment, to ease the improvement of ageing models. This is already ongoing in the Oil & Gas domain: within the so-called OREDA initiative (Off-shore REliability DAta), great petroleum companies [Sintef] share information about observed failures, increasing the accuracy of corresponding predictive models.

This approach could be extended to other domains. End users in a dedicated segment could arrange agreements to share data on failures which occurred in their respective plants, giving way for a high improvement in ageing models, with the help of manufacturers, specialists of their assets. Of course, this will take time, but the digitization trend is already there, enabling this perspective.

Schneider Electric has already started to build optimal ageing models, and to value them for providing new services to its customers. In the next section, we introduce the methodology used to describe the models, the high-level principles to improve them (relying on data), the way they are used, and finally, examples of new services for maintenance relying in ageing information.

3.3 Methodology

The methodology we have developed has been patented in 2018 [Patent]. It mixes a physical approach based on First Principles, providing information about the level of degradation for each asset part, then we transform this information into failure rate values, relying on statistical models.

As stated in [Chevalier-Boutin], we differentiate between:

- a) degradation modes, which we link with the physical evolution of the asset with respect to time, environment factors and usage conditions (but which do not lead immediately to a failure),
 - b) failure modes, which characterizes an asset failure, and may result as the outcome of one or several degradation mode(s).
- In practice, for each type of assets, we defined six steps, which read as follows.

1. Identify sub-assemblies. These sub-assemblies are elements that follow a dedicated and common life cycle (i.e. they can be replaced independently from the other parts) or are optional. They may or may not match the functional sub-assemblies. For a car, it may be the tires, consistent parts of the engine, or mechanics;
2. List degradation modes. With this list, we will also list the influencing factors for the evolution of each degradation mode. For the engine of a car, it may be the wear and the corrosion, and the influencing factors the driven distance, the temperature and hygrometry, and the average speed selected by usual driver;
3. Match degradation modes and failure modes. When a given sub-assembly is entirely degraded with respect to one of its degradation modes, identify which outcome it will have on the asset itself, as perceived by the customer. For a car tire, the outcome of wear may be an explosion, a puncture, or some adhesion loss;
4. Set failure prioritization. When we only have a global failure rate collected on field data, allocate the frequency of appearance for each degradation mode of each sub-assembly, according to asset technical experts. This step can also be fed by figures provided by the results of dedicated tests, or accurate feedbacks from customers. Obviously, with data coming from the field, these figures are likely to be revised on a regular basis;
5. Reallocate failure rates on degradation modes. Using the results of previous step on a reverse way, specifically when we only have one figure for the global failure rate of the asset, we can reversely allocate figures for each degradation mode of each sub-assembly, the sum of which matching the 100% figure representing the observed global asset failure rate. If we have more detailed figures on failure rates, the method applies even more easily. A figure for

global failure rate can be broken down for each failure mode, and then for each degradation mode. For a car motor, if we have only a global failure rate, we will allocate it on the different degradation modes according to the experiments or experts' opinion;

6. Attach a kinetic law for each degradation mode. Here, and using the influencing factors defined in the step N° 2, we define the way the subassembly ages, such as Arrhenius law for temperature, or Coffin-Manson model for temperature and humidity, or other specific models extracted from technical literature, experimental works, or standards. For specific parts, ageing factors are already listed in dedicated standards, which have been re-used here.

3.4 Model's improvement relying on data

Thanks to digitization and the data it provides, we will be able to use them to improve the forecasting of ageing assets in four directions:

- 1) Confirm the ageing factors corresponding to the different levels of degradation, both from environment and usage conditions;
- 2) Refine the breakdown of final degradation outcomes with respect to the different failure modes;
- 3) Improve accuracy of the values for final failure rates, as observed in the real conditions of the customers;
- 4) Precise the mission profile for each customer segment.

To do this, we can also rely on different sources of data:

- a) From the field returns (customer sites), which will provide both information on the failures, but also on the different factors used to predict them (related to environment and usage);
- b) From the minutes of maintenance operations performed by the company's maintenance services on customers' plants;
- c) From the failures collected on connected products, thanks to digitization;
- d) From the results of tests in manufacturer's laboratories.

Additionally, machine Learning techniques are also available, which could help extracting valuable information from huge masses of data.

Validating the models require simultaneously good knowledge of physics underlying the degradation processes, expert skills to criticize the forecasting induced by the models when applied on "what-if" scenarios, and statistical methods to compute an average failure rate from known returned field products. These three approaches are combined in that initiative, which relies on the good cooperation of people working in different parts of the company organization.

Obviously, the operating phase of this project will provide new field data, which will be used to tune up the ageing models, ensuring the best fit between predicted and field figures.

3.5 How the models are used

The models we have developed in our company are ready to be used by any internal application which needs to compute ageing indicators on our products, enabling maintenance optimization. This is done through a web-service, fed by the current ageing indicators (calculated in a previous step) and by the time series reflecting the values of all environmental and usage variables influencing the evolution of the ageing factors since previous computation (wherever available). The result of this computation is the updated value of the ageing indicators, including an estimation of the remaining useful life, and an

estimation for partial and global failure rates. This data will be used in the same way for the next iteration.

Products which are not permanently connected to the cloud can also benefit from this methodology, using several possible mechanisms: from embedded models and computation, to local data storage and computation thanks to mobile phones (which are able to get data directly from the product, using wireless techniques). In this last case, the results of computation will be made available both for the user on his mobile, but also pushed to the cloud whenever possible, to compile all information related to a given customer site and enable the use of these data to improve the models and build relevant trends.

Another major point concerning these models, is that they can consider the effects of a maintenance action, thus reducing the corresponding ageing indicator. Therefore, on each sub-assembly, and depending on the completeness of maintenance, we can either start again from the beginning (when changing a coil, for example), or reduce the ageing indicator by a given factor (for example when a renewal of grease is performed on a circuit breaker).

The computation is done iteratively, since this ageing engine is stateless: as inputs, it welcomes the current state of the considered asset and the current values of environment and usage factors, then provides the updated state of the asset. It can be called as often as the client wants it, this frequency depending for example on the criticality of the asset. "Stateless" means that if the customer wants to study the real evolution of his assets' health, he (or we) has (have) to store the data. This procedure is also used to forecast the health of the assets, playing what-if scenarios, hence enabling the pre-validation of models by comparing the predicted results with the opinion of experts in terms of failure frequency.

3.6 New services for design and maintenance

Multiple actors benefit from these models and the corresponding infrastructure. First, the end users get a better view of the impact of environment and usage conditions on the lifetime of their products. They can therefore put in place relevant actions to improve these conditions: for example, they would protect critical products with the adequate material, e.g. to reduce temperature (by adding fans whenever relevant), or to reduce the influence of corrosive gases (by adding dedicated hoods for corresponding products).

They also have the right information to optimize their maintenance (OPEX) and renewal (CAPEX) plans, and they can also compute what-if scenarios, such as: "What if the global temperature increases in my region? Which impact on my assets?" or "What if I buy dedicated equipment to provide air-condition to my assets? Will it be worthwhile in terms of life duration, given the expected Return on Investment?". The ageing estimation will also enable them to compare the performances of different customer sites (fostering benchmarking).

On another side, the manufacturer will also be able to improve the design of its products, since we now can get accurate information on the weakest parts of the assets (the ones more likely to be affected by frequent harsh environmental or usage conditions). This knowledge, well-centered on what the customers need from their assets, will help define more robust products within the next years.

In parallel, the knowledge of the customers' mission profiles will also be highly improved, which will help a better focus on real technical performance and endurance needed for the assets.

In addition, accurate information on ageing conditions enables field service teams to deliver more efficiently their maintenance offer.

And finally, segment end user communities could leverage this new information to make their internal operational standard evolve consequently. This could be done by sharing information among the communities, taking the example of the OREDA initiative [Sintef].

4 CONCLUSION

Digitization now brings a lot of data. This unlocks new possibilities for asset monitoring, both for short-term anomaly detection and long-term ageing estimation. For short-term, Machine Learning techniques have shown to be as efficient as gray-box physical models to tackle the problem of anomaly detection for a range of electrical assets, in particular for detecting thermal anomalies.

For long-term analysis, the quantity of collected data unlocks new models for ageing, new services, and new ways of considering the exploitation of products on site. Additionally, the improved knowledge on customer mission profiles will also enable designing products more adapted to the real usage of customers, and new data will also provide better estimations for failure rates and their evolution in time, under given conditions of environment and usage.

The systematic application of these types of models in the electrical field is rather new and will provide significant added-value in the field of assets and system maintenance. Finally, mutualizing these machine learning and ageing engines for various applications will ensure consistency, as viewed by the customers. And when it comes to sharing all these data among different actors, it will make their precision much higher than it used to be.

5 REFERENCES

- Schneider Electric (2019a). Exchange, a platform for analytics sharing, <https://exchange.se.com>.
- Schneider Electric (2019b). Electrical fire prevention, <https://www.se.com/ww/en/work/solutions/safety/electrical-fire-prevention/>.
- Hirschbold M., & Chabert D. (2018). How thermal monitoring reduces risk of fire more effectively than IR thermography.
- Hoffman M., Wildermuth S., Gitzel R., Boyaci A., Gebhardt J., Kaul H., Amihai I., Forg B., Suriyah M., Leibfried T., Stich V., Hicking J., Bremer M., Kaminski L., Beverungen D., zur Heiden P., Tornede T. (2020). Integration of Novel Sensors and Machine Learning for Predictive Maintenance in Medium Voltage Switchgear to Enable the Energy and Mobility Revolutions. *Sensors (Basel)*, 2020;20(7):2099. 10.3390/s20072099.
- Curreri F., Fiumara G., & Xibilia M. (2020). Input Selection Methods for Soft Sensor Design: A Survey. *Future Internet* 12(6), 97, 10.3390/fi12060097.
- Souza, F.A.A., Araújo, R., & Mendes, J. (2016). Review of Soft Sensors Methods for Regression Applications. *Chemometrics and Intelligent Laboratory Systems*, Vol 152, 69–79, 10.1016/j.chemolab.2015.12.011.
- Meinshausen, N. (2006). Quantile Regression Forests. *Journal of Machine Learning Research*, Volume 7, Issue. 12/1/2006, 983–999, 10.5555/1248547.1248582.
- Gao, T., Boguslawski, B., Marié, S., Béguery, P., Thebault, S., & Lecoeuche, S. (2019). Data mining and data-driven modelling for Air Handling Unit fault detection. *E3S Web Conf., Volume 111, CLIMA 2019 Congress, Bucharest, Romania*, 10.1051/e3sconf/201911105009.
- Pagès A, Gondran M. (1980), *Fiabilité des systèmes*, Paris, éditions Eyrolles.
- Sintef (2015), *Offshore and Onshore Reliability Data*, 6th edition.
- Patent, "General Frame Product Health Estimation", Marcel Chevalier / Veronique Boutin / Michel Cercueil, INPI N° 1761110.
- Chevalier-Boutin, "Heading to Models of Failure Rate Evolution, with Respect to Environmental and Usage Conditions in Time", Marcel Chevalier / Veronique Boutin, RAMS Congress 2019, Orlando.

Incertitudes du pronostic de durée de vie de systèmes de production par IOHMM

KAMRUL-ISLAM SHAHIN, CHRISTOPHE SIMON, PHILIPPE WEBER

Université de Lorraine, CNRS, CRAN
F-54000 Nancy, France
christophe.simon@univ-lorraine.fr

Résumé – Dans cet article, nous prenons en compte l’incertitude de modèle et de paramètres dans le pronostic de l’état de santé d’un système de production. Un réseau bayésien dynamique représente sous une forme compacte un modèle de Markov caché à Entrées-Sorties utilisé pour le pronostic de l’état de santé d’un système de production. Une méthode de Bootstrap permet l’estimation des paramètres du modèle et une simulation de Monte-Carlo à 2 niveaux permet la considération des incertitudes dans la phase de pronostic en ligne de durée de vie résiduelle.

Abstract – In this article, we consider the model and the parameter uncertainties in the health state prognosis of a production system. A dynamic Bayesian network as a compact form of an Input-Output Hidden Markov model is used. A Bootstrap method for the parametric estimation of the model and a 2-level Monte Carlo simulation allows the consideration of uncertainties in the on-line prognosis phase through estimating the remaining useful life.

Mots clés - dégradation, état de santé, conditions opérationnelles, incertitude.

Keywords – degradation, Health State, Operational conditions, uncertainty.

1. INTRODUCTION

Dans le cadre de la maintenance conditionnelle et en cohérence avec les attentes de l’industrie 4.0, la question de l’estimation de la durée de vie résiduelle et plus globalement du pronostic est un problème difficile. C’est d’autant plus le cas lorsque l’élément, sur lequel le pronostic est à réaliser, se trouve soumis à de multiples conditions opérationnelles. Le pronostic de l’état de santé est alors fortement dépendant de la performance du diagnostic courant et de la connaissance obtenue sur l’évolution de l’état de santé passé (enregistré dans la base de données de retour d’expériences).

Le diagnostic est alors un point clé pour le calcul du pronostic de l’état de santé du système. Par ailleurs, un système même en apparence simple devient rapidement complexe lorsqu’il s’agit de pronostiquer sa durée de vie résiduelle, car il est soumis à de nombreuses interactions avec son milieu, et subit de nombreux mécanismes de dégradation concourants sous différentes contraintes opérationnelles qui sont autant d’agresseurs pour son état de santé, et ceci tout au long de sa durée de vie. Aussi, un diagnostic optimal est difficile à obtenir si on considère l’ensemble des contraintes de fonctionnement et leur impact sur la dégradation de l’état de santé.

La qualité du pronostic d’évolution de l’état de santé sera impactée aussi bien par la qualité du diagnostic, que par la qualité de la modélisation et l’estimation de ses paramètres.

Les modèles de diagnostic/pronostic connus dans la littérature relèvent principalement des approches basées modèles ou basées données et également des approches hybrides [Tidri et al. 2016]. Dans le cas d’un processus de dégradation, il peut être difficile de formaliser un modèle quantitatif, d’une part parce qu’il est difficile d’accéder aux données de dégradation, d’autre part parce que le processus est complexe. En outre, le concept de dégradation est complexe à appréhender et ce sont ses effets qui sont mesurés

à travers d’éventuel mode de défaillances et non une observation directe de la dégradation. Les modèles qualitatifs sont également assez difficiles à mettre en œuvre, car ils nécessitent généralement une bonne connaissance de la physique de dégradation et des interactions entre processus concourants. Si le modèle de dégradation est compliqué, la dégradation non mesurable, les probabilités sont un langage pertinent de modélisation. Aussi, une approche basée sur les données et un modèle de type stochastique représentent une bonne alternative.

Les méthodes stochastiques sont nombreuses dans le domaine de la sûreté de fonctionnement et les modèles markoviens sont largement utilisés [Mechri et al. 2013], mais aussi les réseaux de Petri stochastiques [Aubry et al. 2015] même si chacun d’entre eux apporte son lot d’avantages et de contraintes. Les réseaux bayésiens statiques et dynamiques représentent une évolution intéressante [Weber et Simon 2016] des chaînes de Markov, car ils en sont une forme compacte. Les modèles stochastiques présentent donc l’intérêt de considérer l’incertitude inhérente aux données et à notre incapacité à élaborer un modèle déterministe exact.

Finalement, pour modéliser une évolution de nature stochastique tout en considérant que cette évolution n’est pas mesurée, les modèles de la famille des Modèles de Markov Cachés semblent pertinents et peuvent trouver une correspondance élégante dans les réseaux bayésiens sur lesquels nous avons une expertise [Weber et Simon 2016]. Il existe bien évidemment d’autres modèles stochastiques comme les processus Gamma, Wiener ... [Baussaron 2011].

Les modèles HMM ont été introduits par Baum au début des années 1970 [Baum et Petrie 1966], puis développés dans [Rabiner 1989] pour la reconnaissance de la parole. Leur usage à des fins de diagnostic et de pronostic a été proposé dans [Ying et al. 2000]. Le diagnostic d’un outil de coupe par HMM a été développé dans [Kumar et al. 2019]. Une approche avancée de diagnostic pour une turbine à gaz a été

proposée dans [Roemer et Kacprzynski 2000] et l'application sur une foreuse a été proposée dans [Ertunc et Loparo 2001]. Dans [Li et al. 2005], une méthode de diagnostic appliquée à une machine tournante est discutée. Deux autres modèles HMM sont implémentés dans [Zhang et al. 2005] et [Xinmin et al. 2007] pour la détection des défauts de roulements. Des applications de diagnostic dans un contexte d'usinage sont décrites dans [Tobon-Mejia et al. 2011 ; Cameci et Chinnam 2006 ; Baruah et Chinnam 2005]. Pour l'ensemble de ces applications, seuls des modèles classiques de HMM sont employés. L'hypothèse sous-jacente est donc que le processus est homogène et unique, car indépendant des conditions opérationnelles.

Certains modèles modifiés tels que le modèle semi-Markov caché (HsMM) sont utilisés dans [Dong et Peng 2012] et dans le cadre d'un diagnostic dans [Dong et He 2007a]. Un modèle de Markov caché à mélange gaussien (MogHMM) est utilisé pour le diagnostic de défauts [Jianmin et Xinghui 2012] et un modèle de Markov semi-caché non stationnaire (NSHSM) [Dong et He 2007b] est utilisé pour le pronostic de santé de pompes. Mais les limites inhérentes aux HMM, *c'est-à-dire* pour l'essentiel l'indépendance aux conditions opérationnelles existe toujours. Or, si un système est soumis à des conditions opérationnelles différentes, celles-ci ne sont pas considérées et le modèle est un modèle moyen ou indépendant à ces conditions. Par ailleurs, les modèles de type HMM sont généralement basés sur une seule variable d'observation, or les conséquences d'une dégradation peuvent s'observer indirectement sur un ensemble de variables. Ainsi, un HMM ordinaire ne peut être utile pour réaliser un diagnostic dès lors que le problème considère des conditions de fonctionnement multiples et des observations multiples. Dans [Le 2015 ; Le et al. 2015], un modèle HMM à multiples branches a été exploité. Il s'agit en fait d'un ensemble de HMM classiques entraînées chacune avec des conditions opérationnelles fixes. L'usage en diagnostic puis en pronostic suppose de sélectionner le bon modèle HMM ou de commuter entre les modèles. La difficulté majeure réside dans la construction de lots de données conformes pour l'apprentissage de chaque HMM puis une sélection du modèle le plus approprié pour réaliser un bon diagnostic et un meilleur pronostic.

Pour tenir compte des conditions opérationnelles dans un processus markovien caché, mais également pour modéliser les multiples variables d'observations, dans [Bengio et Frasconi 1994], les auteurs ont proposé l'Input-Output Hidden Markov Model (IOHMM). Dans [Weber et al. 2004; Ben Salem et al. 2006] les auteurs ont travaillé avec un IOHMM pour la fiabilité du système à l'aide d'un réseau bayésien dynamique (RBD) à variables exogènes qui sont des formes compactes d'IOHMM [Weber et Simon 2016]. Ce type d'IOHMM sous forme de RBD a été utilisé pour diagnostiquer un générateur diesel dans [Klingelschmidt et al. 2017]. Dans [Shahin et al. 2019], nous avons proposé une méthode d'estimation des paramètres d'un IOHMM à partir de séquence de données d'observation de manière non supervisée. Dans [Shahin et al. 2020], nous avons également exploité une technique Bootstrap pour considérer l'incertitude de modèle dans le processus d'estimation.

Le travail proposé ici est de considérer d'une part l'incertitude de modèle et d'autre part l'incertitude de données pour le pronostic de l'évolution de l'état de santé d'un système. Dans cet objectif, l'article est structuré comme suit :

la section 2 donne la base théorique des IOHMM et les algorithmes fondamentaux d'apprentissage avec Bootstrap. La section 3 s'intéresse au pronostic et la section 4 montre les résultats de la prise en compte de l'incertitude. Enfin, la section 5 dresse quelques conclusions et perspectives.

2. ÉLÉMENTS DE BASE DES IOHMM

Un Hidden Markov Model (HMM) est un processus stochastique Markovien couplé à un second processus stochastique qui produit des données d'observation. Un IOHMM est une généralisation d'un HMM au sens où il s'agit d'un HMM comportant des entrées permettant la commutation du processus stochastique Markovien et des processus d'émissions de multiples sorties. Leur complexité les rend difficilement représentables, mais ils peuvent être modélisés par des réseaux Bayésiens dynamiques à deux pas de temps [Weber et Simon 2016].

Revenons à une description plus formelle : supposons que X_k représente l'état caché non observable du niveau de dégradation ou ce que nous appelons aussi l'état de santé d'un système. Le modèle de dégradation sous forme d'un processus stochastique du système est donné par la chaîne de Markov $P(X_k|X_{k-1})$ tel que présenté à la Fig. 1.a. Si les probabilités de transition entre états sont fixes alors la chaîne de Markov est homogène. Il est alors très facile de modéliser cette chaîne comme un réseau bayésien dynamique à deux pas de temps (2TBN) où n'est considérée que la distribution sur les états de la variable aléatoire X pour deux pas temps consécutifs (Fig. 1.b) [Weber et Simon 2016].

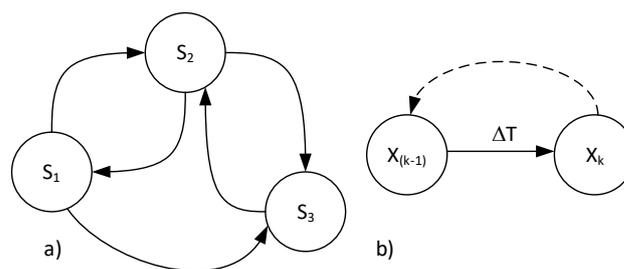


Figure 1. Chaîne de Markov à 3 états

Si l'état du système ne peut pas être observé directement, mais il est visible au travers d'une variable d'observation Y telle que $B = P(Y_k|X_k)$ alors la chaîne de Markov devient un HMM telle que représentée sous la forme d'un 2TBN à la Fig. 2.a. Par ailleurs, si la variable X à chaque instant k est conditionnée par l'état de la variable exogène U , traduisant les conditions opérationnelles à l'instant $k - 1$ alors il s'agit d'un IOHMM dont le modèle 2TBN est donné à la Fig. 2.b. Nous pouvons alors généraliser le modèle en considérant plusieurs sorties observées indépendantes ou non selon la complexité du modèle que nous souhaitons prendre en compte. Nous pourrions généraliser encore le modèle en considérant une dépendance entre les observations Y et les facteurs exogènes U .

Dans l'étude menée ici, les observations Y sont considérées mutuellement indépendantes et, la dépendance entre Y et U n'est pas considérée. La représentation compacte du IOHMM considéré est fournie à la Fig. 2.b.

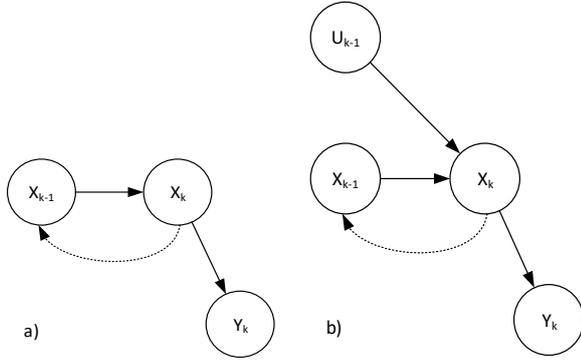


Figure 2. Structure de Réseaux Bayésiens Dynamiques (2TBN) équivalente aux HMM et IOHMM

La variable exogène U_k est connue ou observée et représente une condition opérationnelle (une commande ou une condition environnementale) conditionnant l'évolution du processus de dégradation (le processus Markovien). Il est à noter que U peut représenter un ensemble de conditions. Ce processus markovien caché s'écrit alors $A_N = P(X_k | X_{k-1}, U_{k-1})$ où N représente les états possibles de U_k [Ben Salem et al. 2006].

Formellement, un IOHMM est bien adapté à la modélisation de processus stochastiques complexes tels que la probabilité jointe est $P(Y_k, X_k, U_{k-1})$ [Klingelschmidt et al. 2017]. Si on tient compte de la séquence d'états cachés $X_{1:k-1}$ et de la séquence d'entrée $U_{1:k-1}$, il est possible de définir la probabilité jointe à partir de la distribution initiale des états cachés $\pi = P(X_1)$. Un IOHMM est donc complètement défini par le triplet : $\Lambda = (A_N, B, \pi)$. Si la variable Y est un vecteur de variables d'observation indépendantes, alors B est un ensemble de matrices de probabilités $B^M = P(Y_k^M | X_k)$ avec M l'indice de la variable d'observation.

La commutation du processus caché se fait sur la base du mécanisme de commutation comme proposé dans [Kuan 2013]. Cette commutation induit directement un problème algorithmique d'apprentissage des paramètres à partir de séquences multiples. Les algorithmes d'apprentissage de paramètres proposés par Baum et Welch [Baum et Petrie 1966], tout comme l'algorithme de Viterbi [Viterbi 1967] pour connaître l'état caché, sont dédiés aux HMM et ne peuvent s'appliquer directement aux IOHMM et demandent une adaptation exposée ci-après.

1.1 Algorithme d'apprentissage non supervisé

Les algorithmes Forward et Backward combinés à l'algorithme de Baum-Welch sont largement utilisés dans les problèmes de HMM. Adaptés au contexte des IOHMM [Shahin 2020], ils permettent de calculer efficacement la probabilité qu'une séquence ait été générée par un IOHMM. Nous pouvons alors étendre l'estimation à une série de séquences d'observations. Les algorithmes Forward et Backward supposent que le modèle $\Lambda = (A_N, B, \pi)$ soit connu. Si la séquence d'observations de la variable Y est donnée alors les algorithmes peuvent calculer $P(X|Y, U)$ à l'aide des algorithmes suivants sachant qu'ici une récurrence est proposée pour réduire le coût calculatoire :

Forward

- *Initialisation* : $\alpha(X_1) = \prod_{i=1}^p P(Y^i | X_1) P(X_1)$
- *Récurrence* :

$$\alpha(X_k) = \sum_{X_{k-1}=1}^N \alpha(X_{k-1}) P(X_k | X_{k-1}, U_{k-1}) \prod_{i=1}^p P(Y_k^i | X_k) P(X_k)$$

pour $2 \leq k \leq K$.

Backward

- *Initialisation* : $\beta(X_K = s_i) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, pour $k = K$.
- *Récurrence* :

$$\beta(X_k) = \sum_{X_{k+1}=1}^N \beta(X_{k+1}) \cdot \prod_{i=1}^p P(Y_{k+1}^i | X_{k+1}) P(X_{k+1} | X_k, U_k)$$

pour $(K-1) \leq k \leq 1$.

Dans les équations précédentes, $\prod_{i=1}^p P(Y_{k+1}^i | X_{k+1})$ est le produit de Hadamard des distributions de probabilités et U_k est l'état combiné des entrées.

Sur cette base algorithmique, la probabilité que le HMM ait généré la séquence d'observation avec l'équation 1 est calculée comme suit :

$$P(Y_k | \Lambda) = \sum_{i=1}^N \alpha_i(k) \beta_i(k) ; \forall k \quad (1)$$

Ces deux algorithmes évaluent la probabilité que le HMM ait généré la séquence d'observations de la variable Y . Toutefois, une seule variable est considérée et il n'y a pas de conditionnement aux variables d'entrée, ce qui demande des modifications ad hoc pour l'application aux IOHMM.

Une fois capable de calculer $P(Y_k | \Lambda)$, il faut déterminer les paramètres $\Lambda = (A_N, B, \pi)$ sachant les séquences d'observation o . Il s'agit de trouver Λ qui maximise $P(o | \Lambda)$. Dans ce cas, l'algorithme de Baum-Welch est utilisé. Il s'agit d'un algorithme d'estimation-maximisation de type programmation dynamique. Il peut éventuellement converger vers un minimum local. Il existe d'autres algorithmes, notamment par métaheuristique [Aupetit 2005], qui pourront être exploités.

Le problème de maximisation est de grande complexité algorithmique. À l'aide des algorithmes précédents, nous calculons la probabilité que le HMM Λ ait généré l'ensemble des séquences Y grâce à l'équation 2. La variable auxiliaire $\varepsilon_k(i, j)$ définie par l'équation 3 où i, j, k représente respectivement les indices des états cachés, les symboles d'émission et le mode de la variable exogène U .

$$P(Y_k | \Lambda) = \sum_{i=1}^N \alpha_i(X_k) \beta_i(X_k) ; \forall k \quad (2)$$

$$\varepsilon_k(i, j) = \frac{\alpha_i(X_k) \cdot \alpha_{ij}^p \cdot b_{jk}^q \cdot \beta_j(X_{k+1})}{P(Y_k | \Lambda)} \quad (3)$$

Ainsi, l'algorithme suivant est défini :

1. Distribution initiale : $\pi_i = \varepsilon_1(i, j)$, où $1 \leq i \leq N$
2. Probabilité de transition :

$$a_{ij}^p = \frac{\sum_{k=1}^{K-1} \varepsilon_k(i, j) \cdot \mathbf{1}_{U_k=p}}{\sum_{k=1}^{K-1} \gamma_k(j) \cdot \mathbf{1}_{U_k=p}}$$

3. Probabilité d'émission :

$$b_{jk}^q = \frac{\sum_{k=1}^K \gamma_k(j) \cdot 1_{Y_k^q=v_m}}{\sum_{k=1}^K \gamma_k(j)}$$

Ces étapes sont répétées jusqu'à ce que les changements entre deux résultats convergent.

$1_{X_k(U_k=p)}$ est la fonction indicatrice telle que : $1_{X_k(U=p)} = \begin{cases} 1 & \text{if } U_k = p \\ 0 & \text{sinon} \end{cases}$

De manière similaire, on a : $1_{Y_k^q=v_m} = \begin{cases} 1 & \text{if } Y_k^q = v_m \\ 0 & \text{sinon} \end{cases}$.

1.2 Diagnostic de l'état caché : Algorithme de Viterbi

Cet algorithme permet le calcul de la séquence la plus probable d'états cachés étant donné les observations, ce qui est nommé le chemin de Viterbi. L'état le plus probable à chaque instant est obtenu par l'équation (eq.4). Ceci constitue le diagnostic attendu du système. L'algorithme se définit ainsi :

$$\begin{aligned} P(X|Y^1, \dots, Y^q, U) &\propto P(X, Y^1, \dots, Y^q, U) \\ &= \max_{(X_{1:k})} P(X|Y^1, \dots, Y^q, U) \\ &= \max_{(X_{1:k})} P(X, Y^1, \dots, Y^q, U) \end{aligned} \quad (4)$$

- *Initialisation :*

$$\omega(X_1) = P(X_1, Y_1^1, \dots, Y_1^q)$$

- *Réurrence :*

$$\begin{aligned} \omega(X_k) &= \max_{(X_{1:k-1})} P(X_{1:k}, Y_{1:k}^1, \dots, Y_{1:k}^q, U_{1:k}) \\ &= \max_{(X_{k-1})} \sum_{i=1}^p P(Y_k^i | X_k) P(X_k | X_{k-1}, U_{k-1}) \omega(X_{k-1}) \end{aligned}$$

À chaque nouvelle observation Y_k , il est donc possible de réestimer le chemin de Viterbi et notamment l'état de la chaîne cachée X_k . Cela constitue une forme de diagnostic en ligne de l'état de santé ou du niveau de dégradation du système modélisé par le processus caché. Il est très important de noter ici que le chemin de Viterbi à l'instant k peut être différent de celui à l'instant $k-1$ grâce à la connaissance apportée par la nouvelle observation.

1.3 Estimation de l'incertitude de modèle - Bootstrap

La méthode Bootstrap est une méthode d'échantillonnage de population utilisée ici pour estimer des indicateurs statistiques (moyenne, variance ...) sur les paramètres estimés d'un modèle IOHMM. Il s'agit d'une méthode d'échantillonnage statistique avec remise. L'échantillonnage avec remise sélectionne les séquences $Y-U$ dans le jeu de données originales pour calculer les paramètres puis remet les données pour un prochain tirage. Une séquence peut donc être tirée plusieurs fois. Cette technique permet de maintenir la structure du jeu de données tout en remaniant les valeurs pour extrapoler la population. La méthode Bootstrap est utile même si les données ne suivent pas une loi normale, si elles ont des propriétés statistiques inconnues et permet des calculs des indicateurs d'incertitude tels les intervalles de confiance, etc.

Quelques éléments sont importants pour comprendre la méthode Bootstrap :

- **Intervalle de Confiance (IC) :** l'intervalle de confiance est estimé à partir des données statistiques observées qui peuvent contenir une population de paramètres inconnus de la population pour donner une information statistique de la précision de l'estimation. La taille de l'intervalle dépend des variations dans la population et de la taille de l'échantillon.
- **Variation de la population :** Si les valeurs des données de la population sont semblables, alors l'échantillon possède également une variation faible et l'intervalle de confiance sera naturellement faible. Si au contraire, les données sont très variables alors les échantillons auront une variation surestimée, ce qui conduira à une estimation moins fiable de la moyenne et l'intervalle de confiance sera plus grand.
- **Taille de l'échantillon :** La taille de l'échantillon de la population influence également la largeur de l'intervalle de confiance. Une taille faible induit une variation plus forte et un intervalle de confiance plus large. Une taille d'échantillon élevé induit moins de variabilité inter-échantillon et crée un intervalle de confiance plus petit, car les échantillons présentent moins d'information.
- **Calcul des intervalles de confiance :** L'intervalle de confiance pour une moyenne utilise l'équation 5 :

$$CI = \bar{X} \pm t \frac{s}{\sqrt{n}} \quad (5)$$

où \bar{X} est la moyenne estimée, t est la distribution de Student qui dépend de la taille de l'échantillon et du niveau de risque choisi, s est la distance à la moyenne estimée et n la taille de l'échantillon.

1.4 Pronostic de durée de vie par simulation de Monte-Carlo

Le pronostic par un IOHMM consiste à déterminer l'instant où l'état caché indésirable X est atteint sachant que l'état actuel diagnostique est défini par l'équation 3. Comme seule l'évolution de l'état de santé nous intéresse, seule la partie cachée est utile à l'estimation de la durée d'atteinte de l'état considéré. Néanmoins, les IOHMM intègrent en entrée du processus caché, les entrées exogènes U . Il faut donc déterminer ces entrées U dans le futur pour déterminer le temps d'atteinte. Il y a donc un nombre très important de séquences $U_{k:K}$ possibles. Dans le cadre d'une simulation de Monte-Carlo de l'évolution de la chaîne cachée conditionnée par l'entrée U , nous pouvons réduire le nombre de valeurs possibles d'entrées par exemple en considérant des valeurs constantes ou en fixant des propriétés statistiques d'évolution des valeurs de cette entrée.

Partant d'un état courant X_k donné par l'équation 3, nous générons une trajectoire d'états $X_{k:K}$ où K représente l'instant où la trajectoire atteint l'état final. À chaque instant k , la distribution de l'état futur est déterminée par $P(X_{k+1} | X_k = S_{i \neq n})$ où S_n et l'état final. Cette simulation est répétée jusqu'à ce que la distribution du temps d'atteinte soit stable ou à défaut pour un nombre important de simulations.

En parallèle de l'incertitude paramétrique dans l'estimation de la durée de vie résiduelle, nous pouvons également considérer l'incertitude de modèle liée à l'imprécision sur les paramètres de la chaîne de Markov cachée que nous avons déterminée à l'aide de la méthode Bootstrap précédente.

L'algorithme de simulation est donc le suivant :

1. Choisir un jeu de paramètre de $P(X_{k+1}|X_k, U_k)$
2. Simuler une trajectoire jusqu'à $X_{k1} = S_n$ et collecter $k1$
3. Répéter 1-2 jusqu'à $P(k1)$ stable

3. ILLUSTRATION NUMERIQUE (SIMULATION)

Pour illustrer les éléments définis précédemment, nous utilisons l'application numérique décrite ci-après.

Considérons un processus à 5 états cachés, 2 entrées U_1 et U_2 à 2 états chacune, et 2 sorties Y_1 et Y_2 possédant respectivement 3 et 4 états. Le modèle original est défini par :

Le modèle estimé est donné ci-dessous :

$$A_1 = \begin{pmatrix} 0.999 & 0.001 & 0 & 0 & 0 \\ 0 & 0.999 & 0.001 & 0 & 0 \\ 0 & 0 & 0.999 & 0.001 & 0 \\ 0 & 0 & 0 & 0.999 & 0.001 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 0.996 & 0.004 & 0 & 0 & 0 \\ 0 & 0.996 & 0.004 & 0 & 0 \\ 0 & 0 & 0.999 & 0.001 & 0 \\ 0 & 0 & 0 & 0.999 & 0.001 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0.999 & 0.001 & 0 & 0 & 0 \\ 0 & 0.998 & 0.002 & 0 & 0 \\ 0 & 0 & 0.996 & 0.004 & 0 \\ 0 & 0 & 0 & 0.998 & 0.002 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0.993 & 0.007 & 0 & 0 & 0 \\ 0 & 0.996 & 0.004 & 0 & 0 \\ 0 & 0 & 0.996 & 0.004 & 0 \\ 0 & 0 & 0 & 0.996 & 0.004 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 0.5 & 0.4 & 0.1 & 0 \\ 0.2 & 0.4 & 0.2 & 0.2 \\ 0 & 0.1 & 0.3 & 0.6 \\ 0 & 0.2 & 0.4 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \end{pmatrix}, B_2 = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0.4 & 0.6 \\ 0 & 0.6 & 0.4 \\ 0 & 0.8 & 0.2 \end{pmatrix}$$

Pour appliquer la méthode Bootstrap et les algorithmes d'estimation des paramètres du IOHMM, le lot d'apprentissage est constitué de 100 séquences $Y-U$. La méthode Bootstrap utilise 50 groupes de 20 séquences choisies aléatoirement parmi les 100 séquences.

La distribution des valeurs estimées des paramètres du processus caché conditionné par l'entrée $U=2$ est illustrée à la figure 3 i.e. A_2 . Nous n'avons illustré que A_2 pour des raisons de place. On notera que pour chaque paramètre, on peut également observer la valeur réelle (en vert) et le paramètre moyen estimé (en rouge). L'écart entre la valeur réelle et la valeur estimée donne un aperçu de la précision obtenue. Les intervalles de confiance n'ont pas été tracés ici, mais auraient pu être donnés sachant toutefois les limites

d'usage de la méthode Bootstrap pour leur estimation. De la même manière, la figure 4 illustre l'estimation des probabilités d'émission B_1 .

À partir des paramètres estimés et de leur distribution et avec une séquence de données inconnue, l'algorithme modifié de Baum-Welch permet l'estimation des états cachés et le pronostic de durée de vie résiduel peut être estimé via une méthode de Monte-Carlo comme énoncé précédemment. La figure 5 illustre le résultat obtenu à partir d'une séquence qui n'a pas participé à l'estimation de paramètres du modèle IOHMM. Les 2 courbes du haut présentent les données de sortie Y_1 et Y_2 (bleu et rouge) et l'entrée U issue de la combinaison des 2×2 états des entrées U_1 et U_2 . La troisième courbe à partir du haut montre la comparaison entre l'état caché et l'état caché estimé à partir du modèle moyen. Pour cette séquence, il n'y a pas de différence notable, mais il arrive que les instants de changement d'état soient différents. La courbe du bas illustre le pronostic moyen en ligne pour les 4 niveaux d'entrée futurs possibles c'est-à-dire en considérant que les entrées futures resteront constantes. Pour chaque courbe, il s'agit d'un tirage de Monte-Carlo à 2 niveaux à partir des distributions de paramètres puis des trajectoires d'états. L'instant de première atteinte de l'état de défaillance sert au calcul de la valeur moyenne représentée ici.

La défaillance réelle du système a lieu à l'instant $k1=1227$. En début de fonctionnement, le pronostic en ligne surestime assez fortement la durée de vie du système si l'entrée $U=1$ et sous-estime plus légèrement la durée de vie si $U=4$ puisque globalement la séquence d'entrée U est principalement faible en début de séquence et est majoritairement élevée en fin de séquence. Plus le processus avance dans le temps, plus le pronostic de la durée de vie résiduelle est proche de la réalité, ce qui est relativement logique. Toutefois, nous remarquons bien l'importance de connaître la séquence d'entrée, car cette séquence a bien une influence très importante. Cette information ne pourra être prise en compte que si le plan directeur de production est connu ou prédéfini.

Les transitions d'états de la chaîne cachée sont les moments où l'estimation de durée de vie prend plus de sens et offre une meilleure correspondance avec la réalité. Le choix d'un modèle caché à 5 états (4 transitions) est discutable par rapport à cette dernière remarque. Il s'agit souvent d'un compromis à faire entre le volume et la qualité du jeu de données et le nombre de paramètres à estimer, mais aussi de la précision que l'on cherche à obtenir sur le pronostic.

4. CONCLUSION

Dans cet article, nous avons proposé d'exploiter un réseau bayésien dynamique représentant une forme évoluée de modèle de Markov caché pour représenter le processus d'évolution de l'état de santé d'un système de production. Ce modèle correspond bien au principe d'évolution d'un élément (état de santé) complexe et non observable directement, mais au travers de variables d'observation. Sa nature stochastique permet la prise en compte de l'incertitude. La méthode Bootstrap nous a permis d'avoir une estimation satisfaisante de la valeur des paramètres du modèle, mais aussi la distribution de ces paramètres. Par ailleurs, le pronostic par méthode de Monte-Carlo à 2 niveaux avec tirage des paramètres du modèle et des trajectoires d'états cachés donne une estimation cohérente de la durée de vie résiduelle.

Dans ce pronostic, nous retrouvons des plateaux caractéristiques d'estimations de la durée résiduelle. Cela est dû à une bonne estimation des états cachés avec des probabilités élevées. Les instants de transition des états cachés sont des instants caractéristiques où l'estimation est assez fidèle à la réalité. L'augmentation du nombre d'états

cachés pourrait nous donner plus d'éléments de durée de vie résiduelle, mais une étude s'impose pour confirmer la validité des résultats obtenus. Les HsMM combinés avec des entrées-sorties seraient sur ce point relativement intéressant avec une gradualité plus fine bénéfique à un usage en pilotage de l'état de santé par exemple.

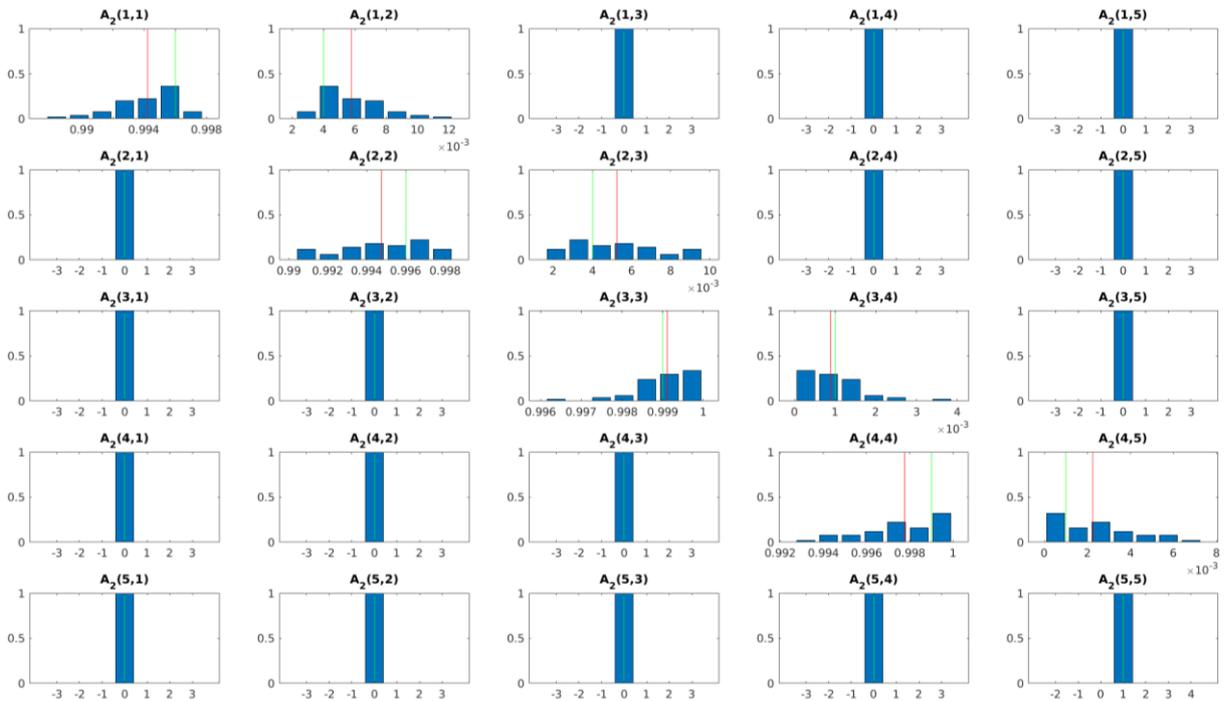


Figure 3. Distribution des paramètres estimés de A_2

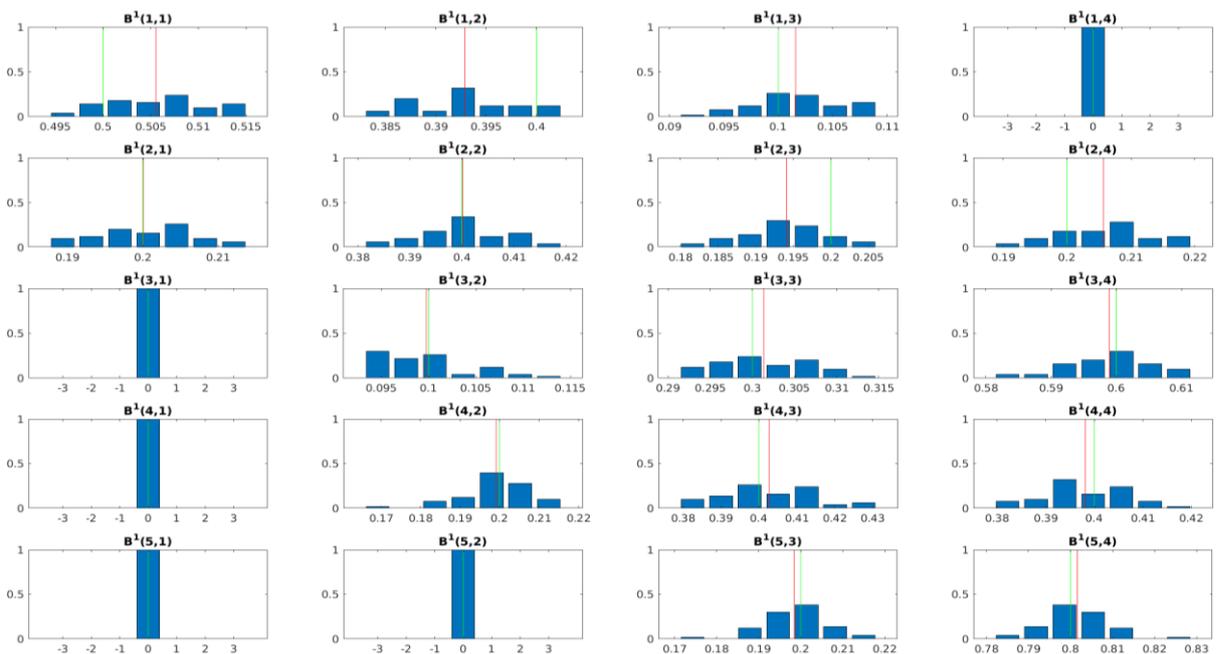


Figure 4. Distribution des paramètres estimés de B_1

5. BIBLIOGRAPHIE

- Aubry, Jean-François, Nicolae Brinzei, et Mohammed-Habib Mazouni. 2015. *Benefits of Petri Nets Models*. Vol. 1. Systems Dependability Assessment SET. ISTE WILEY.
- Aupetit, Sébastien. 2005. « Contributions aux Modèles de Markov Cachés : métaheuristiques d'apprentissage, nouveaux modèles et visualisation de dissimilarité ». Doctorat, Tours : Université François Rabelais.
- Baruah, P., et R. B. Chinnam. 2005. « HMMs for diagnostics and prognostics in machining processes ». *International Journal of Production Research* 43 (6): 1275-93.
- Baum, Leonard E., et Ted Petrie. 1966. « Statistical Inference for Probabilistic Functions of Finite State Markov Chains ». *The Annals of Mathematical Statistics* 37 (6) : 1554-63.
- Baussaron, Julien. 2011. « Mise au point de modèles prédictifs de fiabilité dans un contexte de dégradation associé à des profils de mission ». Doctorat, Angers : Université d'Angers.
- Ben Salem, Abdeljabbar, Alexandre Muller, et Philippe Weber. 2006. « Dynamic Bayesian Networks in System Reliability Analysis ». IFAC Proceedings Volumes, 6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, 39 (13): 444-49.
- Bengio, Yoshua, et Paolo Frasconi. 1994. « An Input Output HMM Architecture ». In Proceedings of the 7th International Conference on Neural Information Processing Systems, 427-434. NIPS'94. Cambridge, MA, USA : MIT Press.
- Camci, F., et R.B. Chinnam. 2006. « Hierarchical HMMs for Autonomous Diagnostics and Prognostics ». In 2006 IEEE International Joint Conference on Neural Network Proceedings, 2445-52.
- Dong, Ming, et David He. 2007a. « Hidden Semi-Markov Model-Based Methodology for Multi-Sensor Equipment Health Diagnosis and Prognosis ». *European Journal of Operational Research* 178 (3): 858-78.
- Dong, Ming, et David He. 2007b. « A Segmental Hidden Semi-Markov Model (HSMM)-Based Diagnostics and Prognostics Framework and Methodology ». *Mechanical Systems and Signal Processing* 21 (5): 2248-66.
- Ertunc, H. M, et K. A Loparo. 2001. « A Decision Fusion Algorithm for Tool Wear Condition Monitoring in Drilling ». *International Journal of Machine Tools and Manufacture* 41 (9): 1347-62.
- Jianmin, CD, et Z. Xinghui. 2012. « Fault Diagnosis of Gearbox based on EEMD and MGHMM ». *Journal of Mechanical Transmission* 11 (8).
- Jie Ying, T. Kirubarajan, K.R. Pattipati, et A. Patterson-Hine. 2000. « A hidden Markov model-based algorithm for fault diagnosis with partial and imperfect tests ». *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 30 (4): 463-73.
- Klingelschmidt, T., P. Weber, C. Simon, D. Theilliol, et F. Peysson. 2017. « Fault diagnosis and prognosis by using Input-Output Hidden Markov Models applied to a diesel generator ». In 25th Mediterranean Conference on Control and Automation (MED), 1326-31.
- Kuan, Chung-Ming. 2013. « Markov Switching Model ». *Quantile*, no 11: 13-40.
- Kumar, Akhilesh, Ratna Babu Chinnam, et Finn Tseng. 2019. « An HMM and Polynomial Regression Based Approach for Remaining Useful Life and Health State Estimation of Cutting Tools ». *Computers & Industrial Engineering* 128 : 1008-14.
- La, T.M.H, K.T. Huynh, Yves Langeron, et Antoine Grall. 2018. « Dégradation d'un système dynamique : Modélisation dans le plan complexe ». In 21e congrès de Maitrise des Risques et Sécurité de Fonctionnement. Reims, France.
- Le, T. T. 2015. « Contribution to deterioration modeling and residual life estimation based on condition monitoring data ». Doctorat, Grenoble Alpes.
- Le, T. T., Christophe Bérenguer, et F. Chatelain. 2015. « Prognosis Based on Multi-Branch Hidden Semi-Markov Models: A Case Study ». IFAC-PapersOnLine, 9th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS 2015, 48 (21): 91-96.
- Li, Zhinong, Zhaotong Wu, Yongyong He, et Fulei Chu. 2005. « Hidden Markov model-based fault diagnostics method in speed-up and speed-down process for rotating machinery ». *Mechanical Systems and Signal Processing* 19: 329-39.
- Liu, Qinming, Ming Dong, et Ying Peng. 2012. « A Novel Method for Online Health Prognosis of Equipment Based on Hidden Semi-Markov Model Using Sequential Monte Carlo Methods ». *Mechanical Systems and Signal Processing, Uncertainties in Structural Dynamics*, 32: 331-48.
- Mechri, Walid, Christophe Simon, Frédérique Bicking, et Kamel Ben Othman. 2013. « Probability of failure on demand of safety systems by Multiphase Markov Chains ». In 2013 Conference on Control and Fault-Tolerant Systems (SysTol), 98-103.
- Rabiner, L.R. 1989. « A tutorial on hidden Markov models and selected applications in speech recognition ». *Proceedings of the IEEE* 77(2): 257-86.
- Roemer, M.J., et G.J. Kacprzyński. 2000. « Advanced diagnostics and prognostics for gas turbine engine risk assessment ». In 2000 IEEE Aerospace Conference. Proceedings (Cat. No.00TH8484), 6:345-53 vol.6.
- Shahin, Kamrul Islam, Christophe Simon, et Philippe Weber. 2019. « Input-output hidden Markov model for diagnosis of complex system ». In 13ème Conférence internationale de CIGI QUALITA. Montréal, Canada.
- Shahin, Kamrul Islam, Christophe Simon, et Philippe Weber. 2020. « Bootstrap confidence interval on IOHMM parameters for system health diagnostic under multiple operating conditions ». In 21st IFAC World Congress, IFAC 2020. Berlin, Germany.
- Shahin, Kamrul-Islam. 2020. « Modèle graphique

probabiliste appliqué au diagnostic de l'état de santé des systèmes, au pronostic et à l'estimation de la durée de vie résiduelle ». Doctorat, Nancy : Université de Lorraine.

Tidiri, Khaoula, Nizar Chatti, Sylvain Verron, et Teodor Tiplica. 2016. « Bridging Data-Driven and Model-Based Approaches for Process Fault Diagnosis and Health Monitoring: A Review of Researches and Future Challenges ». *Annual Reviews in Control* 42: 63-81.

Tobon-Mejia, Diego, Kamal Medjaher, Noureddine Zerhouni, et Gérard Tripot. 2011. « Hidden Markov models for failure diagnostic and prognostic. » In *Prognostics and System Health Management Conference, PHM'11.*, 1-8. Shenzhen, China: IEEE Catalog Number : CFP1161H-PRT.

Viterbi, Andrew. 1967. « Error bounds for convolutional codes and an asymptotically optimum decoding algorithm ». *IEEE Transactions on Information Theory*.

Weber, P., P. Munteanu, et L. Jouffe. 2004. « Dynamic Bayesian Networks modelling the dependability of systems with degradations and exogenous constraints ». In *INCOM'04*.

Weber, Philippe, et Christophe Simon. 2016. *Benefits of Bayesian Network Models*. Vol. 2. Systems Dependability Assessment SET. ISTE WILEY.

Xiaodong Zhang, R. Xu, Chiman Kwan, S.Y. Liang, Qiulin Xie, et L. Haynes. 2005. « An integrated approach to bearing fault diagnostics and prognostics ». In *Proceedings of the 2005, American Control Conference, 2005.*, 2750-55 vol. 4.

Xinmin, Tao, Du Baoxiang, et Xu Yong. 2007. « Bearings Fault Diagnosis Based on HMM and Fractal Dimensions Spectrum ». In *2007 International Conference on Mechatronics and Automation*, 1671-76.

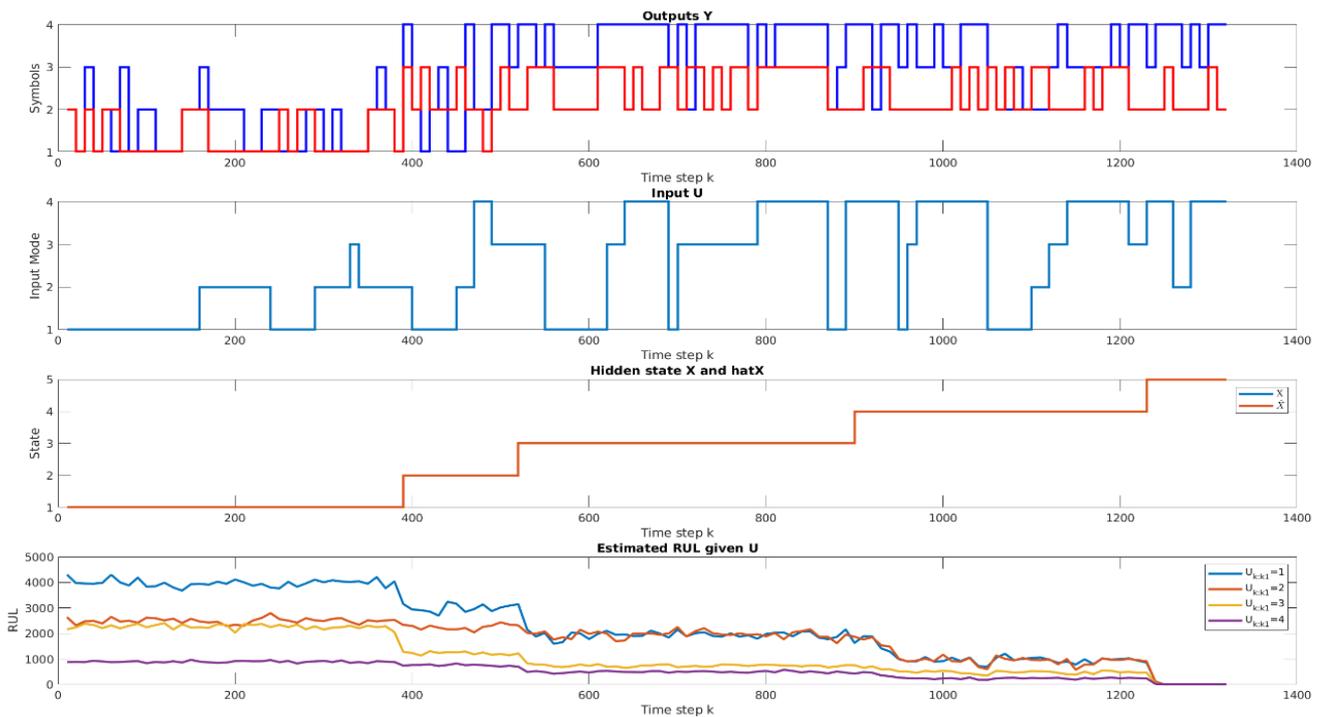


Figure 5. Pronostic en ligne par simulation de Monte-Carlo

**Session 2.4 - Session spéciale 12 :
Ingénierie des exigences et sûreté de
fonctionnement**

Chairman : Abdessamad Kobi - Georges Habchi

L'Excellence Opérationnelle au service de l'Ingénierie Système : Etat de l'art

JOSEPH KOKOLO¹, BENOIT EYNARD¹

¹LABORATOIRE ROBERVAL
UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE
Rue du Docteur Schweitzer, CS 60319
60203 Compiègne cedex, France
benoit.eynard@utc.fr

Résumé – L'excellence opérationnelle (EO) est de plus en plus présente dans l'actualité managériale ainsi que scientifique. La concurrence croissante, Les événements de plus en plus incertains, les exigences clients et sociétales de plus en plus pressantes, l'évolution des systèmes vers des systèmes IoT, poussent les organisations à adapter leurs méthodologies d'ingénierie. L'excellence opérationnelle (EO) est l'une des réponses proposées par la littérature scientifique pour rendre les organisations d'ingénierie plus flexibles, plus réactives, plus efficaces donc plus compétitives. Dans cet article, nous souhaitons partager un état de l'art de l'excellence opérationnelle (EO) dans l'ingénierie système (IS) à travers ses méthodologies les plus actuelles: le Lean Six Sigma (LSS), la Théorie des Contraintes (TC) et l'Agilité (A) avec une approche opérationnelle incluant la responsabilité sociale et sociétale via le pentagone Qualité-Coût-Délai-Sécurité-Environnement (QCDSE).

Abstract – Operational excellence (OE) is increasingly present in scientific and managerial news. Increasing competition, increasingly uncertain events, demands customers and society increasingly pressing, the evolution of systems towards cyberphysics systems, push organizations to adapt their engineering methodologies. Operational excellence (OE) is one of the answers proposed by the scientific literature to make engineering organizations more flexible, more responsive, more efficient and therefore more competitive. In this article, we want to share a state of the art on operational excellence (OE) in systems engineering (SE) through its most modern methodologies: the Lean Six Sigma (LSS), Theory of Constraints (TOC) and Agility (A) with an operational approach including social and societal responsibilities via the Quality-Cost-Delivery-Safety-Environment (QCDSE).

Mots clés – Excellence Opérationnelle, Ingénierie Système, Lean Six Sigma, Agilité, Qualité-Coût-Délai-Sécurité-Environnement.

Keywords – Operational Excellence, Systems Engineering, Lean Six Sigma, Agility, Quality-Cost-Delivery-Safety-Environment.

1 INTRODUCTION

L'ingénierie des systèmes prend un tournant technologique significatif ces dernières années. Nous nous dirigeons progressivement vers les systèmes IoT, interconnectés, d'une complexité exponentielle dont nous ne mesurons, naturellement, pas encore tous les aboutissants. Une chose est certaine, nos méthodologies d'ingénierie vont devoir s'adapter progressivement à cette réalité pour pouvoir rester compétitives. La littérature scientifique nous montre que seules des méthodologies adaptatives, apprenantes, englobantes répondent efficacement à des problématiques en perpétuel mouvement. A ce titre, l'excellence opérationnelle (EO) qui est une approche scientifique, managériale, opérationnelle, globale, devrait permettre aux organisations d'ingénierie de s'adapter continuellement à ce changement et aux changements à venir. Nombreux sont les ouvrages et les travaux à en faire état [Gershon, 2010] [Girard, 2017] [Jombart, 2016] [Rossi et al., 2016] [Meyer, 2014] [Ghavami, 2008]. La démarche d'excellence opérationnelle trouve naissance à la fin du XIXe

siècle, dans l'industrie américaine, avec le taylorisme. Henry Ford introduit ensuite la notion de *mouvement continu* dans les années 20. La *satisfaction du client* apparaît autour des années 1930-1940 grâce à Alfred Sloan. Puis, Edward Deming, entre 1945-1960, est le premier à introduire la notion de statistique dans le contrôle qualité, il donne naissance à la roue de l'amélioration continue [Jombart, 2016]. Le terme «Excellence opérationnelle» est un terme récent qui englobe toutes les méthodologies permettant de rendre les organisations continuellement plus performantes.

Dans les lignes qui suivent, nous vous proposons un état de l'art de l'EO à travers ces méthodologies les plus actuelles: le Lean Six Sigma (LSS), la Théorie des Contraintes (TC) et l'Agilité (A). L'EO étant un terme récent, nous vous proposons un état de l'art adapté à l'intérêt croissant de la recherche scientifique sur cette problématique. Nous introduirons le sujet avec une analyse statistique avant d'aborder la définition et les caractéristiques de chacune de ces méthodologies. Ensuite, nous parlerons de la complémentarité de ces méthodologies actuelles à travers les solutions de synergie: Le TLS (Théorie

des contraintes-Lean-Six Sigma) et le LAS (Lean Six Sigma-Agilité-Ingénierie Système). Puis, nous terminerons par une mise en perspective laissant envisager notre contribution future sur une méthodologie d'excellence opérationnelle destinée à l'ingénierie des systèmes embarqués, cyberphysiques et d'internet des objets [Hehenberger et al., 2016]. La Figure 1, ci-dessous, est une version adaptée de celle de [Morris, 2012]. Elle illustre l'évolution des méthodologies actuelles d'EO au fil du temps. Nous avons choisi d'y intégrer l'historique de l'ingénierie système pour mieux contextualiser nos travaux.

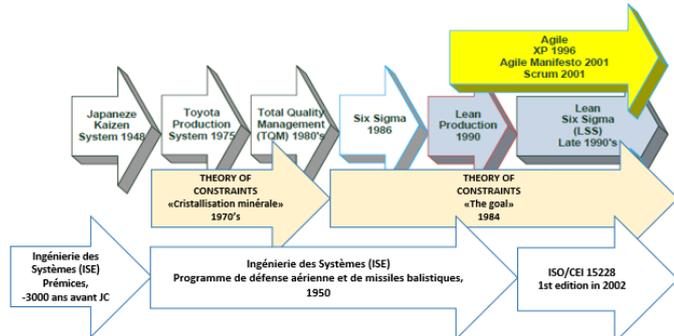


Figure 1. Historique de développements comparés de l'ingénierie système et de l'excellence opérationnelle.

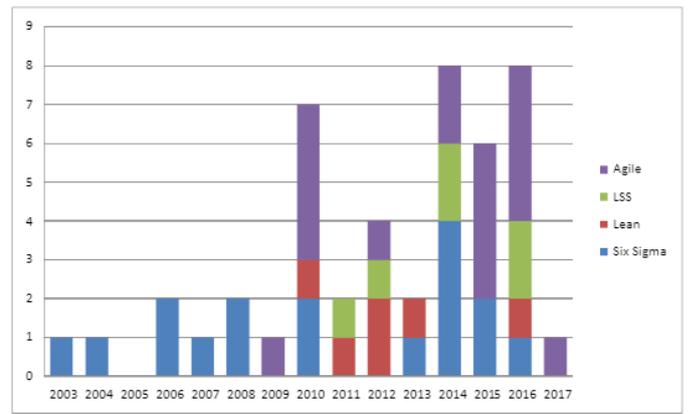


Figure 3. Répartition des publications sur les méthodologies abordées par ordre chronologique.

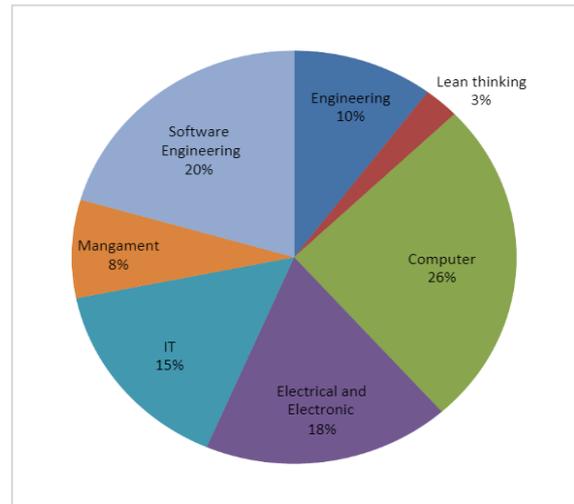


Figure 4. Répartition des publications académiques par domaine d'ingénierie système.

2 ETAT DE L'ART

2.1 Quelques chiffres

Dans cette section, nous vous présentons quelques données statistiques mettant en exergue l'importance croissante des méthodologies actuelles de l'EO dans la littérature scientifique via les travaux de [Badwe et Erkan, 2018]. En effet, dans leur article, Badwe et Erkan ont effectué une analyse statistique des publications académiques sur le Lean, le Six sigma, le Lean Six Sigma (LSS) et l'agilité (A). Cette analyse statistique a été conduite sur une période de plus de dix ans, de 2003 à 2017. Elle démontre, factuellement, le poids croissant que prennent les méthodologies actuelles d'excellence opérationnelle dans l'ingénierie des systèmes. Les Figures 2, 3, 4 sont des transcriptions graphiques des données obtenues grâce à cette analyse. Elles sont complémentaires et, ont été extraites des travaux de [Badwe et Erkan, 2018].

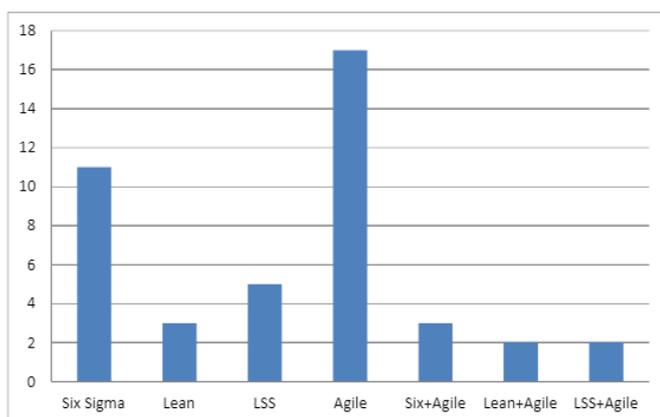


Figure 2. Occurrence des méthodologies abordées dans les publications académiques.

L'étude bibliographique faite par Badwe et Erkan ainsi que nos expériences professionnelles en ingénierie système, nous ont poussé à orienter nos travaux vers ces méthodologies actuelles de l'EO. Dans la section suivante, nous nous attarderons sur leurs définitions, leurs caractéristiques, leurs complémentarités et leurs antagonismes.

2.2 Définitions

L'excellence opérationnelle (EO). [Girard, 2017] définit l'excellence opérationnelle comme *un élément de la culture d'entreprise visant à favoriser des comportements et à mettre en œuvre une organisation et des outils pour améliorer continuellement la performance d'une organisation*. Dans sa définition, Girard met en exergue cinq (5) notions clés: La culture, les comportements, l'organisation, les outils et l'amélioration continue de la performance. *La culture* implique d'envisager l'excellence opérationnelle comme un objectif à long terme qui touche à l'ADN de l'entreprise. *Les comportements* impliquent que la mise en place de la culture se traduise dans des comportements précis conformes à la grandeur de l'entreprise. *L'organisation* implique que la culture mise en place et les comportements adoptés soient identifiables via une organisation. *Les outils* seront nécessaires pour le déploiement et l'application de cette nouvelle culture d'entreprise. *L'amélioration continue de la performance* est la finalité de l'excellence opérationnelle. Il s'agit, à terme, d'amener continuellement l'entreprise à un haut niveau de performance, de compétitivité. A la définition managériale de

Girard, [Jombart, 2016] donne une définition opérationnelle, complémentaire, plus visuelle. La définition de Jombart met en évidence les impératifs opérationnels d'une organisation: Comblent le client, embarquer les parties prenantes, progresser en permanence, valoriser l'homme et partager la vision. A travers leurs définitions, Girard et Jombart posent le cadre des objectifs que doivent viser une démarche d'excellence opérationnelle efficace dans le temps. Nous nous sommes appuyés sur ces objectifs identifiés pour proposer l'adoption d'un outil d'évaluation opérationnelle à cinq dimensions: le QCDSE (Qualité-Coût-Délai-Sécurité-Environnement) que nous définissons, ci-dessous, dans une section dédiée.

Le Lean Six sigma (LSS). Le Lean Six Sigma est la jonction de deux méthodologies éprouvées: le Lean et le Six Sigma. Il intègre tous les points forts des deux méthodologies et présente tous leurs avantages. Depuis son introduction par [George, 2002], Le LSS s'est répandu dans le monde comme outil pour améliorer les performances des organisations. Le Tableau 1 présente les apports complémentaires du Lean Management et du Six Sigma [Nave, 2002] [Volck, 2009] [Munteanu, 2017] [Pacheco, 2015]. [Rossi et al., 2016].

Tableau 1. Le Lean Six Sigma: Principales singularités.

Méthodologie	Six Sigma	Lean	Lean Six Sigma
Proposé en	1980	1990	2002
Objectif	Stabiliser Réduire la variabilité (satisfaire les exigences client)	Simplifier Supprimer le gaspillage (Améliorer la productivité)	Stabiliser et Simplifier
Outils	DMAIC, DMADV, DFSS (mesures statistiques de la variabilité)	Diagramme d'Ishikawa, 5W, 5S, Kanban, SIPOC, VSM, Poka Yoke (outils visuels)	Jonction des outils du Lean et des outils du Six sigma
Cible	Problème	Flux	Performances
Déploiement	Peu difficile	Facile	Peu difficile
Résultats directs	Uniformisation des processus	Réduction de la durée du flux	Résultats cumulés du Lean et du Six sigma.
Résultats indirectes	Moins de variance, moins de stocks, mesure de la performance (variabilité), amélioration de la qualité.	Moins de gaspillage, débit rapide, moins de stocks, mesure de la performance (flux).	
Limites	Interactions système non pris en compte (amélioration individuel de chaque processus).	Analyse statistique moins valorisée.	Limites cumulés du Lean et du Six sigma.

La Théorie des Contraintes (TC). La théorie des contraintes a été initiée par [Goldratt, 2014]. Dans son ouvrage, il propose une méthodologie visant à traiter les goulots d'étranglements inévitables dans l'exécution de tâches d'une organisation systémique. Goldratt va du principe que la présence de goulots d'étranglement est inévitable dans toute activité. Plutôt que de limiter les contraintes présentes dans une organisation, autant les identifier et exécuter toutes les tâches en fonction de ces contraintes afin de maximiser les performances du système. [Gershon, 2010], [Nave, 2002] et [Pacheco, 2015], dans leurs travaux respectifs, mettent en exergue les éléments essentiels de la TC. Nous nous sommes inspirés des travaux de ces travaux pour vous proposer le Tableau 2.

Tableau 2. La Théorie des Contraintes: Principales singularités.

Méthodologie	Théorie des contraintes
Proposé en	1984
Objectif	Synchroniser Gérer les goulots d'étranglement (Réduire l'impact des contraintes)
Outils	Neuf (9) règles de la TOC (outils statistiques et non statistiques)
Cible	Contrainte
Déploiement	Difficile
Résultats directs	Débit rapide
Résultats indirectes	Moins de dépassement de couts, moins de stocks, mesure de la performance (débit), amélioration de la qualité.
Limites	Méthodologie peu connue, contribution de l'équipe réduite au minimum, analyse de données non valorisée, abstraction d'une partie de l'organisation au profit de la gestion de la contrainte.

L'agilité (A). L'agilité est la méthodologie d'excellence opérationnelle d'actualité, elle est appliquée par un grand nombre d'organisations [Badwe et Erkan, 2018]. En 2001, un groupe de dix-sept passionnés du développement logiciel initie le manifeste agile [Beck et al., 2001]: un ensemble de pratiques de pilotage de projet qui se veut être plus pragmatique que les méthodologies traditionnelles. La méthodologie agile permet à une organisation d'accroître sa rapidité d'exécution, sa flexibilité au changement et de réduire les coûts d'exécution en mettant au centre des préoccupations la satisfaction du client, l'intelligence collective, l'implication quotidienne de l'équipe, l'amélioration continue des pratiques [Badwe et Erkan, 2018] [Morris, 2012] [Munteanu, 2017]. Le tableau 3 énumère les principales caractéristiques de l'agilité, extraites des travaux de [Badwe et Erkan, 2018], [Morris, 2012] et [Munteanu, 2017].

Tableau 3. L'agilité: Principales singularités.

Méthodologie	Agilité
Proposé en	2001
Objectif	Prioriser Ordonner et cadencer le traitement des besoins client (Satisfaire les exigences client, améliorer la productivité)
Outils	Workflow kanban, démarche scrum, project backlog, sprint backlog, sprint (pratiques itératives et incrémentales)
Cible	Système
Déploiement	Facile
Résultats directs	Conception rapide
Résultats indirectes	Transparence, intelligence collective, réduire tout surcharge délai et coût des processus, inspection en continue, améliore la flexibilité.
Limites	Qualité souvent mis à l'écart, peu de place pour la documentation, nécessite la disponibilité du client.

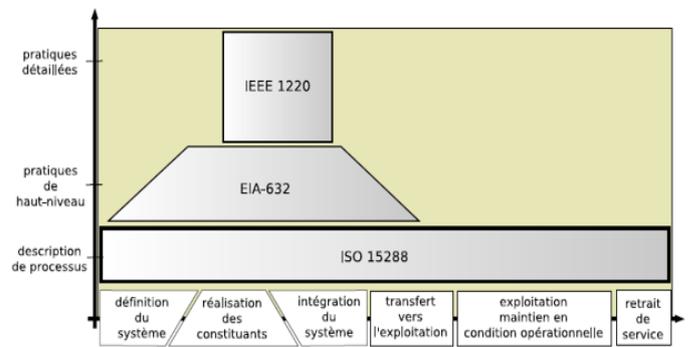


Figure 5. Ingénierie Système: Les normes existantes et définition de leur périmètre par rapport au cycle de vie système.

Dans les grandes organisations, la norme ISO/IEC 15288 d'ingénierie système est toujours associée à des méthodologies complémentaires pour parfaire la démarche d'amélioration continue des performances. [Hehenberger et al., 2016] rappellent que la complexité de démarche d'ingénierie système dépend directement de la complexité du système à concevoir. Autrement dit, plus notre système est complexe, plus ses sous-systèmes sont complexes, plus le nombre et la diversité des interfaces augmentent, plus la démarche d'ingénierie système devra traiter des problématiques d'intégration et d'interaction système. Par effet de causalité, ce niveau de complexité croissant des systèmes nécessitera une évolution de la norme ISO/IEC 15288 ainsi que l'adoption d'une ou de plusieurs nouvelles méthodologies complémentaires. Dans nos travaux futurs, nous souhaitons apporter notre contribution sur une nouvelle méthodologie, complémentaire à la norme ISO/IEC 15288, axée vers l'excellence opérationnelle, permettant de prendre en compte la conception des systèmes actuelles de plus en plus complexes. La figure 6 vient illustrer les niveaux de complexité dont parle [Hehenberger et al., 2016], elle est adaptée d'une illustration présente dans leur article.

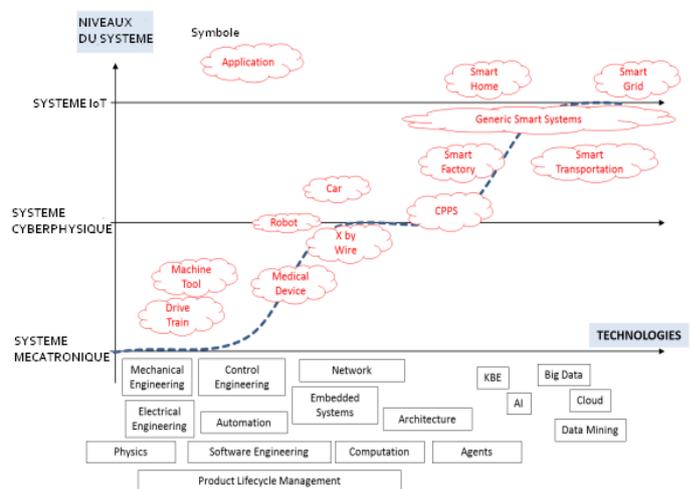


Figure 6. L'ingénierie système: Du système mécanique au système IoT en passant par les systèmes cyberphysiques.

Le QCDSE. L'EO tel que définit managérialement par [Girard, 2017] et, opérationnellement par [Jombart, 2016], vise à améliorer les performances globales d'une organisation. Nous nous sommes interrogées sur les critères pouvant garantir qu'une organisation satisfasse aux définitions de Girard et Jombart. Dans la littérature managériale, le triangle Qualité-Cout-Délai (QCD) est la solution la plus communément utilisée pour évaluer opérationnellement les activités d'une

organisation d'ingénierie. Cependant, il trouve ses limites dans l'absence des principes de responsabilités sociale et sociétale prônés par Jombart. [Gillet-Goinard et Maimi, 2015] propose de passer des critères à trois dimensions (QCD) à des critères à cinq dimensions (QCDSE). Ils mettent en exergue le caractère global des critères d'évaluation Qualité-Coût-Délai-Sécurité-Environnement qui sied à toute organisation de production grâce à l'intégration des principes de responsabilité sociale et sociétale via deux critères supplémentaires: la sécurité et l'environnement. Il nous est alors paru clair que le triangle Qualité-Coût-Délai (QCD) devait évoluer vers une définition plus aboutie. Nous opterons donc pour le QCDSE pour pouvoir évaluer, dans son ensemble, les performances de notre contribution future sur les méthodologies d'excellence opérationnelle. La Figure 7 indique, plus en détail, en quoi consiste chacun des critères du QCDSE. Elle est extraite des travaux de [Gillet-Goinard et Maimi, 2015]. Nous la proposons pour vous indiquer les objectifs visés à travers chacune des lettres du QCDSE.

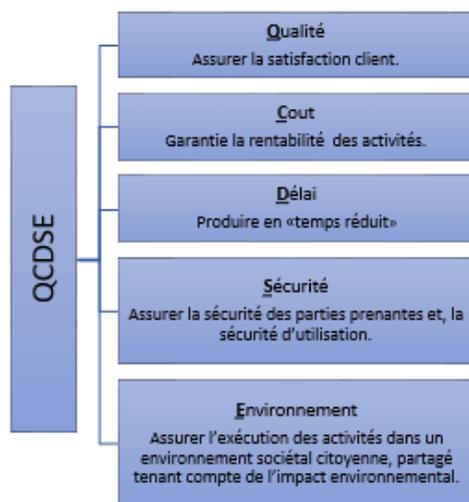


Figure 7. QCDSE: Définition et principes

3 LES SOLUTIONS DE SYNERGIE EXISTANTES

La littérature scientifique et managériale fait état de plusieurs travaux sur des solutions de synergie regroupant les méthodologies actuelles d'EO [Pacheco, 2015] [Hohmann et al., 2016] [O'Rourke, 2005] [Sreedharan et Raju, 2016] [Huang, 2002] [Gubinelli, 2019]. De toutes les solutions de synergie, deux solutions se distinguent pour leur complétude, leur complémentarité, leur faculté de répondre pleinement aux définitions de [Girard, 2017] et [Jombart, 2016]: la TLS et le LAS.

La solution TLS (Théorie des contraintes- Lean- Six sigma). [Hohmann et al., 2016], dans leurs travaux communs, propose une "méta-méthodologie" regroupant la Théorie des contraintes, le Lean et le Six sigma qui emploierait tout ou une partie de chacune de ses méthodologies pour démultiplier leurs effets. Ils énoncent une démarche de déploiement possible en deux (2) étapes. Préalablement, se focaliser sur le goulot dont l'identification puis la caractérisation serait possible avec la théorie de contrainte. Puis, une fois le goulot identifié et caractérisé, se servir des outils et méthodes du Lean ou du Six sigma pour une exploitation efficiente. Toutefois, comme le souligne également [Hohmann et al., 2016], il revient à chaque organisation d'adapter l'ordonnancement de la démarche de déploiement de la TLS en fonction des besoins prioritaires de

l'organisation. Recommandation confirmée par [Pacheco, 2015]. Meyer a démontré l'applicabilité de la TLS pour les entreprises de service [Meyer, 2014] cependant, la solution TLS trouve sa limite à travers la théorie des contraintes, connue pour être une méthodologie d'une complexité scientifique certaine, difficile à déployer dans une organisation.

La solution LAS (Lean Six Sigma- Agilité- Ingénierie Système). [Morris, 2012], [Cesarotti et al., 2019] proposent une solution de synergie Lean Six sigma- Agile- Méthodologie d'ingénierie système (respectivement le CMMI et PMP). Ils énoncent un cadre faisant jonction entre les processus d'ingénierie système (respectivement CMMI et PMP) et la démarche DMAIC du LSS permettant l'exécution des activités d'ingénierie système dans un environnement d'amélioration continue des processus. Pour chaque phase du cycle de vie, ils préconisent l'application des méthodologies Agile afin de profiter des apports de l'agilité. Ceci favorisera l'installation d'une culture de la discipline, essentielle pour assurer le déploiement et l'application de la démarche d'EO. [Ghavami, 2008] fait état de l'applicabilité de la LAS dans l'ingénierie des technologies de l'information. Il met en exergue ses gains en performances conduisant vers une recherche continue de l'excellence. Les travaux de Ghavami sont ceux qui rapprochent le plus de la solution de synergie que nous proposerons.

4 HYPOTHESES POUR UNE BOUCLE D'EXCELLENCE OPERATIONNELLE

Les travaux que nous avons mené dans les sections précédentes, nous ont permis de consolider quatre (4) hypothèses qui constitueront un socle pour proposer une solution menant à l'excellence opérationnelle:

1. **L'approche par processus d'ingénierie système:** Toutes les grandes organisations industrielles utilisent les processus de la norme ISO/IEC 15288 comme méthodologie d'ingénierie système. [Galinié et al., 2017] proposent à toute organisation, quel que soit sa taille ou son secteur d'activité, des outils pour adopter les processus d'ingénierie système. Notre première hypothèse sera donc, de supposer que les organisations utilisent les processus de l'ingénierie système pour concevoir leurs systèmes.

2. **La modélisation de processus pour l'amélioration continue des performances:** Dans son ouvrage, [Gramdi, 2013] propose un "méta-concept" analogue à l'approche processus. Il estime que chaque processus peut être modélisé sous forme d'un tuyau dont l'optimisation des performances serait possible en jouant sur ces caractéristiques: sa longueur, sa section, son épaisseur, sa porosité. La proposition de Gramdi a retenu notre attention pour son approche, à priori simpliste, mais qui s'avère global car systémique, fractale, adaptative, évolutive. Les méthodologies actuelles d'excellence opérationnelle permettent d'agir sur chacune des caractéristiques du tube: l'A permet réduire la longueur du tube, le LSS permet réduire la section et la longueur du tube (Tableau 4). De ce fait, notre seconde hypothèse sera, de prioriser les méthodologies actuelles qui permettent de réduire continuellement la longueur, l'épaisseur, la section, la porosité d'un tuyau afin de garantir d'améliorer continuellement les performances des processus.

3. **L'excellence opérationnelle selon Girard et Jombart:** La solution de synergie que nous proposerons vise à l'excellence opérationnelle telle que définit opérationnellement et managérialement par [Girard, 2017] et [Jombart, 2016].

Pour cela, toute organisation devra être évaluée sur des critères à cinq dimensions: le QCDSE. Notre troisième hypothèse sera donc de proposer une solution de synergie nous permettant d'être conforme aux critères QCDSE.

4. **La vitesse de déploiement:** L'un des enseignements de nos expériences dans l'ingénierie système est que l'adhésion des parties prenantes (corps managérial et corps opérationnel), la simplicité et l'efficacité d'une solution sont des facteurs essentiels pour garantir l'adoption puis le succès d'une démarche d'excellence opérationnelle dans une organisation. Notre quatrième hypothèse sera donc, par anticipation, de privilégier une solution de synergie avec un déploiement facile voire peu difficile dont, les résultats seront disponibles rapidement, périodiquement, à portée de tous.

Les quatre (4) hypothèses énoncées ci-dessus, associées à la solution LAS, constitueront notre socle pour vous faire notre proposition. Le tableau 4, ci-dessous, met en exergue la complétude et la complémentarité de la solution LAS. Elle a été établie grâce aux travaux de [Morris, 2012], [Munteanu, 2017] et [Ghavami, 2008].

Tableau 4. La solution LAS: le LSS associé à l'A.

Méthodologie	Lean Six Sigma	Agilité
Proposé en	2002	2001
Objectif	Stabiliser et Simplifier Réduire la variabilité (satisfaire les exigences client) Supprimer le gaspillage (Améliorer la productivité)	Prioriser Ordonnancer et cadencer le traitement des besoins client (Satisfaire les exigences client, améliorer la productivité)
Outils	DMAIC, DMADV, DFSS (mesures statistiques de la variabilité) Diagramme d'Ishikawa, 5W, 5S, Kanban, SIPOC, VSM, Poka Yoke (outils visuels)	Workflow kanban, démarche scrum, project backlog, sprint backlog, sprint (pratiques itératives et incrémentales)
Cible	Performances	Système
Déploiement	Peu difficile	Facile
Résultats directs	Uniformisation des processus et Réduction de la durée du flux.	Conception rapide
Résultats indirectes	Moins de variance, moins de gaspillage, moins de stocks, mesure de la performance (variabilité, flux), amélioration de la qualité.	Transparence, intelligence collective, réduire tout surcharge délai et coût des processus, inspection en continue, améliore la flexibilité.
Limites	Limites cumulés du Lean et du Six sigma.	Qualité souvent mis à l'écart, peu de place pour la documentation, nécessite la disponibilité du client.

Dans nos futurs travaux, nous détaillerons notre solution d'excellence opérationnelle dont l'amélioration continue sera rendue possible via une boucle de régulation.

5 CONCLUSION ET PERSPECTIVES

Dans cet article, nous avons présenté un état de l'art des méthodologies actuelles d'excellence opérationnelle. Nous avons commencé, en premier lieu, par mettre en évidence leur importance croissante via une étude bibliographique. Ensuite, nous avons donné la définition ainsi que les singularités de chaque des méthodes actuelles d'excellence opérationnelle. Nous avons pris le temps de situer nos travaux, de détailler nos objectifs à travers les définitions de Girard et Jombard. En second lieu, nous avons présenté les deux solutions de synergie proposées par la littérature managériale et scientifique en adéquation avec nos objectifs. De ces travaux découlent quatre (4) hypothèses, la solution de synergie LAS puis, les critères d'évaluation à cinq dimensions QCDSE qui nous serviront de socle, dans nos futurs travaux, pour proposer une solution d'excellence opérationnelle globale destinée à l'ingénierie système: de l'ingénierie des systèmes embarqués à l'ingénierie des systèmes d'internet des objets [Hehenberger et al., 2016] via *La boucle de l'excellence opérationnelle*.

Nous avons opté pour une solution par boucle d'intégrations (outils et méthodologie) afin de rendre possible son adaptabilité en fonction de la complexité du système à concevoir (des systèmes embarqués aux systèmes IoT), du domaine d'application concerné (produit industriel ou service) et de la pondération souhaitée de chacun des critères d'évaluation QCDSE. Nous détaillerons notre proposition dans un prochain article.

6 REMERCIEMENTS

Cet article a été rédigé avec la sollicitude du service de Formation Continue de l'Université de Technologie de Compiègne. Nous tenons à remercier toutes les personnes qui, de près ou de loin, nous ont permis de proposer cet article.

7 REFERENCES

- Badwe, S. & Erkan T.E. (2018). A Taxonomy of Lean Six Sigma and Agile Methodologies used in Software Development. *International Journal of Engineering Research and Technology*. ISSN 0974-3154 Volume 11, Number 5 (2018), 725-753.
- Gershon, M. (2010). Choosing Which Process Improvement Methodology to Implement. *Journal of Applied Business & Economics* (Vol. 10), Issue 5, 61.
- Girard, M. (2017). *L'Excellence opérationnelle-L'arme secrète des Entreprises à Succès*. Printed in Great Britain by Amazon, ISBN 9781521056646.
- Jombard, P. (2016). *L'Excellence opérationnelle- Piloter l'entreprise 5.0*. Edition l'Archipel, ISBN 978-2-8098-1887-1.
- Rossi, M., Cattaneo, L., Duigou, J., Terzi, S., Fugier-Garrel, S., Eynard, B. (2016). Lean Product Development and the Role of PLM. *13th IFIP International Conference on Product Lifecycle Management (PLM)*, 183-192.
- George, M. (2002). *Lean Six Sigma: Combining Six Sigma Quality with Lean Production Speed*. McGraw-Hill Professional, ISBN 9780071385213.
- Nave, D. (2002). How To Compare Lean and the Theory of Constraints, *Quality Progress* (Vol. 35), No.3, 73-9.
- Volck, N. (2009). *Déployer et exploiter Lean six sigma*. Editions d'Organisation- Groupe Eyrolles, ISBN 978-2-212-54334-6.

- Munteanu, A. (2017). Comparative Analysis between Lean, Six Sigma and Lean Six Sigma Concepts. *Management and Economics Review* (Vol. 2), Issue 1.
- Goldratt, E (2014). *The Goal: A Process of Ongoing Improvement*, 3rd Revised Edition. North River Press, ISBN 9780884271956.
- Beck, K., Grenning, J., Martin, R., Beedle, M., Highsmith, J., Mellor, S., Van Bennekum, A., Hunt, A., Schwaber, K., Cockburn, A., Jeffries, R., Sutherland, J., Cunningham, W., Kern, J., Thomas, D., Fowler, M., Marick, B. (2001), <https://agilemanifesto.org/iso/fr/manifesto.html>.
- Morris, P. (2012). The perfect process storm: Integration of CMMI, agile, and lean six sigma. *CrossTalk*, 39-45.
- Galinier, M., Galinier, S., Laporte, C (2017). *Les clés de l'ingénierie des systèmes*. Cépaduès-Éditions, ISBN 978.2.36493.597.6.
- Luzeaux, D., Ruault, J-R. (2013). *L'ingénierie système, 100 questions pour comprendre et agir*. AFNOR Editions, ISBN 978-2-12-465406-2.
- INCOSE. (2015). *INCOSE systems engineering handbook: a guide for system Life cycle processes and activities*. John Wiley & Sons.
- ISO/IEC 15288. (2002). *Systems engineering-system life cycle processes*. International Standardization Organization/International Electrotechnical Commission.
- Hohmann, C. (2009). *Techniques de productivité: Comment gagner des points de performance*. Editions d'Organisation- Groupe Eyrolles, ISBN 2-212-54295-X.
- Gillet-Goinard, F., Maimi, L. (2015). *Toute la fonction production*, 2e édition. Dunod, ISBN 2100723626.
- Hohmann, C. (2016) <http://christian.hohmann.free.fr/index.php/synergie-tls-toc-lean-six-sigma>.
- Pacheco, D. (2015). TOC, lean and six sigma: The missing link to increase productivity? *African Journal of Business Management*, Vol 9 (12), 513-520.
- O'Rourke, P. M. (2005). A Multiple-Case analysis of lean six sigma deployment and implementation strategies. *Theses and Dissertations*. 3765, <https://scholar.afil.edu/etd/3765>.
- Sreedharan, R. V., Raju, R. (2016). A systematic literature review of Lean Six Sigma in different industries. *International Journal of Lean Six Sigma*, Volume 7, N°4, 430-466.
- Huang, H. (2002). Integrated production model in agile manufacturing systems. *International Journal of Advanced Manufacturing Technology*, 20, 515-525.
- Meyer, F. (2014). *Appliquer le ToC Lean Six Sigma dans les services*. Lexitis Editions, ISBN 978-2-36233-129-9.
- Gramdi, J. (2013). *La Boucle Vertueuse de l'Excellence-Comment mettre harmonieusement en synergie le lean management, le six sigma et la théorie des contraintes pour enfin de vrais résultats en rupture*. Lexitis Editions, ISBN 978-2-36233-115-2.
- Gubinelli, S., Cesarotti, V., Introna, V. (2019). The evolution of Project Management (PM): How Agile, Lean and Six Sigma are changing PM. *The Journal of Modern Project Management* (Vol. 7), N°3.
- Ghavami, P. (2008). *Lean, Agile and Six Sigma Information Technology Management: New Stratagems to Achieve Perfection*. CreateSpace Independent Publishing Platform, ISBN: 978-1440478123.
- Hehenberger, P., Vogel-Heuser, B., Bradley, D., Eynard, B., Tomiyama, T., Achiche, S. (2016). Design, Modelling, Simulation and Integration of Cyber Physical Systems: Methods and Applications. *Computers in Industry* (Vol. 82), 273-289.

L'ingénierie dirigée par les modèles pour assurer la sécurité des logiciels embarqués en automobile.

Proposition méthodologique et étude de cas.

YANDIKA SIRGABSOU^{1,2,4,5}, CLAUDE BARON^{2,3,4,5}, LORENZO GRENIER^{4,5}, LAURENT PAHUN¹, PHILIPPE ESTEBAN^{2,4,6}

¹ RENAULT SOFTWARE FACTORY
Toulouse, France
laurent.pahun@renault.com

³ ISAE-SUPAERO
Toulouse, France
claude.baron@laas.fr

⁵ INSA-TOULOUSE
Toulouse, France
grenier@etud.insa-toulouse.fr

² LAAS-CNRS, CNRS
Toulouse, France
ysirgabsou@laas.fr

⁴ UNIVERSITE DE TOULOUSE
Toulouse, France

⁶ UNIVERSITE TOULOUSE 3
Toulouse, France
esteban@laas.fr

Résumé – L'avènement des véhicules autonomes a été longtemps perçu comme un événement qui allait révolutionner le monde du transport, améliorer le confort et la sécurité des usagers. Bien que cet optimisme soit partagé par la plupart et que d'importants progrès technologiques aient été fait dans ce sens, le chemin ne reste pas pour le moins jonché de nombreux défis, notamment celui de la sécurité, liée à l'utilisation croissante de logiciels embarqués devenus très complexes. Face à cette complexité grandissante, les méthodes manuelles traditionnelles d'analyse de la sécurité basées sur des documents ont atteint leur limite et il est temps de repenser les approches. Nous proposons pour cela de nous appuyer sur l'ingénierie dirigée par les modèles pour mener les analyses de sécurité non plus directement à partir des documents de conception mais d'un modèle de l'architecture du logiciel embarqué. Nous faisons en ce sens une proposition méthodologique, qui, validée sur l'étude d'un composant logiciel automobile, améliore les pratiques actuelles en temps, en qualité des analyses et en réutilisation.

Abstract – The advent of autonomous vehicles was long seen as an event that would revolutionize the world of transportation and improve the comfort and safety of users. Although this optimism is shared by most people and significant technological progress has been made in this direction, the road ahead offers many challenges, particularly in the area of safety, due to the increasing use of embedded software that has become very complex. Faced with this growing complexity, traditional manual, document-based methods of safety analysis have reached their limits and it is time to rethink approaches. We propose to rely on model-driven engineering to conduct safety analyses no longer directly from design documents but from a model of the embedded software architecture. To this goal, we make a methodological proposal, which, validated on the study of an automotive software component, improves current practices in time, analysis quality and reuse.

Mots clés – ISO 26262, Logiciel Embarqué, MBSA, IDM, Sécurité

Keywords – ISO 26262, Embedded Software, MBSA, MBSE, Safety

1 INTRODUCTION

Avec l'avènement des véhicules autonomes, l'industrie automobile a de plus en plus recours aux logiciels embarqués pour assurer diverses fonctionnalités. Cela n'est pas sans conséquence pour la sécurité (au sens de safety, capacité du système à se prévenir de défaillances pouvant causer blessures et dommages) (DoD, 2012), qui doit être assurée avec rigueur. Être en mesure de fournir des garanties de la sécurité sera un facteur déterminant pour l'acceptation des véhicules autonomes par la société. Chez Renault, les pratiques en matière d'ingénierie logicielle sont bien encadrées par divers standards dont les principaux sont l'ASPICE (VDA QMC, 2015) qui définit l'utilisation et l'évaluation des processus d'ingénierie, et l'ISO 26262 (ISO 26262-1, 2018) qui adresse les aspects liés à la sécurité dans le développement du système et des logiciels. Le processus de développement suit un cycle en V classique, qui couvre l'élicitation des exigences, l'analyse, la conception d'architecture, jusqu'à l'implémentation, sur la partie dite descendante, ainsi que les tests et l'intégration, sur la partie remontante.

Conformément aux normes citées, la traçabilité verticale (des exigences vers les composants les implémentant) et horizontale (des tests vers les exigences) doivent être maintenues entre les artefacts produits aux différentes étapes du développement. Or,

celles-ci reposent principalement sur des documents textuels, même s'ils sont supportés par des outils comme Excel ou DOORS. Avec la complexité croissante des logiciels, et l'augmentation du volume de documents associés, il n'est plus possible aujourd'hui de garantir la qualité, de maintenir la traçabilité et être en conformité avec les standards en utilisant les méthodes actuelles.

Par ailleurs, la norme ISO 26262 (ISO 26262-1, 2018) requiert déjà d'effectuer des analyses de sécurité et d'en apporter des preuves. Or, la complexité actuelle des applications embarquées dans un véhicule automobile rend extrêmement difficile d'assurer la safety. De plus, la complexité des architectures automobiles nécessite aujourd'hui un travail fortement collaboratif entre plusieurs équipes, constituées de spécialistes de domaines métiers différents, qui ont besoin d'utiliser de méthodes et de moyens efficaces pour travailler ensemble. En outre, il devient très difficile de maintenir à jour les analyses de sécurité par rapport aux évolutions, parfois rapides en cas de développement agile, des artefacts d'ingénierie.

Face à cette utilisation croissante de logiciel dans les véhicules, à la difficulté d'en évaluer et garantir la safety, et au vu du besoin accru d'échanger, au sein des équipes et entre équipes, des données d'ingénierie consolidées, simulables, vérifiables, il est aujourd'hui nécessaire de changer de paradigme, et de passer

des approches basées-documents à des méthodes basées-modèles pour les activités de développement de systèmes et de logiciels, ainsi que celles d'évaluation de leur sécurité. Ces approches favorisent la communication entre experts et équipes, et améliorent la capitalisation à travers la réutilisation. N'échappant pas à cette tendance, l'industrie des systèmes embarqués critiques, incluant l'automobile, considère également avec beaucoup d'intérêt ces approches basées modèles.

La proposition, dans cet article, est de nous appuyer sur l'ingénierie dirigée par les modèles pour définir une méthode pour l'évaluation de la safety des architectures logicielles en automobile, et de faire une proposition d'outillage. Les travaux présentés dans cet article résultent de notre retour d'expérience sur des études de cas menées dans l'entreprise Renault Software Labs dans l'objectif de déployer et de faire évaluer plus largement la méthodologie et son outillage par les experts de l'entreprise.

En section 2, l'article commence par explorer les méthodes basées sur l'ingénierie dirigée par les modèles (IDM) en conception système et logicielle, ainsi que pour l'évaluation de la safety. La Section 3 fait ensuite une proposition méthodologique visant à aider les experts safety à construire, à partir des artefacts d'ingénierie, des modèles que l'on appelle dysfonctionnels, à partir desquels mener des analyses de safety; celle-ci étend une première proposition, décrite dans (Sirgabsou et al., 2020). La Section 4 applique pas à pas la méthodologie sur l'une des études de cas, le contrôle longitudinal, une application embarquée d'aide à la gestion de la vitesse du véhicule compte tenu de la signalisation et des conditions de circulation. La Section 5 discute les résultats obtenus. L'article conclut et indique comment il serait intéressant de faire évoluer l'outillage de la méthodologie compte tenu des limitations observées, ainsi que des options alternatives d'outillage.

2 DEVELOPPEMENT DES APPROCHES BASEES MODELES

On observe depuis quelques temps cette évolution des pratiques industrielles vers des approches basées modèles en ingénierie système et logicielle ; on parle de MBSSE (Model Based System and Safety Engineering). Il s'agit d'un ensemble de pratiques fondées sur le concept de modèle conceptuel de domaine qui ont entre autres pour but d'automatiser la production de logiciels (NDIA, 2011). Le processus de conception peut alors être vu comme un ensemble ordonné de transformations de modèles, jusqu'à obtention d'artefacts exploitables, ce qui favorise notamment réutilisation et les vérifications précoces. Ce processus consiste à tisser différents modèles entre eux (fonctionnel, organique, ou adressant différentes exigences extra-fonctionnelles à garantir, comme la sécurité, la fiabilité, la performance ...). Plus récemment, et toujours à l'état de recherche, on considère l'adoption des approches basées modèles pour mener des analyses de safety, pour le moment limitées au niveau système ; on parle de MBSA (Model Based Safety Assessment).

2.1 Model Based System and Software Engineering

Avant d'être adoptées en ingénierie système, les approches basées modèles sont apparues en ingénierie du logiciel (Estefan, 2008). Elles permettent d'assurer une certaine continuité entre les différentes étapes du développement de systèmes et de logiciels (modélisation des exigences, de l'architecture logique, de l'architecture physique, jusqu'à l'implémentation), en assurant une traçabilité durant les transitions entre modèles. Elles reposent souvent sur l'utilisation de langages et d'outils.

Ainsi, les langages UML (Unified Modeling Language) et SysML (Systems Modeling Language) sont couramment utilisés pour modéliser les architectures fonctionnelles et organiques, et le comportement des logiciels et systèmes. Les diagrammes structuraux (diagramme de classe, de paquetage) sont par exemple utilisés pour modéliser l'architecture organique ; les diagrammes de cas d'utilisation, de séquence, d'activité, sont utilisés pour modéliser des scénarii comportementaux.

Il existe également des langages de description d'architecture, qui supportent spécifiquement l'étape de conception. Citons par exemple le langage AADL (Architecture Analysis and Description Language), un standard de la SAE (Society of Automotive Engineers) initialement conçu pour l'avionique (Feiler et al., 2003), qui permet de concevoir et d'analyser l'architecture des systèmes embarqués. Certains langages de description d'architecture sont spécifiques à un domaine. En automobile, c'est par exemple le cas de EAST-ADL (Cuenot et al., 2007), un méta-modèle qui permet de modéliser l'environnement du système, le système lui-même à 5 différents niveaux de détails, et qui offre des extensions permettant de modéliser les propriétés du système à chaque niveau.

Toujours en conception, des sémantiques plus proches du code telles que celles des modèles SIMULINK et SCADE sont utilisées pour le prototypage rapide des architectures détaillées du logiciel et pour la génération de code. SCADE, beaucoup utilisé en avionique (Colaco et al., 2017), permet de modéliser des systèmes synchrones tels que ceux de commandes de vol ayant des contraintes de temps réel et ainsi d'anticiper certaines problématiques relatives à la safety propres aux systèmes concurrents ; il permet également l'automatisation des tests. Dans d'autres secteurs, comme c'est le cas pour l'automobile, ce sont les modèles Simulink qui sont de plus en plus utilisés.

Il apparait donc que différents modèles/langages sont utilisés à différentes étapes du cycle de développement d'un système ou d'un logiciel, selon le point de vue que l'on souhaite représenter. Une conception supportée par des modèles offre plusieurs avantages, tels qu'une meilleure communication, efficacité et réutilisation. La difficulté reste d'assurer une continuité entre ces différents artefacts de conception, par des transformations de modèles, dans un objectif de maintien de la cohérence globale. Une autre difficulté, qui devient majeure dans le contexte actuel de systèmes critiques de plus en plus autonomes, consiste à savoir tisser ces modèles de conception avec des modèles d'analyses spécifiques, comme les analyses de safety.

2.2 Model-Based Safety Assesment

Comme le MBSE, le MBSA résulte de la tendance en ingénierie de l'abandon des approches basées-documents au profit de celles basées-modèles. Il ne vient pas remplacer les modèles traditionnels sur lesquels reposent les analyses de la safety (telles que les arbres de fautes (Vesely et al., 1981), coupes minimales (Center for Chemical Process Safety, 2010) et les AMDE(C) (Analyse des Modes de Défaillance et de leurs Effets (Critiques) (SAE, 2015), mais vise plutôt à établir des ponts entre ces modèles et les modèles d'ingénierie (résultant du MBSE), favorisant ainsi une meilleure cohérence et traçabilité entre modèles, ainsi qu'une justesse des analyses. Une autre motivation du MBSA est également d'unifier les modèles classiques d'analyses de la safety, considérée au sens large (incluant fiabilité et maintenabilité) dans un seul modèle dysfonctionnel (Lisagor et al., 2011).

Plusieurs approches de MBSA existent, telles que le FPTN (Fenelon & McDermid, 1993), HiP-HOPS (Papadopoulos &

McDermid, 1999), FSAP/NuSVM (Bozzano & Villafiorita, 2007) ou encore AltaRica (Point & Rauzy, 1999). Sans rentrer dans les détails, ces approches se distinguent principalement par la façon dont le modèle dysfonctionnel est construit : celui-ci peut être étendu à partir du modèle fonctionnel (dit nominal) issu de la conception, ou être distinct (Lisagor et al., 2011). Dans le cas d'un modèle étendu, le modèle du système sera enrichi par des informations de défaillance ; un modèle dédié sera construit manuellement par l'expert safety.

Les approches offrent chacune des avantages et des inconvénients. Utiliser un modèle étendu permet de garantir naturellement la cohérence entre le modèle du système et le modèle de safety dérivé, sans avoir besoin de recourir à des mécanismes supplémentaires. De plus, le concepteur et l'expert safety peuvent utiliser le même environnement et outil de modélisation. Cependant, si l'idéal est d'avoir ainsi un seul modèle capable d'intégrer les éléments fonctionnels et dysfonctionnels, suffisamment structuré pour en déduire les analyses de safety, en pratique, les sémantiques des modèles de conception sont souvent encore peu formelles et peu structurées. Une alternative satisfaisante consiste alors dans ce cas à utiliser un modèle dédié. L'un des intérêts de cette approche est qu'elle permet d'avoir une indépendance entre les activités de conception et d'analyse de la safety, ce qui peut constituer un atout important, par exemple dans un contexte de certification. Cependant, utiliser un modèle dédié pose la question du temps conséquent nécessaire pour construire ce modèle. Plusieurs travaux portent sur des passerelles de construction d'un modèle MBSA à partir d'un modèle MBSE (David et al., 2009), (Nguyen et al., 2018) et souligne l'effort à faire pour établir celles-ci. D'autres travaux tels que ceux de (Batteux et al., 2019) au (Legendre et al., 2016) s'intéressent au maintien de la cohérence des modèles MBSE et MBSA à travers la synchronisation.

Parmi les approches mentionnées plus tôt, nous nous intéressons à AltaRica (Point & Rauzy, 1999), qui aujourd'hui est une approche pionnière dans l'écosystème des systèmes critiques Européen. AltaRica est un langage de haut niveau conçu pour la modélisation des systèmes. Un modèle AltaRica décrit un système à travers des domaines, des variables et une hiérarchie de nœuds où chaque composant peut incorporer plusieurs sous-nœuds. Un domaine décrit un ensemble de valeurs booléenne, entières, énumérées ou abstraites, qu'une variable peut prendre. Les nœuds décrivent les comportements des composants du système à travers des variables d'état, des événements, des transitions et des assertions.

Les transitions décrivent les changements d'état. Un nœud interagit avec l'environnement de deux manières : 1) à travers des événements et 2) à travers des variables de flux. À l'exemple d'une fonction de transfert, les assertions décrivent les relations entre les flux d'entrée, les variables d'état et les flux de sortie. L'intérêt d'AltaRica, réside surtout dans sa sémantique à la fois formelle et proche des systèmes qu'elle décrit. Cette dualité, difficile à obtenir avec les autres approches, justifie le choix d'AltaRica pour notre démarche.

3 PROPOSITION METHODOLOGIQUE

Notre ambition et l'originalité de notre contribution est de proposer une méthodologie adaptant les concepts, principes et méthodes du MBSA, généralement appliqués au niveau d'un système, afin d'améliorer les pratiques d'analyse de la sécurité des logiciels. Cette proposition est notamment importante dans le contexte des logiciels embarqués en automobile, qui sont critiques du point de vue de la sécurité.

Parmi les approches possibles, nous avons choisi dans cet article d'explorer celle d'un modèle dysfonctionnel dédié, basé sur une méthode de modélisation des logiques de propagation de défaillance et utilisant le langage AltaRica.

Comme spécifié plus tôt, cette méthodologie étend et adapte une première proposition, décrite dans (Sirgabsou et al., 2020), qui permettait de construire un modèle dysfonctionnel de l'architecture logicielle et de l'utiliser comme base pour conduire les analyses safety en trois étapes, modélisation dysfonctionnelle, traduction des logiques et analyse de safety. La nouvelle proposition introduit une étape préliminaire de choix des données d'entrée, qui adresse spécifiquement la difficulté d'avoir de bonnes entrées pour la construction du modèle. La méthodologie procède donc en trois nouvelles étapes, comme illustré en Figure 1: 1) le choix des données d'entrée à partir desquelles sera construit le modèle dysfonctionnel, 2) la construction du modèle dysfonctionnel du logiciel, à partir duquel 3) les coupes minimales, les arbres de fautes et les AMDE pourront être générés automatiquement.

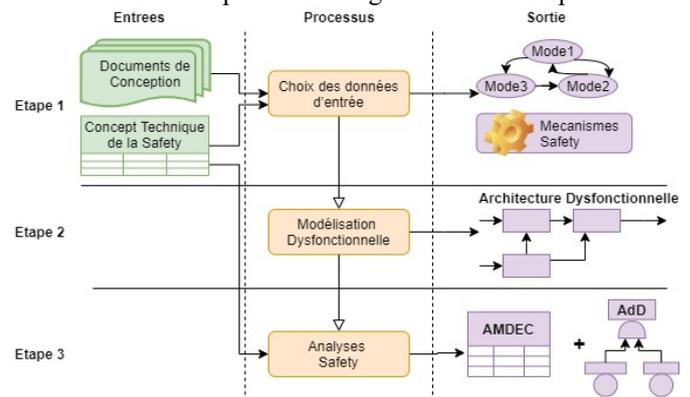


Figure 1. Les 3 grandes étapes de la méthode

3.1 Étape 1 : Choix des données d'entrée

Ayant opté pour une approche reposant sur un modèle dédié, nous devons tout d'abord identifier les informations nécessaires pour sa construction : nous partons des documents de conception (architecture) et d'un 'Technical Safety Concept' (TSC). Le TSC, défini dans l'ISO 26262 (ISO 26262-1, 2018), est un document qui contient les spécifications des exigences de safety issues du système tels que les safety goals (exigences de safety de haut niveau résultant de l'analyse préliminaire des risques au niveau véhicule) (cf. ISO 26262-1 3.139), leur degré de criticité ou ASIL (Automotive Safety Integrity Level) ainsi que leurs allocations aux composants matériels ou logiciels.

Le TSC contient aussi des informations qui précisent quelles mesures (actions préventives) doivent être prises durant le développement et quels mécanismes de sécurité (actions curatives) doivent être implémentés dans l'architecture. Ainsi, grâce au TSC nous pouvons identifier les composants et interfaces à modéliser ainsi que les exigences et mécanismes à évaluer dans le cadre de l'architecture dysfonctionnelle, et ainsi éviter de surcharger le modèle dysfonctionnel avec des éléments inutiles à la safety.

A partir des documents d'architecture, nous identifions le périmètre du système, ses interactions, ses modes de fonctionnement (états, normaux ou anormaux) et les conditions de changement d'état de ses composants ; nous pouvons ainsi avoir une compréhension de son fonctionnement. La démarche suivie à l'étape 1 est illustrée sur la Figure 2

Cette compréhension peut être complétée, si nécessaire, en consultant le concepteur (architecte logiciel) ou encore en exploitant des modèles de design plus détaillés s'ils existent. Par exemple, des modèles fonctionnels Simulink peuvent être

exploités pour une compréhension de la logique fonctionnelle et pour déterminer comment les mécanismes de safety sont implémentés.

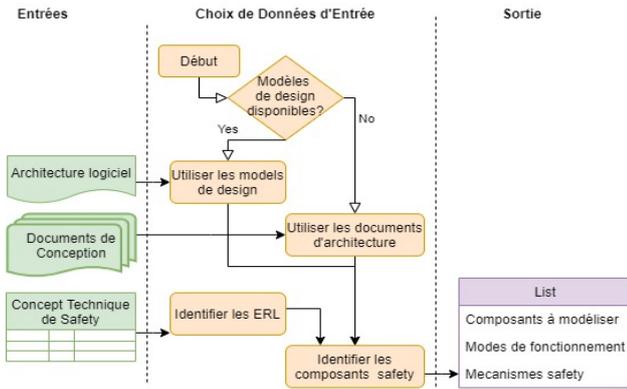


Figure 2. Étape 1, Choix des données d'entrée

3.2 Étape 2 : Modélisation dysfonctionnelle

Elle consiste en la modélisation du comportement dysfonctionnel (en présence de défaillance) des composants logiciels ainsi que leurs interactions. Comme illustrée sur la Figure 3, elle se déroule en 2 sous-étapes : 1) la modélisation par des machines à états du comportement des composants, et 2) la modélisation des propagations des défaillances. Nous les décrirons dans les 2 sous-sections suivantes.

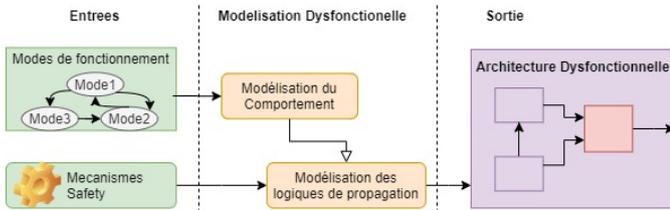


Figure 3. Étape 2, Modélisation dysfonctionnelle

3.2.1 Modélisation du comportement des composants

Elle consiste en la modélisation du comportement dysfonctionnel interne des composants du système grâce aux données collectées à l'étape 1 (modes de fonctionnement, mécanismes de safety). Elle se focalise sur les composants et signaux en lien avec la safety. Nous modélisons le comportement interne des composants grâce à des machines à états (états et transitions). Les états ne sont rien d'autre que les modes de fonctionnement (que nous avons identifiés), qu'ils soient nominaux ou défaillants (ON, OFF, défaillant, dégradé...). Au travers des transitions, nous exprimons les conditions logiques conditionnant le passage d'un état à un autre. Ces conditions peuvent être liées à des événements aléatoires et guidées par une loi de probabilité (stochastique), ou dépendantes de l'état des entrées (déterministes). De cette modélisation, il résulte des composants logiciels ramenés à des machines d'états.

3.2.2 Modélisation des interfaces

La modélisation des interfaces complète celle du comportement des composants avec des logiques de propagation de défaillance. Il s'agit de modéliser les interactions dysfonctionnelles entre les composants (comment les défaillances se propagent d'un composant à un autre). Dans un premier temps, nous modélisons les signaux d'entrée et de sortie de chaque composant en leur attribuant des états discrets représentant les classes de valeurs qu'ils ne peuvent pas prendre. Par exemple, d'un point de vue dysfonctionnel, une donnée d'entrée peut être correcte, erronée, en retard ou absente (voir l'annexe E de la partie 6 de ISO 26262, qui liste les erreurs

possibles relatives à l'exécution et à l'échange d'information entre composants logiciels).

Dans un deuxième temps, nous écrivons des logiques nous permettant de déduire les états de sortie en fonction de l'état interne du composant et de ses entrées. Par exemple, prenons une simple fonction ADD qui additionne deux variables (entree1 et entree2). Pour cette fonction nous définissons une variable d'état 'state' pouvant prendre trois valeurs : ok, erronée ou perdue. Pour les 2 entrées et la sortie, nous définissons également 3 états : correcte, erronée, absente). Nous pouvons alors écrire l'expression de la sortie en fonction des entrées et de l'état de la fonction :

```

if (state = ok) then
  if ((entree1 = correcte) and (entree2 = correcte)) then
    sortie : = correcte
  Else if ((entree1 = absente) or (entree2 = absente)) then
    sortie : = absent
  else
    sortie : = erronée
  end if
else //state not ok
  sortie : = erronée
end if
  
```

Il existe aujourd'hui des formalismes supportant une telle approche, comme AltaRica ou celui de l'annexe d'erreur d'AADL.

En réitérant cette démarche pour tous les composants, nous aboutissons à une architecture dysfonctionnelle bien définie, dans une sémantique formelle; ceci constitue une base nécessaire pour la génération des analyses de safety.

3.3 Étape 3 : Analyses de Safety

Le but de l'étape 3 est de dériver des modèles de safety classiques à partir du modèle dysfonctionnel construit. Pour cela, il est avant tout nécessaire d'ajouter des Événements Redoutés Logiciels (ERL) au modèle dysfonctionnel. Pour ce faire, il est nécessaire d'exprimer les ERL par des combinaisons booléennes et de les associer aux composants concernés. Après cela, il sera alors possible d'utiliser le modèle résultant pour produire les AMDE et arbres de fautes pour faire des analyses grâce à la vérification de modèle. Cela est rendu possible au travers de l'ensemble du comportement dysfonctionnel capturé grâce aux modes de défaillances individuels ainsi qu'aux événements associés modélisés au niveau de chaque composant. De plus, l'ajout des événement redoutés à l'architecture permet, au travers de la logique de propagation modélisée, de les associer ou non à certains composants. Ainsi plusieurs ERL peuvent être ajoutés, évalués sur la base de la même architecture dysfonctionnelle.

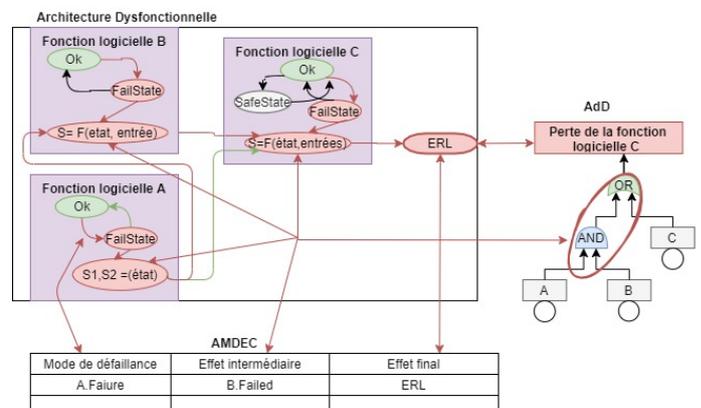


Figure 4. Étape 3 : Génération des AMDE et arbres des fautes à partir de l'architecture dysfonctionnelle

Une illustration faisant apparaître les correspondances entre les éléments du modèle dysfonctionnel et ceux des AMDE et arbre des fautes est présentée sur la Figure 4. Sur cette base, on peut donc générer, à partir de l'architecture dysfonctionnelle, des arbres des fautes ayant comme sommet l'un des événements redoutés modélisés, et dont les branches découlent (en fonction de la logique de propagation) des transitions vers des états défaillants tels que modélisés individuellement par les machines à états au niveau des composants.

Cette même logique peut s'appliquer aux AMDE, où les transitions vers des états défaillants deviendront des modes de défaillances, alors que les événements redoutés deviendront des effets finaux.

4 ÉTUDE DE CAS

Nous proposons d'outiller la méthode décrite précédemment avec le logiciel SimfiaNeo et de l'appliquer sur une étude de cas, l'analyse d'un composant logiciel assurant le contrôle longitudinal d'un véhicule. Elle est complémentaire à l'étude sur le contrôle latéral, décrite dans (Sirgabsou et al., 2020). L'objectif est d'évaluer l'architecture du logiciel, grâce à un modèle dysfonctionnel, pour vérifier si elle satisfait un certain nombre d'exigences de safety. Nous évaluerons par la même occasion les performances de l'outil SimfiaNeo. Le périmètre de la nouvelle étude inclut un safety concept incorporant des modèles informels expliquant les mécanismes safety proposés. Ceci nous permettra de simplifier le processus de modélisation de l'architecture dysfonctionnelle.

4.1 Présentation de SimfiaNeo

SimfiaNeo est un outil de MBSA basé sur le langage AltaRica développé par APSYS-Airbus. Il offre une interface de modélisation graphique basé sur Eclipse qui est plus conviviale comparée aux autres éditeurs AltaRica présents sur le marché tels que OCAS, Open AltaRica ou AltaRica Studio. SimfiaNeo permet de construire graphiquement le modèle AltaRica d'un système et de générer directement des coupes, des arbres de fautes et des AMDE à partir du modèle AltaRica. Il est donc compatible avec notre démarche. Pour notre étude nous utiliserons la version 2.0.

4.2 Présentation du système étudié

Le contrôle longitudinal est une fonction du système ADAS (Advanced Driver Assistance Systems) d'aide à la conduite. Concrètement, il s'agit d'un composant logiciel dont le but est d'assurer le contrôle de vitesse et du freinage en mode de conduite autonome. Il est construit autour de l'ACC (Adaptive Cruise Control), un système de régulation de la vitesse et de la distance. L'ACC calcule la vitesse que peut avoir le véhicule tout en restant en sécurité par rapport à certains événements prédéfinis (virages, bouchons, stop...). Un certain nombre d'exigences de safety lui sont associées. Parmi elles figurent celles liés aux ERL, telles que le freinage intempestif et l'accélération non contrôlée émanant de l'ACC. Il existe également d'autres contraintes ; par exemple, l'utilisateur doit toujours pouvoir désactiver l'ACC à tout moment.

Nous allons modéliser ce système en ayant à l'esprit ces exigences, tout en accordant une attention particulière aux mécanismes de safety proposés dans l'architecture du logiciel pour leur implémentation (limites de l'accélération, conditions d'activation de l'ACC). Ainsi, pour modéliser l'ACC, il faut non seulement en considérer les éléments internes mais aussi les autres éléments avec qui l'ACC est en interaction. Cela nous permettra de délimiter le périmètre du système et favorisera une meilleure interprétation des interfaces. Dans la section suivante,

nous identifierons les composants en lien avec le contrôleur longitudinal en partant du TSC et de la documentation générale de l'ADAS.

4.3 Application des trois étapes de la méthodologie

4.3.1 Choix de données d'entrée

Ayant opté dans notre démarche pour la construction d'un modèle dysfonctionnel manuellement, il fallait tout d'abord recueillir un ensemble d'informations de conception. Nous avons donc consulté les livrables du projet : documents de conception, planches des présentations officielles, tableaux Excel des exigences, safety concept. Devant le manque de clarté dans l'expression des modes de haut niveau, nous avons dû compléter notre compréhension du système en interprétant les modèles Simulink de l'ACC. Grâce aux éléments contenus dans le safety concept, nous avons pu identifier les safety goals, les composants (et sous-composants) à modéliser, ainsi que les signaux d'entrée associés à prendre en considération. De même, à partir du document de conception, nous avons identifié les modes de fonctionnement des composants ainsi que leurs conditions d'activation. Nous avons ainsi déterminé les domaines nécessaires à la modélisation du comportement des composants. Après avoir étudié les composants, nous avons créé 5 domaines particuliers avec différents états. Grâce à ces domaines, nous avons constitué des machines à états, à partir desquelles nous allons pouvoir modéliser le comportement dysfonctionnel des composants dans la partie suivante.

4.3.2 Modélisation

Le but est de modéliser le comportement dysfonctionnel des composants prenant part dans la fonction de l'ACC. Pour ce faire, dans l'outil SimfiaNeo, nous avons commencé par déclarer les domaines AltaRica (ensemble des états possibles à attribuer aux composants et aux interfaces) en nous appuyant sur les modes de fonctionnement retenu à l'étape précédente. Ainsi, nous avons créé les domaines à la fois pour représenter l'état des composant mais aussi ceux de leurs entrées et sorties. Ces informations nous permettent de modéliser facilement le comportement et les interfaces de chaque composant.

Modélisation du comportement des composant

Pour modéliser les composants et leurs états dans l'outil, nous utilisons les briques de modèle disponibles, puis nous leur attribuons les domaines préalablement créés. Il faut ensuite modéliser les conditions de franchissement des transitions en utilisant les équations identifiées dans les machines d'états. Cela se fait dans SimfiaNeo à travers la création des événements AltaRica. Un événement AltaRica est caractérisé par une garde (condition à remplir avant déclenchement du changement d'état), un effet (action résultant du changement d'état), et une loi (exponentielle, Dirac etc.). Cependant, si l'utilisation de taux de défaillances dans la modélisation des transitions est bien maîtrisée dans le cas des composants matériels, ce n'est pas le cas pour les logiciels qui sont de nature déterministe et pour lesquels il n'existe pas aujourd'hui de modèle de défaillance probabiliste assez mature. Ainsi notre démarche se différencie d'une utilisation classique d'AltaRica par la manière dont les gardes des transitions sont construites ; au lieu de nous limiter à une modélisation des transitions par les lois de probabilistes citées ci haut, nous utilisons aussi des combinaisons des états d'entrée. Traditionnellement, en AltaRica, les gardes ne font apparaître que l'état dans lequel le composant doit être pour que la transition soit franchissable ; l'événement est alors déclenché suivant la valeur d'une loi de probabilité spécifiée. Cependant, en associant les entrées aux gardes, nous pouvons aller au-delà de ce paradigme et faire en sorte que les gardes puissent

dépendre non seulement des lois de probabilité (si applicable) mais également des entrées. Cela nous permet d'avoir un changement d'état déterministe, qui ne dépend plus non seulement des lois de probabilité mais aussi de l'état des entrées. L'approche convient donc bien pour la modélisation des défaillances déterministes comme ceux des logiciels.

Modélisation des interfaces

Pour chacun des composants, nous avons écrit des logiques de propagation des défaillances liant les entrées et les sorties correspondantes. Pour ce faire, nous avons étudié des fichiers Simulink qui précisent les logiques fonctionnelles implémentées ; nous avons également trié et choisi les éléments à prendre en considération car toutes les entrées et sorties n'étaient pas utiles pour notre modèle dysfonctionnel, car n'affectant pas les mécanismes safety associés. Ensuite, nous avons décrit les états des entrées et sortie en utilisant le domaine AltaRica que nous avons nommé « Generic Data » et qui incorpore 4 états : Ok (la donnée est bonne), Loss (la donnée est perdue), Delay (la donnée arrive tardivement), Out of Range (la valeur est hors des limites acceptable). Grâce à ces états, nous étions alors en mesure de modéliser les flux d'information dysfonctionnelle entre les composants.

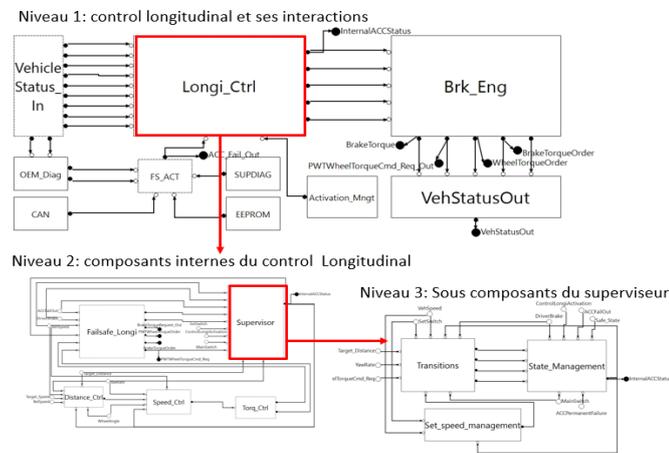


Figure 5. Trois niveaux dans la modélisation

Du fait du niveau de détail auquel il était nécessaire de descendre pour modéliser le comportement du système, nous avons obtenu un modèle en 3 niveaux tel qu'illustré en Figure 5. Ainsi, si nous nous focalisons sur l'ACC, le niveau 1 représente le contrôleur longitudinal et les composants entrant en interaction avec lui ; le niveau 2, les composants internes du contrôleur longitudinal et le niveau 3, les sous-composants des éléments du contrôleur longitudinal.

Scenario	Local effect	Local effect value	Intermediate effect	Intermediate effect value	Final effect	Final effect value
SELECT_ALL	SELECT_ALL	SELECT_ALL	SELECT_ALL	SELECT_ALL	SELECT_ALL	SELECT_ALL
	VehicleStatus_In.YawRate.YawRate	Loss	VehicleStatus_In.YawRate.YawRate	Loss	InternalACCStatus	Loss
			VehicleStatus_In.YawRate	Loss	VehStatusOut	Loss
					BrakeTorqueOrder	Loss
					PWTWheelTorqueCmd_Req_Out	Loss
					WheelTorqueOrder	Loss
					BrakeTorque	Loss
Longi_Ctrl.Supervisor.State_Management.failure	VehicleStatus_In.YawRate.YawRate	Loss	VehicleStatus_In.YawRate.YawRate	Loss	InternalACCStatus	Loss
			VehicleStatus_In.YawRate	Loss	VehStatusOut	Loss
					BrakeTorqueOrder	Loss
					PWTWheelTorqueCmd_Req_Out	Loss
					WheelTorqueOrder	Loss
					BrakeTorque	Loss
Longi_Ctrl.Supervisor.State_Management.failure	VehicleStatus_In.YawRate.YawRate	Loss	VehicleStatus_In.YawRate.YawRate	Loss	InternalACCStatus	Loss
Longi_Ctrl.Supervisor.State_Management.failure	VehicleStatus_In.YawRate.YawRate	Loss	VehicleStatus_In.YawRate	Loss	VehStatusOut	Loss
					BrakeTorqueOrder	Loss

Figure 6. AMDE

4.4 Analyses de safety

Après la modélisation complète de l'ACC et des composants qui interagissent avec lui, l'objectif était de faire des analyses de

safety à partir du modèle dysfonctionnel. Pour cela, nous avons mis en place des observateurs AltaRica sur les sorties qui nous intéressaient. Un observateur AltaRica est un indicateur qui peut être associé à une défaillance que l'on souhaite observer. En guise d'exemple, intéressons-nous à l'envoi d'une commande de gestion du moteur (PWTWheelTorqueCmd). Il s'agit d'un scénario qui peut affecter plusieurs safety goals. Il peut conduire à : 1) une accélération ou un freinage moteur intempestifs, 2) l'application d'une commande moteur alors que le statut de l'ACC ou celui du PWTWheelTorqueOrder ne le permettent pas et 3) l'application d'une commande conflictuelle avec la commande de freinage. L'objectif va donc être de vérifier, grâce à la simulation et à partir des AMDE, coupes minimales et arbres de fautes, si l'on peut déterminer les événements ou défaillances pouvant mener à l'envoi de cette commande erronée.

4.4.1 Analyse des Modes de Défaillance et leurs Effets

Nous avons ensuite utilisé SimfiaNeo pour générer des tables d'AMDE à partir du modèle dysfonctionnel construit. Les AMDE permettent de lister tous les événements provoquant la violation d'un safety goal (ou d'un observateur créé), et cela pour chaque composant du modèle. Ainsi, on retrouve en Figure 6 un certain nombre d'éléments de base d'une AMDE.

En première colonne (Scenario) figurent les événements provoquant la violation. Dans les colonnes suivantes, nous retrouvons : Local Effect (effet de l'événement sur la sortie du composant initial), Intermediate Effect (effet de l'événement sur tous les composants intermédiaires entre le composant initial et l'observateur final) et enfin Final Effect (effet sur la sortie du modèle - ici il s'agit de l'effet sur certains des observateurs décrits dans les parties précédentes).

SimfiaNeo permet d'exporter ce document sous la forme d'un tableur Excel, permettant de mieux traiter les données et de les partager avec d'autres personnes. Cela constitue un atout de l'outil, quand on sait que les outils MBSA ne sont pas forcément utilisés par un grand nombre. Cependant la Figure 6 ne représente qu'un petit extrait de l'AMDE initiale, qui présente plus de 18000 lignes. De ce fait, si le but est de d'avoir des détails exploitables ou de faire une synthèse, cette représentation n'est certes pas idéale.

4.4.2 Coupes minimales

La génération des coupes minimales passe avant tout par la configuration et la génération des séquences de calcul qui proposent plusieurs possibilités pour un même observateur. Ainsi il est possible de choisir l'ordre de la coupe, correspondant au nombre maximal de combinaisons possibles entre

événements élémentaires amenant à l'événement redouté. Pour notre étude de cas, nous avons choisi l'ordre 3 (choix lié à la performance de l'outil), car plus l'ordre est grand, plus la génération de la coupe prendra du temps et plus elle nécessitera

une puissance élevée. Un ordre de 3 constitue donc un bon compromis entre temps de calcul et précision. Le choix du type de filtre (coupes minimales ou séquences minimales) aussi est important. Tandis que les séquences donnent les combinaisons d'événements dans un ordre précis, les coupes minimales renvoient les plus petites combinaisons possibles pouvant mener à l'événement redouté. Nous avons choisi l'option « coupes minimale » car grâce à elle il sera ensuite possible de générer des arbres de fautes.

□ PWT_Torque_Order (erroneous)

Elements	Order	Probability
1 ==> Longi_Ctrl.FailSafe_Longi.SafeState.TempFailure	1	1.0E-5
2 ==> Longi_Ctrl.Torq_Ctrl.Brake_Torque.error	1	1.0E-5
3 ==> Activation_Mngt.error	1	1.0E-5
4 ==> FS_ACT_ACC_OutputGen.error	1	1.0E-5
5 ==> Longi_Ctrl.Distance_Ctrl.Rate_Limiters.error & Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error	2	1.0E-10
6 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.V_kph.VehicleSpeed_Est_comp.error	2	1.0E-10
7 ==> Longi_Ctrl.Distance_Ctrl.DynamicSaturation.error & Longi_Ctrl.Speed_Ctrl.FFD_Loop_FBK_Loop_Corrections.error	2	1.0E-10
8 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & Longi_Ctrl.Speed_Ctrl.Internal_Model_Accelerator_Ctrl.error	2	1.0E-10
9 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.Target_Distance.error	2	1.0E-10
10 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & Longi_Ctrl.Torq_Ctrl.Preprocessing.error	2	1.0E-10
11 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.RelSpeed.error	2	1.0E-10
12 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.Target_Speed.error	2	1.0E-10
13 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.WheelAngle.error	2	1.0E-10
14 ==> Longi_Ctrl.Distance_Ctrl.Feedback_Module.error & Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error	2	1.0E-10
15 ==> Longi_Ctrl.Distance_Ctrl.Vstop_Vmin_Strategies.error & Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error	2	1.0E-10
16 ==> Longi_Ctrl.Torq_Ctrl.DynamicSaturation.error & VehicleStatus_In.YawRate.error	2	1.0E-10
17 ==> Longi_Ctrl.Distance_Ctrl.Cut_In_Cut_Out.error & Longi_Ctrl.Distance_Ctrl.Following_Module.error & Longi_C...	3	1.0E-15

Figure 7. Coupe minimale

A titre d'exemple, une coupe minimale pour l'ERL « commande erronée de la gestion du moteur » est présentée en Figure 7. Elle montre des combinaisons (d'ordre 1, 2, 3) d'événements de base pouvant causer l'événement redouté spécifié (PWT_Torque_Order(erroneous)) ainsi que des probabilités associées.

Par ailleurs, on peut voir que la coupe met en avant les événements et les composants hiérarchiques, permettant d'avoir la traçabilité des composants à haut niveau comme montré en Figure 7. Pour des modèles dysfonctionnels ayant plusieurs sous-composants à nomenclature identiques, cela permet d'identifier clairement d'où provient chaque événement.

Cependant, plusieurs erreurs de compilation sont survenues, dont certaines du fait de la présence de boucles dans la chaîne de propagation des défaillances. Pour résoudre ce problème, nous avons modifié les assertions des observateurs.

La large palette de choix qu'offre SimfiaNeo à travers différentes options de configuration permet d'affiner les calculs et les simulations en fonction du besoin du projet.

4.4.3 Arbres de fautes à partir des coupes minimales

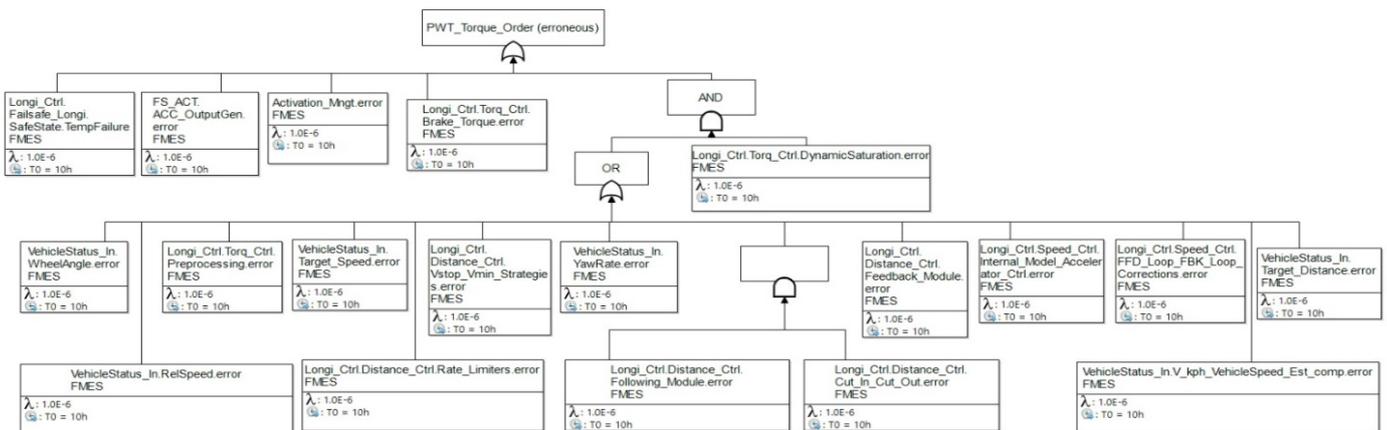


Figure 8. Arbre de faute

Nous avons pu générer des arbres de fautes à partir des coupes minimales présentées plus tôt. Comme le montre la Figure 8, les arbres de fautes permettent de voir, à travers une arborescence graphique, la chaîne de causalité entre les événements de base

au niveau des composants et l'événement redouté de haut niveau.

En modifiant les assertions des observateurs, il est possible de combiner les événements pour chercher à reproduire des événements redoutés liants plusieurs composants. La Figure 8 représente les événements pouvant amener à un ordre de gestion du moteur **erroné** et pourtant à une commande du moteur **Ok**. Ainsi, en fonction de l'analyse safety nécessaire, SimfiaNeo et AltaRica permettent de développer des assertions pour représenter des situations très diverses.

Toutefois il est parfois difficile de traduire certaines exigences safety (comme celles ayant des contraintes de temps) en expression AltaRica. Il s'agit d'un problème spécifique de la version 2 d'AltaRica implantée dans l'outil SimfiaNeo. Il s'agit donc là d'une limite de l'approche, cependant liée à la version utilisée du langage.

5 DISCUSSION DES RESULTATS

L'application de la proposition méthodologique à l'étude de cas montre qu'il est possible d'appliquer au logiciel une approche de type MBSA, jusqu'à présent orientée système. A travers un bon choix des données d'entrée et un effort de modélisation, notre approche permet d'obtenir un modèle dysfonctionnel représentatif du système réel, ce qui améliore la qualité des analyses de safety. Grâce à cette démarche, il est donc possible d'unifier différents modèles de safety classiques dans un modèle dysfonctionnel unique. De plus la démarche offre un énorme atout pour la réutilisation. En effet, pour évaluer de nouveaux ERL, il suffira de les décomposer en des combinaisons logiques et de les associer à l'architecture dysfonctionnelle déjà modélisée. Une fois le modèle dysfonctionnel construit, celui-ci permettra de mener des analyses avec des paramètres différents et ce pour un grand nombre d'ERL car il suffira d'exprimer ces derniers en utilisant des combinaisons booléennes appropriées. La proposition méthodologique améliore ainsi les pratiques actuelles. Grâce à l'approche décrite et à l'utilisation d'outils comme SimfiaNeo, il est possible de déplacer la tâche de l'expert safety, de la construction manuelle des AMDE et arbres de fautes, à la construction d'une architecture dysfonctionnelle. En focalisant son énergie sur la modélisation dysfonctionnelle, la méthode rapproche les modèles safety le plus proche possible des modèles du design, ce qui améliore de ce fait la qualité des analyses.

Ainsi, le gain se mesure en temps, en qualité des analyses et en possibilités de réutilisation des modèles.

En outre, l'utilisation de la sémantique formelle d'AltaRica permet de représenter le comportement du système sans

ambiguïté, qui en fait un candidat possible dans un contexte de certification.

L'ensemble de ces éléments permettent d'améliorer significativement les pratiques actuelles en matière d'analyse de la safety. Cependant, l'une des difficultés avec un modèle dysfonctionnel dédié, est la problématique du maintien en cohérence de celui-ci avec le modèle de conception lorsque ce dernier évolue. Il est donc nécessaire de mettre en place des mesures supplémentaires pour garantir la cohérence. Ainsi nous avons relevé certaines limitations durant l'étude de cas. La modélisation peut prendre du temps ; il faut se placer au juste niveau de détail dans la représentation du système. Nous proposons de clairement définir les éléments d'entrée (composants, signaux, mécanismes safety) à étudier à l'étape 1. Grâce à cette étape préliminaire et en nous basant sur les informations du TSC (mécanismes safety, criticités), nous sommes en mesure de nous assurer que les éléments liés à la safety ne sont pas ignorés. Une autre limitation constatée est relative à l'incapacité de AltaRica 2.0 à gérer les boucles ; notre solution a été de modifier certaines assertions de manière à supprimer ces boucles. Une dernière limitation, relative aux performances de SimfiaNeo, est son besoin important en ressources (mémoire et processeur) ; ainsi, l'un des problèmes rencontrés a été la perte inexplicable d'une partie du fichier source de notre modèle et nous avons dû le reconstruire manuellement. C'est ce qui explique notre choix de nous limiter à l'ordre 3 pour les coupes. Cela soulève des questions de fiabilité de l'outil. Enfin, l'étude de cas concernant un système de complexité réduite, le passage à l'échelle reste à évaluer.

6 CONCLUSION

Nous avons fait une proposition méthodologique reposant sur l'ingénierie dirigée par les modèles qui permet, en trois étapes, de construire un modèle dysfonctionnel à partir duquel il est possible de dériver des modèles classiques safety.

Grâce à l'outil SimfiaNeo et en utilisant le langage AltaRica, nous avons appliqué la méthodologie sur une étude de cas, en construisant un modèle dysfonctionnel d'un logiciel à partir duquel nous avons pu générer des AMDE, coupes minimales et arbres de fautes. Ces résultats sont encourageants. Ils montrent qu'il est possible d'appliquer une approche MBSA pour évaluer la safety des logiciels, notamment en automobile. Ils démontrent également les bénéfices d'une génération des analyses de safety à partir d'un modèle dysfonctionnel (gain en temps, qualité des analyses, réutilisation).

Forts de ces résultats, l'étude doit maintenant se poursuivre, afin d'évaluer la complexité des systèmes pour lesquels la méthode et l'outillage peuvent être raisonnablement appliqués. De plus, dans la mesure où la proposition de cet article s'appuie sur un modèle dysfonctionnel dédié, il faudra également adjoindre à la méthode un mécanisme de mise en cohérence entre les modèles de conception et les modèles de safety.

7 REFERENCES

- Batteux, M. B., Prosvirnova, T., & Rauzy, A. (2019, October). Model synchronization: A formal framework for the management of heterogeneous models. *International Symposium on Model Based Safety Assessment, IMBSA 2019*. https://doi.org/10.1007/978-3-030-32872-6_11
- Bozzano, M., & Villaflorita, A. (2007). The FSAP/NuSMV-SA Safety Analysis Platform. *International Journal on Software Tools for Technology Transfer*, 9(1), 5–24. <https://doi.org/10.1007/s10009-006-0001-2>
- Center for Chemical Process Safety. (2010). Appendix D: Minimal Cut Set Analysis. In *Guidelines for Chemical Process Quantitative Risk Analysis* (pp. 661–670). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470935422.app4>
- Colaco, J.-L., Pagano, B., Pouzet, M., & Pouzet, M. (2017). Scade 6: A Formal Language for Embedded Critical Software Development. 10.
- Cuenot, P., Frey, P., Johansson, R., Lönn, H., Papadopoulos, Y., Reiser, M.-O., Sandberg, A., Servat, D., Kolagari, R. T., Törngren, M., & Weber, M. (2007). The EAST-ADL architecture description language for automotive embedded software. *Proceedings of the 2007 International Dagstuhl Conference on Model-Based Engineering of Embedded Real-Time Systems*, 297–307.
- David, P., Idasiak, V., & Kratz, F. (2009). Automating the synthesis of AltaRica Data-Flow models from SysML. In M. & G. S. Briš (Ed.), *ESREL 2009* (p. 8). Taylor & Francis Group. <https://hal.archives-ouvertes.fr/hal-00579532>
- DoD. (2012). MIL-STD-882E "Department of Defense Standard Practice System Safety". <https://www.dau.edu/cop/armyesh/DAU%20Sponsored%20Documents/MIL-STD-882E.pdf>
- Estefan, J. A. (2008). *Survey of Model-Based Systems Engineering (MBSE) Methodologies*. 70.
- Feiler, P. H., Lewis, B., & Vestal, S. (2003). The SAE Avionics Architecture Description Language (AADL) Standard: A Basis for Model-Based Architecture-Driven Embedded Systems Engineering: Defense Technical Information Center. <https://doi.org/10.21236/ADA612735>
- Fenelon, P., & McDermid, J. A. (1993). An integrated tool set for software safety analysis. *Journal of Systems and Software*, 21(3), 279–290. [https://doi.org/10.1016/0164-1212\(93\)90029-W](https://doi.org/10.1016/0164-1212(93)90029-W)
- ISO 26262-1_2018_Ed2. (n.d.).
- Legendre, A., Lanusse, A., & Rauzy, A. (2016, December). Synchronisation des modèles d'architecture et d'analyse de risques: Quel gain, comment et pourquoi? *Congrès Lambda Mu 20 de Maîtrise des Risques et de Sécurité de Fonctionnement. Congrès Lambda Mu 20 de Maîtrise des Risques et de Sécurité de Fonctionnement*, 11-13 Octobre 2016, Saint Malo, France. <https://doi.org/10.4267/2042/61813>
- Lisagor, O., Kelly, T., & Niu, R. (2011). Model-based safety assessment: Review of the discipline and its challenges. *The Proceedings of 2011 9th International Conference on Reliability, Maintainability and Safety*, 625–632. <https://doi.org/10.1109/ICRMS.2011.5979344>
- Nguyen, N., Mhenni, F., & Choley, J.-Y. (2018). AltaRica 3.0 code generation from SysML models (pp. 2435–2440). <https://doi.org/10.1201/9781351174664-306>
- Papadopoulos, Y., & McDermid, J. A. (1999). Hierarchically Performed Hazard Origin and Propagation Studies. *Computer Safety, Reliability and Security*, 139–152. https://doi.org/10.1007/3-540-48249-0_13
- Point, G., & Rauzy, A. (1999). AltaRica: Constraint automata as a description language. <http://www.altarica-association.org/ressources/ARBib/PointRauzy1999-AltRicaConstraintLanguage.pdf>
- SAE. (2015). *Potential Failure Mode and Effect Analysis Reference Manual*. https://www.lehigh.edu/~intribos/Resources/SAE_FMEA.pdf
- Sirgabsou, Y., Pahun, L., Baron, C., Grenier, L., Bonnard, C., & Esteban, P. (2020). Investigating the use of a model-based approach to assess automotive embedded software safety. 9.
- VDA QMC. (2015). *Automotive SPICE Process Assessment / Reference Model*. http://www.automotivespice.com/fileadmin/software-download/Automotive_SPICE_PAM_30.pdf
- Vesely, W. E., Goldberg, F. F., Roberts, N. H., & Haasl, D. F. (1981). NUREG-0492, "Fault Tree Handbook". 209.

Conception sûre des systèmes de production : intégration de la santé/sécurité des opérateurs à travers l'analyse des usages

SERGIO O. ARIAS C.*¹, BRUNO DAILLE-LEFEVRE¹, XAVIER GODOT², PATRICK MARTIN²

¹Institut national de recherche et de sécurité (INRS),
Département Ingénierie des Equipements de Travail,
1 rue du Morvan, 54519 Vandœuvre-Lès-Nancy cedex, France
arias_corzo@hotmail.com
bruno.daille-lefevre@inrs.fr

²Laboratoire de Conception Fabrication Commande,
Arts et Métiers, Université de Lorraine,
4 rue Augustin Fresnel, 57078 Metz, France
xavier.godot@ensam.eu
patrick.martin@ensam.eu

*Auteur correspondant

Résumé – Lors d'une démarche de conception de poste de travail, les leviers d'action du concepteur face à d'éventuels risques dépendront du moment où celui-ci sera en mesure d'identifier ces risques. En effet, s'ils sont découverts trop tard, les retours en arrière sont très difficiles, voire impossible dans certains cas. Quoiqu'il en soit, ils impacteront inmanquablement le projet en termes de coût et de délai.

En cas de poursuite du projet malgré les risques encourus, le poste de travail ne sera pas adapté aux opérateurs, ce qui provoquera une diminution de la productivité et une augmentation inéluctable de situations dangereuses.

Dans le cadre de ce projet, pour pouvoir assurer la santé et la sécurité des opérateurs, nous avons développé une méthodologie qui permettra aux concepteurs d'identifier le plus rapidement possible les usages de poste de travail et les risques associés. Cette méthodologie combine une démarche de conception générique qui s'appuie sur une approche systémique afin d'identifier les usages que les opérateurs feront du système pour assurer les impératifs de production, et une démarche reposant sur une approche énergétique, qui elle, vise à identifier les dangers potentiels.

Abstract – During a workstation design process, the capacity to deal with possible risks will be determined by the moment when the designer is able to identify these risks. Indeed, if they are discovered too late, it's very difficult, and in some cases impossible, to take a step backwards. Nevertheless, they will inevitably impact the project in terms of cost and deadlines.

If the project continues in spite of the risks involved, the workstation will not be suitable for the operators, which will lead to a decrease in productivity and an inevitable increase in dangerous situations.

Within the framework of this project, in order to be able to ensure the health and safety of the operators, we have developed a methodology that will allow the designers to identify as quickly as possible the uses of the workstation and the associated risks. This methodology combines a generic design approach based on a systems approach to identify the uses that operators will make of the system to ensure production requirements, and an approach based on an energy approach, which aims to identify potential dangers.

Mots clés - industrie du futur, analyse fonctionnelle, ingénierie système, volume d'influence, santé-sécurité des opérateurs.

Keywords - factory of the future, functional design, systems engineering, volume of influence, health and safety of operators.

1 INTRODUCTION

À l'ère des réseaux sociaux et d'un monde de plus en plus connecté, chaque individu désire un degré de personnalisation élevé dans ce qu'il achète et ce qu'il consomme, et ce, malgré des ressources qui se font de plus en plus rares. Cette tendance entraîne l'apparition de nouveaux défis sur le plan de « l'industrie du futur », puisqu'il est important de mettre en place des systèmes de production dits flexibles qui répondront plus facilement à cette demande personnalisée et permettront d'offrir une gamme de produits plus large tout en utilisant le moins de ressources possibles. [Bidet-Mayer, 2016], [Evers, 2018], [Østergaard, 2018].

Pour y parvenir tout en garantissant la santé et la sécurité des

opérateurs, il est nécessaire de connaître les caractéristiques du système de production, et plus particulièrement l'usage qu'en feront ces opérateurs pour identifier les risques associés.

Dans ce contexte, nous avons développé, au sein du laboratoire commun sur la « conception sûre de situation de travail », le projet intitulé « aide à la conception de poste de travail par la prise en compte des variabilités d'usages et des exigences de sécurité », afin d'introduire une méthodologie qui pourra aider les concepteurs à identifier dès les premières phases du projet les usages que les opérateurs feront du système pour assurer les impératifs de production et les risques avérés.

2 ETAT DE L'ART

La norme NF EN ISO 6385 (2016), définit un poste de travail comme étant une combinaison et aménagement spatial des équipements de travail, au sein de l'environnement de travail, selon les conditions imposées par les tâches.

Afin d'assurer la santé et sécurité des opérateurs, il faut parvenir à un équilibre entre les capacités de l'opérateur et les exigences du poste de travail.

À cet égard, l'état de l'art a été fait sur les démarches de conception de produits et de système, mais aussi sur la prise en compte de la santé et de la sécurité des opérateurs dans l'activité de conception.

Concernant le premier axe, les seuls objectifs sont d'identifier le moment précis durant lequel le concepteur sera amené à construire l'usage qu'un opérateur fera du système de production en cours de conception, mais aussi sous quelle forme ces usages sont représentés.

Concernant le second axe, le but est de trouver une méthode outillée permettant d'identifier, d'estimer et d'évaluer les phénomènes dangereux auxquels un opérateur est exposé à travers l'exploitation des seules informations liées à l'usage qu'il fera d'un système de production donné.

2.1 Identification d'usages

2.1.1 Les démarches de conception

Il existe de nombreuses démarches de conception qui ambitionnent toutes de formaliser et structurer cette activité. Cependant, elles diffèrent profondément notamment sur leur niveau de complétude, mais aussi sur leur niveau de généricité ou encore leur niveau de détail. [Paynter, 1961], [Cross et Roy, 1989], [Suh, 1990], [Das et Sengupta, 1996], [Pahl et Beitz, 1996], [Ullman, 2003], [Hatchuel et Weil, 2003], [Gero, 2003], [Scaravetti, 2004], [Godot, 2013].

Plusieurs synthèses permettent de les positionner par rapport à l'activité de conception dans son ensemble, c'est-à-dire en partant du besoin brut du client pour aller jusqu'à la définition complète du produit ou du système physique. Ainsi, ces travaux ont permis d'identifier un cadre générique constitué de cinq grandes phases (voir figure 1).



Figure 1. Démarche de conception générique.

2.1.2 Usages du système par l'opérateur

En termes de conception, les usages se traduisent sous forme de « fonctions » et de « contraintes ». Ainsi, nous pouvons définir un usage du point de vue de la conception comme étant un groupe de fonctions organisées de manière cohérente afin d'atteindre un objectif donné tout en respectant un ensemble de contraintes. Par conséquent, nous avons sélectionné les démarches de conception abordant de façon détaillée et structurée les fonctions et les contraintes. Ainsi, nous allons

aborder plus en détails l'Analyse Fonctionnelle du Besoin (AFB) et l'Ingénierie Système (IS).

2.1.3 Analyse Fonctionnelle de Besoin (AFB)

L'AFB, [AFNOR, 1991], [AFNOR, 2009] est une méthode outillée qui permet de décrire une fonction ou un ensemble de fonctions et leurs relations, qui répondent au besoin d'un client, lui-même conditionné par différents facteurs « techniques, économiques, réglementaires, sociologiques, etc. » ; ces fonctions sont systématiquement caractérisées, classées et évaluées.

La démarche à suivre peut-être décrite comme suit :

1. *Identifier les besoins* : ces besoins sont des nécessités ou désirs éprouvés par un utilisateur, exprimés par le demandeur (le client) lesquels vont être adaptées à l'environnement fonctionnel.
2. *Traduire les besoins sous forme de fonctions à remplir par le produit* ; ces fonctions sont décrites à partir des besoins de l'utilisateur en exprimant leur relation avec l'environnement.
3. *Décrire et associer des fonctions services, technique et contrainte à chaque fonction* ; ces fonctions sont décrites à partir des phases du cycle de vie du produit.

Cette démarche peut être représentée comme indiquée en figure 2.



Figure 2. Représentation graphique de la démarche d'AFB

2.1.4 Ingénierie Système (IS)

L'IS [ISO/IEC, 2007], [ISO/IEC/IEEE 24765, 2009], [ISO/IEC/IEEE, 2011], [Faisandier, 2011], [Fiorèse et Meinadier, 2012], [ISO/IEC/IEEE, 2015], est un ensemble de concepts, méthodes et outils reposant sur une approche systémique de l'activité de conception. Ainsi, le produit à concevoir est considéré comme un ensemble d'éléments interagissant de manière coordonnée et cohérente appelé « système ». Sa finalité est de répondre au besoin du demandeur appelé « mission ». Mais l'approche systémique permet également d'aborder le caractère multi-technologique d'un système à travers l'intégration des différents champs disciplinaires requis pour le concevoir.

Du point de vue méthodologique, l'IS est portée par le « Cycle en V » qui décrit et organise chacune des étapes de conception (voir figure 3).

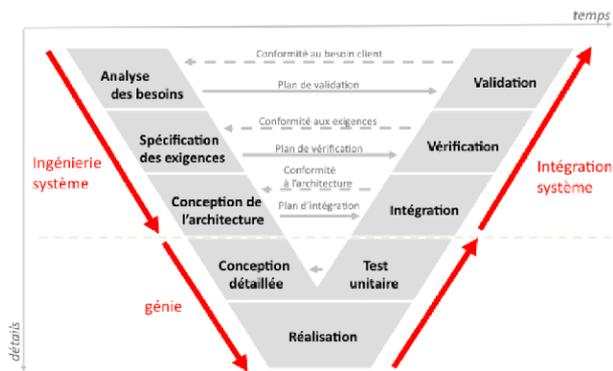


Figure 3. Cycle en V d'IS

De même qu'avec l'AFB, cette démarche peut être représentée comme indiquée en figure 4.



Figure 4. Représentation graphique de la démarche d'IS

Afin d'établir le périmètre et la portée du projet, les démarches représentées en figures 2 et 4 ont été positionnées par rapport au cadre général de conception présenté en figure 1. Il apparaît clairement que l'AFB et l'IS interviennent dans les deux premières phases du cadre général, à savoir « le besoin du client et le contexte » mais aussi « la conception architecturale ». De plus, l'élicitation d'usages tels que nous les avons précédemment définis intervient dans les étapes 2, 3, 4 et 5 que ce soit pour l'AFB ou pour l'IS (voir figure 5).

Les résultats de la comparaison sont présentés ci-dessous.

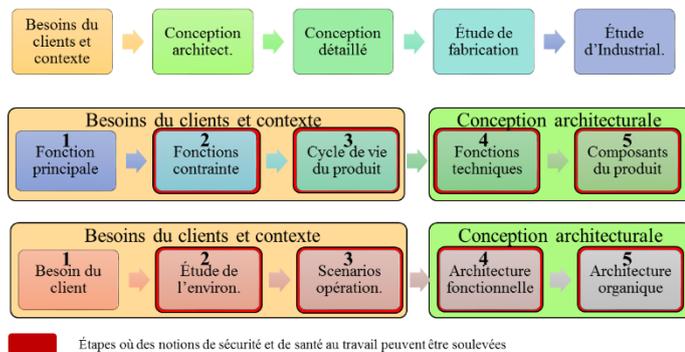


Figure 5. Comparaison entre l'AFB et l'IS sur la base d'une démarche de conception générique

2.1.5 Conclusion concernant les démarches de conception

Dans le cadre de ce projet, nous avons privilégié l'utilisation de l'IS, car elle répond mieux aux besoins de l'étude au niveau d'usages. En effet, l'AFB permet d'identifier les fonctions, les contraintes et les composants du système, mais la nature de leurs interactions n'est que partiellement traitée et l'exhaustivité des éléments collectés n'est pas garantie. Concernant l'IS, ces aspects sont gérés à travers les notions « d'architecture fonctionnelle » (ensemble de fonctions interagissant via des liens d'antériorités et des liens logiques) et « d'exigence » (estimation, évaluation, traçabilité des contraintes et des

niveaux de performance souhaités).

2.2 Prise en compte de la santé et de la sécurité des opérateurs dans l'activité de conception.

Concernant la prise en compte de la santé/sécurité des opérateurs, la méthodologie générale est donnée par la norme. Trois étapes essentielles y sont présentées : l'identification, l'estimation et l'évaluation (voir figure 6). [Marsot, 2001], [Fadier et De la Garza, 2006], [Ghemraoui, 2009], [LE PARLEMENT EUROPÉEN, E. L. C. & DE L'UNION, E. U. R. O., 2009], [ISO, N., 2010], [INRS, 2011].

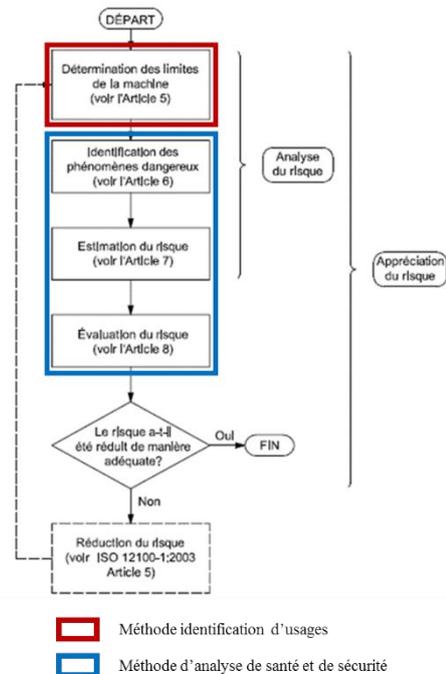


Figure 6. Démarche simplifiée d'identification des risques

Ces trois aspects ont été traités dans plusieurs études qui mettent en évidence le lien direct entre le danger et les flux d'énergies. Parmi celles-ci, se trouvent le modèle MOSTRA, la méthode EZID et la méthode des « Volumes d'Influence ».

2.2.1 Modèle de Situation de Travail (MOSTRA)

MOSTRA [Hasan, 2002], est un modèle permettant de représenter une « situation de travail ». Il suit la logique suivante : dans la « situation de travail », le système remplit une « fonction ». La « fonction » se base sur des « solutions techniques » et est satisfaite par la réalisation des « tâches ». Ces « tâches » nécessitent des « outils » et « consommables » appelés « auxiliaires » et une « équipe de travail ». Par ailleurs, les « solutions techniques » sont source de « phénomènes dangereux » qui engendrent des « risques » dans les « zones dangereuses ». L'existence de « risques » implique de prendre des mesures de sécurité. Enfin, un « événement dangereux » peut déclencher un accident et influencer directement l'indice de risque.

2.2.2 Méthode d'analyse énergétique pour l'identification systématique des dangers pendant la conception (EZID)

EZID [De Galvez, 2016] permet d'identifier, d'estimer et

d'évaluer les phénomènes dangereux auxquels un opérateur sera exposé lors de l'utilisation d'un système de production. Elle s'inscrit dans le cadre de la conception. Elle repose sur une modélisation du système et trois niveaux de traitement.

1. *Modélisation du système* : permet de capitaliser et d'identifier tous les flux et source d'énergies généralisées (potentielle, cinétique, électrique, thermique, chimique, rayonnement, structurelle) issus des pièces et composants du système de production.
2. *Premier niveau de traitement* : répertorier et identifier la nature des phénomènes dangereux à partir des paramètres caractérisant les sources et flux d'énergies (*variables généralisées étendues*).
3. *Second niveau de traitement* : identifier les causes des phénomènes dangereux précédemment identifiés.
4. *Troisième niveau de traitement* : identifier le caractère significatif des phénomènes dangereux (*soit leur capacité à générer un risque*) ; cette identification repose sur la comparaison des niveaux d'énergie de la machine avec des valeurs seuils réglementaires ou normalisés.

2.2.3 Méthode des Volumes d'Influences (VI)

La méthode des VI [Martin et al, 2019] a les mêmes objectifs que la méthode EZID et intervient également lors de la conception. Elle repose sur une analyse spatio-temporelle systématique et structurée du système de production à travers les flux d'énergie pouvant potentiellement exister entre les composants et les opérateurs de celui-ci.

L'identification et l'évaluation des phénomènes dangereux sont réalisées à partir d'éléments spatiaux, temporels et énergétiques issus du système de production comprenant, des opérateurs et des éléments à produire (séquencement des actions et implantation des constituants du système de production). Quant à la quantification des phénomènes dangereux, elle est faite en deux temps. Tout d'abord, elle est traitée via le gradient d'énergie défini par rapport au volume que peut potentiellement occuper cette énergie lorsqu'elle est mise en œuvre (voir figure 7). Enfin, elle est traitée à travers des tables d'équivalences entre le niveau d'énergie auquel est soumis un opérateur et le niveau de blessures que celui-ci peut engendrer.

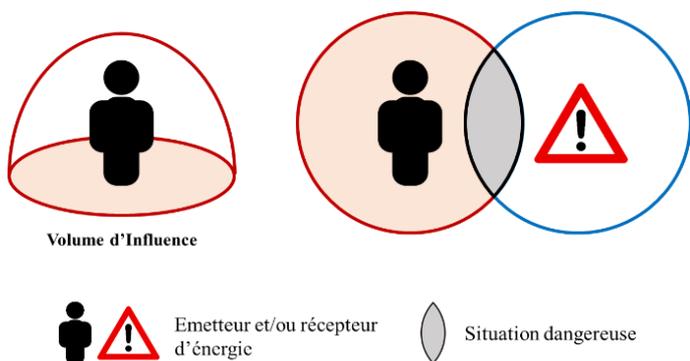


Figure 7. Volumes d'Influence

2.3 Conclusion concernant la prise en compte de la santé et de la sécurité des opérateurs dans l'activité de conception.

MOSTRA identifie l'ensemble des objets nécessaires à la réalisation d'une analyse de santé/sécurité des opérateurs lors de

la conception d'un système. Ces objets concernent le système proprement dit et les phénomènes dangereux. Cependant, il ne répond donc pas à notre besoin, car en dehors d'une logique assurant la cohérence de l'ensemble du modèle, il ne propose pas de méthode systématique permettant d'identifier, d'estimer et d'évaluer les phénomènes dangereux auxquels un opérateur est confronté lors de l'utilisation d'un système de production. Il ne permet donc pas de quantifier les risques associés.

EZID est un modèle de données associé à une méthode complète d'analyse des phénomènes dangereux que peut rencontrer un opérateur de système de production. De ce point de vue, il répond à notre besoin. Néanmoins, il s'appuie sur une description détaillée de ce dernier qui nécessite une mise en œuvre à un stade très tardif de la conception du système (exemple : géométrie des pièces élémentaires). Par conséquent, la méthode EZID ne répond pas à nos besoins sur ce plan précis.

La méthode des VI répond à nos besoins. En effet, elle est structurée et entièrement outillée. De plus, elle s'appuie sur l'exploitation d'informations spatiales et temporelles dont le niveau de détail semble parfaitement compatible avec les données disponibles au moment de l'élaboration des usages que les opérateurs feront du système de production. Cette approche constitue la base de la méthode proposée dans le cadre de cette étude.

3 PRESENTATION ET APPLICATION DE LA METHODOLOGIE

3.1 Présentation de la méthodologie d'aide à la décision

Compte-tenu de l'état de l'art, nous avons suivi une approche permettant de proposer un mécanisme d'intégration de la santé/sécurité au moment où sont élicités les usages lors de la conception du système de production et de le tester immédiatement sur un cas d'étude réel avant d'en tirer toutes les conclusions qui s'imposent. La figure 8 illustre cette progression.

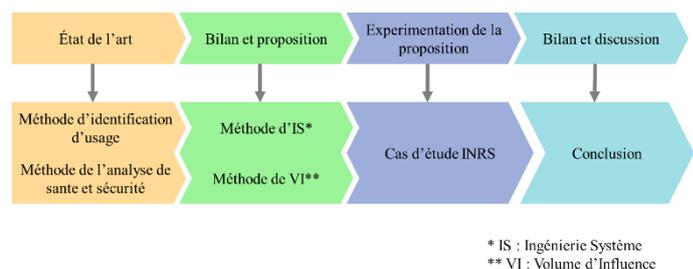


Figure 8. Approche scientifique pour réaliser l'étude

Même si nous nous appuyons sur les concepts et modèles de l'IS, nous devons impérativement détailler la méthode de mise en œuvre opérationnelle afin de nous assurer qu'elle permettra de faire apparaître des données de conception dont le niveau de détail sera compatible avec celles qui sont nécessaires à l'application de la méthode des VI.

Ainsi, cette intégration a nécessité l'identification de la nature et du niveau de granularité des données permettant leur exploitation pertinente tant du point de vue de la conception que de celui de la santé/sécurité. Ce travail a été formalisé à travers une méthode de conception détaillée suivant la démarche de l'IS. Cette méthode est constituée d'un ensemble ordonné

d'étapes, chacune étant caractérisée par ses données d'entrée, ses données de sortie et la manière de les créer. La figure 9 présente le détail de chaque étape est traité dans la section 3.1 de cet article.

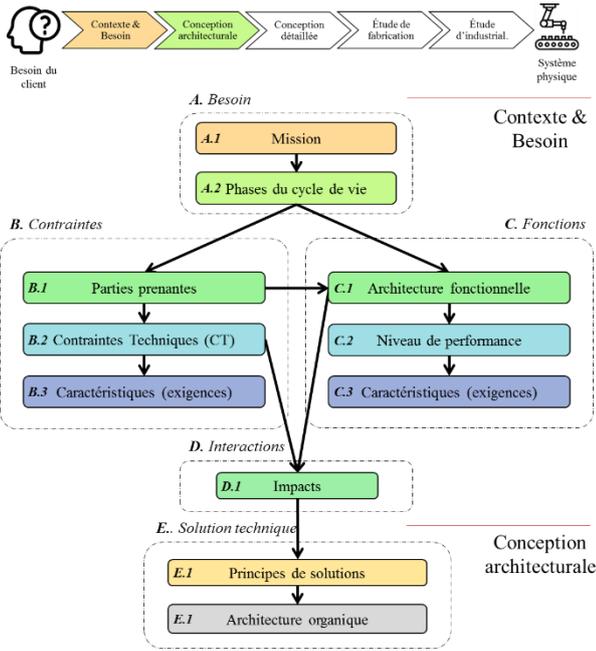


Figure 9. Méthode développée

Concernant la structuration et la formalisation des données de conception, nous avons dû définir de nouveaux outils que ceux préconisés par l'IS (en l'occurrence « SysML ») afin qu'ils soient adaptés à la nature et au niveau de granularité des données de conception issues de la mise en œuvre de la méthode que nous proposons en figure 9. Ainsi, la figure 10 présente l'outil de représentation utilisé pour formaliser les données issues des blocs d'étapes A, B et C de la figure 9. Il s'agit d'une carte heuristique. Quant à la figure 11, elle présente la manière dont sont structurées les données émanant de l'étape D de la figure 9. Il s'agit d'une simple matrice. Enfin, la figure 12 présente la structuration des données du bloc d'étapes D de la figure 9. Il s'agit ici d'une représentation de type « entité/relation ».

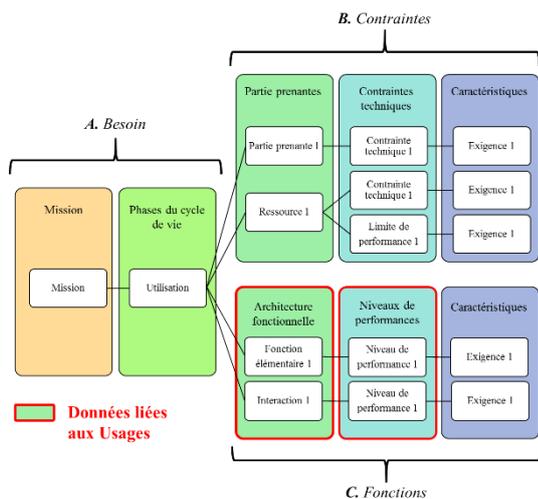


Figure 10. Formalisation et structuration des données de conception (contexte et besoin)
Voir figures 9A, B et C

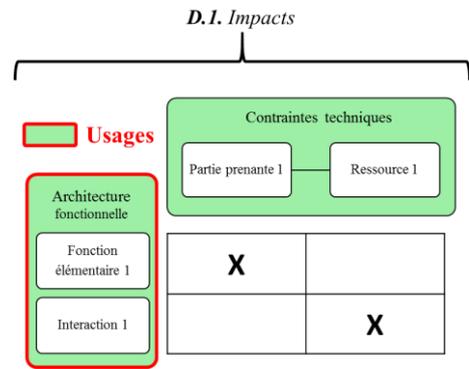


Figure 11. Liens entre les contraintes techniques et l'architecture fonctionnelle
Voir figure 9D

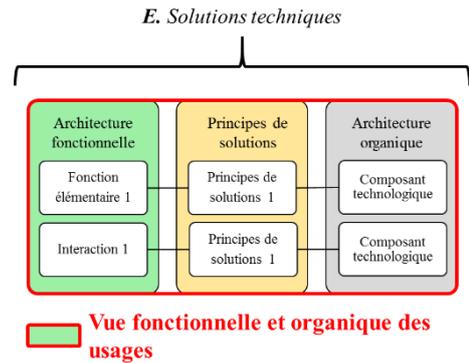


Figure 12. Formalisation et structuration des données de conception (conception architecturale)
Voir figure 9E

3.1 Mise en œuvre opérationnelle de la méthode à travers la réalisation d'une étude de cas

Pour mettre à l'épreuve notre méthode, nous avons réalisé les deux premières phases (contexte et besoin, puis conception architecturale) de la conception d'un nouveau poste de travail destiné à équiper l'atelier de l'Institut National de Recherche et de Sécurité (INRS). Nous n'avons pas encore traité la mise en œuvre de la méthode des VI par manque de temps. Cependant, nous nous sommes assurés que les données de conception créées permettent bien l'application de celle-ci (voir figure 18).

Les principales caractéristiques de cet équipement et de son environnement sont illustrées en figure 13.

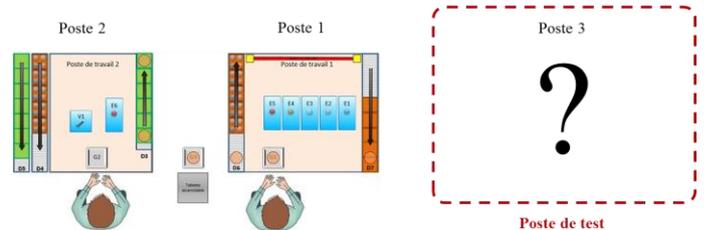


Figure 13. Poste de travail à concevoir

Ce poste de travail s'insère dans une ligne de production déjà constituée de deux autres postes d'assemblages. Il a pour objectif de tester les produits assemblés d'un point de vue

électrique. Dans la suite de cet article, ce poste sera désigné comme « système ».

La première étape de notre méthode consiste à définir la « mission » du système (voir figure 9 A.1). Elle a pour objectif de s'assurer que le client et le concepteur perçoivent le système de la même façon afin d'éviter les malentendus en fin de projet. Pour cela, nous rédigeons une unique phrase constituée d'un seul verbe d'action et de deux compléments : le premier précise le ou les objets sur lesquels cette action portera et le second précis le périmètre de cette action en termes du niveau de performance global. Voici la mission définie pour le système : *contrôler la conformité des produits en faisant des tests électriques de basse et de haute tension pour assurer que 100 % des produits vendus fonctionnent.*

La deuxième étape concerne le cycle de vie du système. L'objectif est de circonscrire le périmètre exact de l'étude afin de ne pas omettre certains usages. Pour cela, nous avons identifié les différents contextes (ou environnements) dans lesquels le système sera plongé. Ces derniers représentent une phase du cycle de vie du système (voir figure 9 A.2). Compte tenu de l'objectif du projet, nous nous sommes intéressés plus particulièrement aux phases d'utilisation, de réglage et de maintenance, qui sont les phases où le système pourra interagir avec ses « utilisateurs » (l'opérateur, la maintenance et des tierces personnes).

La troisième étape a pour objectif de recenser tous les éléments matériels et immatériels, qui interagiront directement avec le système. En outre, les éléments potentiellement utilisables en tant que ressources sont également répertoriés (voir figure 9 B.1).

Les quatrième et cinquième étapes ont pour finalité de détailler la nature et interactions et les quantifier. Leur nature est notamment déterminée à travers les flux d'énergies. Quant à leur dimensionnement, il suit le formalisme des « exigences » telles qu'elles sont décrites en IS. Chacune d'entre elles est donc constituée d'un paramètre, d'une question de validation, d'une valeur, d'une tolérance et d'un moyen de contrôle (voir figure 14). Ces éléments forment alors les « contraintes techniques » du système (voir figure 9 B.2 et B.3).

Contrainte technique					
Partie prenante	Paramètre	Question de validation	Valeur	Tolérance	Moyen de contrôle
Magasinier	Charge physique de travail	La personne a-t-elle un handicap ?	oui/ non	N/A	Certificat RH, diplôme ou certificat d'études
Aire comprimé	Pression	N/A	7 bar	± 0,5 bar	Baromètre

N/A : Non Applicable

Figure 14. Contraintes techniques du système

La sixième étape est très importante puisque c'est à ce moment précis que nous élaborons les différents usages du système. Il s'agit d'imaginer une suite d'actions séquencée de manière logique permettant de répondre entièrement la mission décrite à l'étape. Cet ensemble forme « l'architecture fonctionnelle du système » (voir figure 9 C.1). Chaque action constitue alors une « fonction élémentaire » (voir figure 15). Pour une même mission, il existe généralement plusieurs solutions possibles. Dans ce cas, un choix doit être fait par le concepteur en utilisant par exemple des outils d'analyse de la valeur.

N°	Fonction	Fonc. élémentaire
F10	Positionner le produit sur le système dans la zone de stockage amont	Positionner
F20	Solidariser le produit et le système en vue de son déplacement vers la zone de test	Solidariser
F30	Déplacer le produit de la zone de stockage amont vers la zone de test	Déplacer

Figure 15. Fonctions et fonctions élémentaires du système

Les étapes sept et huit ont le même but et les mêmes modalités de réalisation que les étapes quatre et cinq, à savoir : détailler et quantifier, mais cette fois-ci le niveau de performance des fonctions élémentaires (voir figures 9 C.2 et C.3 et 16).

NIVEAU DE PERFORMANCE					
Fonc. élémentaire	Paramètre	Question de validation	Valeur	Tolérance	Moyen de contrôle
Positionner	Orientation	Le cube est-il centré par rapport à la zone de test ?	oui/ non	N/A	Poka-Yoke
Déplacer	Temps	N/A	t1	± t s	Cronomètre

N/A : Non Applicable

Figure 16. Niveau de performance des fonctions élémentaires du système

La neuvième étape permet de lier les actions et les fonctions élémentaires afin que le concepteur puisse aisément intégrer les contraintes dans les choix techniques et technologiques qu'il sera amené à faire pour répondre aux différentes fonctions du système (voir figure 10 D.1). Une simple matrice de liens permet d'indiquer quelles contraintes techniques s'appliquent sur quelles fonctions élémentaires (voir figure 11). Ici s'achève la première phase : « contexte et besoin ».

Par conséquent, la dixième étape inaugure la phase de « conception architecturale ». Il est important de noter que le concepteur considérera alors le système qu'il est en train de concevoir à travers ses éléments physiques. Cette transition s'opère tout d'abord en associant à chaque fonction élémentaire un principe de solution (voir figures 9 E.1 et 12). Celui-ci est décrit à partir d'une combinaison de transformations énergétiques assurant l'action, la commande et le contrôle de manière à garantir un fonctionnement en boucle fermée de l'ensemble (voir figure 17).

N°	Fonction	Fonc. élémentaire	Sous-fonct.	Solution technique
F10	Positionner le cube sur le système dans la zone de stockage amont	Positionner	Action	Distributeur pneumatique
			Contrôle	Contrôle visuel
			Commande	Commande électrique

Figure 17. Solutions choisies par sous-fonctions des fonctions du système

La onzième et dernière étape permettent dans un second temps d'associer à chaque principe de solution un ensemble de composant techniques standards ou spécifiques. L'assemblage des composants de l'ensemble des fonctions élémentaires forme « l'architecture organique » du système (voir figures 9 E.2 et 12). L'élaboration d'une de l'agencement géométrique des différents composants peut être réalisé à ce stade (mise en œuvre des logiciels de CAO).

Il est important de noter que ces deux dernières étapes peuvent faire apparaître plusieurs solutions. Le concepteur doit procéder à un choix en s'appuyant sur les mêmes outils qu'à l'étape 6.

Au fur et à mesure de l'avancement de la conception du poste de travail, nous avons constaté que nous étions systématiquement confrontés à deux types de situation : celles constant uniquement à collecter des informations (*étapes 1, 2, 3, 4, 5, 7, 8 et 9*) et celles nécessitant également de faire des choix (*étapes 6, 10 et 11*).

Par conséquent, nous avons pu facilement identifier les points d'entrées et points de sorties permettant de mettre en œuvre la méthode des VI suivant le schéma directeur préconisé par la norme. Ainsi, même si les premiers choix apparaissent dès l'étape 6, il n'est pas possible à ce stade de fournir une aide à la décision au concepteur, car le niveau de données minimal pour identifier, estimer et évaluer les risques de dangers pour les opérateurs n'est pas atteint. Ce n'est qu'à l'étape 9 que les premiers résultats pertinents peuvent être donnés au concepteur. La figure 18 illustre ce mécanisme d'intégration.

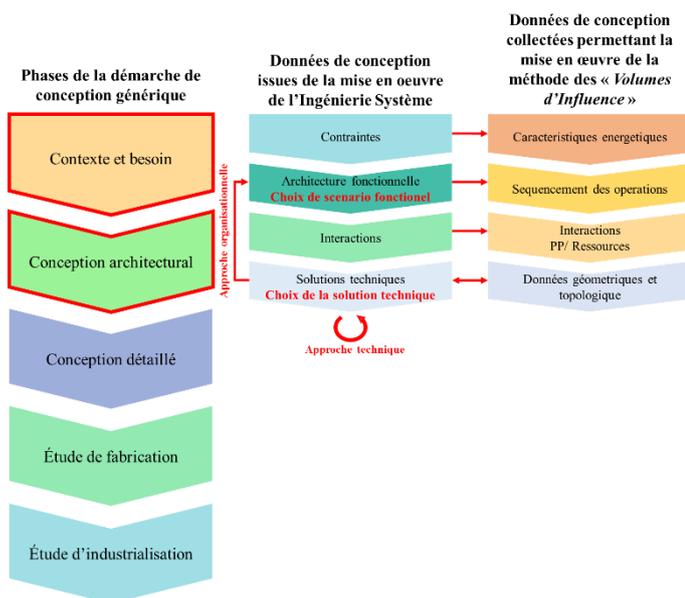


Figure 18. Intégration de la santé/sécurité des opérateurs au processus de conception d'un système

Quant à la figure 19, elle présente un exemple d'identification des risques sur notre étude de cas à travers la mise en œuvre de la méthode des VI.

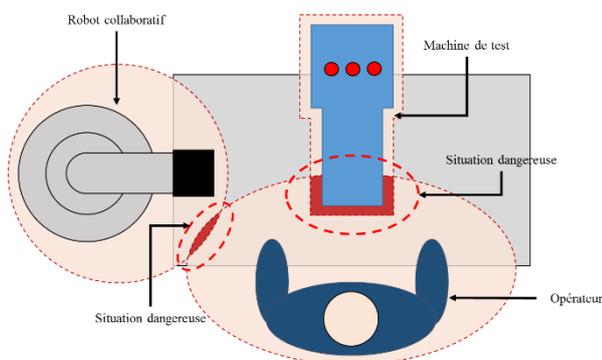


Figure 19. Exemple d'identification des risques de dangers à partir des VI sur notre cas d'étude

À travers cet article, nous avons tout d'abord montré que les usages définis durant la conception d'un système de production constituaient des éléments très importants pour identifier, estimer et évaluer les risques de dangers auxquels sont soumis les opérateurs. En revanche, ils ne permettent pas à eux seuls de mener ce travail à son terme de manière à pouvoir fournir au concepteur une aide à la décision pertinente. Une approche au moindre engagement (parcimonie) nous a permis d'identifier le moment précis où cela était possible.

Mais nous avons également montré qu'il était possible de prendre la santé et la sécurité des opérateurs lors de la conception d'un système de production non pas en utilisant une démarche de conception dédiée, mais en s'appuyant sur une démarche de conception générique, en l'occurrence l'IS. Cette intégration repose sur une structuration pertinente des données produites, la définition de leur nature et de leur niveau de granularité.

Ces deux points essentiels nous amènent aujourd'hui à envisager d'autres travaux complémentaires. Il serait ainsi intéressant d'étudier les points d'intégration possibles de la santé/sécurité des opérateurs dans les autres phases de conception (*conception détaillée, étude de fabrication et d'industrialisation*).

Une autre voie à explorer consisterait à identifier des règles génériques permettant de collecter automatiquement les données nécessaires pour réaliser une analyse de l'impact du système de production en cours de conception sur la santé/sécurité des opérateurs.

Enfin, ces travaux peuvent également être réutilisés pour étudier l'intégration d'autres types d'expertises telle que la prise en compte de l'impact environnementale.

La méthodologie, proposée et la structuration des informations pertinentes associées ont été appliquées sur un cas simple. La mise en œuvre dans des cas industriels nécessitera son informatisation au niveau de la capitalisation et la gestion des données et également par des méthodes avancées de leurs exploitations adaptées au contexte. Mais avant tout, il nous faut finaliser notre cas d'étude en déployant complètement la méthode des « volumes d'influence ».

7. REMERCIEMENTS

Ce projet s'inscrit dans le cadre du laboratoire LC2S (Laboratoire mixte pour la conception sûre de situations de travail) soutenu par l'INRS et le LCFC/ENSAM dont l'objectif est de fournir aux concepteurs d'équipements de travail une démarche structurée et outillée pour passer de la conception de systèmes techniques à la conception des situations de travail pour concilier performance et prévention des risques professionnels.

8. BIBLIOGRAPHIE

- AFNOR, (1991). X 50-151. *Analyse de la Valeur, Analyse Fonctionnelle-Expression fonctionnelle du besoin et cahier des charges fonctionnel*.
- AFNOR, (2009). X 50-153. *Analyse de la valeur- Recommandations pour sa mise en œuvre, European*

standard.

- Bidet-Mayer, T. (2016). Industrie du futur : où en est la France ? *La Fabrique de l'Industrie*. 2-6.
- Cross, N., & Roy, R. (1989). *Engineering design methods* (Vol. 2): Wiley New York.
- Das, B., & Sengupta, A. K. (1996). Industrial workstation design: a systematic ergonomics approach. *Applied ergonomics*, 27(3), 157-163.
- De Galvez, N. (2016). *Sécurité des machines: méthodologie d'identification systématique des phénomènes dangereux en conception* (Thèse de doctorat, ENSAM)
- Evers, V. (2018). De robots y hombres. [Online] *El correo de la UNESCO*, UNESCO. Disponible sur : <https://n9.cl/jtsoi>
- Fadier, E., & De la Garza, C. (2006). Safety design: Towards a new philosophy. *Safety Science*, 44(1), 55-73.
- Faisandier, A. (2011). Formation à l'ingénierie des systèmes-ingénierie des besoins et des exigences. *MAP système*.
- Fiorèse, S. & Meinadier, J.P. (2012). Découvrir et comprendre l'ingénierie système, ed. c. AFIS : *CEPADUES Editions*.
- Gero, J. S. (2003). The situated Function-Behaviour-Structure framework, *Design Studies*, Editions Elsevier.
- Ghemraoui, R. (2009). *Méthodologie de conception innovante intégrant la sécurité des utilisateurs: application aux liaisons tracteur-outils* (Thèse de doctorat, Doctorat Sciences Pratiques Ecole Normale Supérieure de Cachan).
- Godot, X. (2013). *Interactions Projet/Données lors de la conception de produits multi-technologiques en contexte collaboratif* (Thèse de doctorat, Paris, ENSAM).
- Hasan, R. (2002). *Contribution à l'amélioration des performances des systèmes complexes par la prise en compte des aspects socio-techniques dès la conception: proposition d'un modèle original de situation de travail pour une nouvelle approche de conception* (Thèse doctoral, Université Henri Poincaré-Nancy 1).
- Hatchuel, A., & Weil, B. (2003). A new approach of innovative Design: an introduction to CK theory. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design*, Stockholm.
- INRS, (2011). Conception des lieux et des situations de travail. Brochure, ED 950.
- ISO 6385: (2016) Principes ergonomiques de la conception des systèmes de travail
- ISO, N. (2010). 12100–*Sécurité des machines–Principes généraux de conception–Appréciation du risque et réduction du risque*. AFNOR.
- ISO/IEC. (2007). Systems and Software Engineering - Recommended Practice for Architectural Description of Software-Intensive Systems. Geneva, Switzerland.
- ISO/IEC/IEEE 24765, (2009). Systems and Software Engineering - System and Software Engineering Vocabulary Geneva, Switzerland.
- ISO/IEC/IEEE. (2011). Systems and software engineering: Developing user documentation in an agile environment. Geneva, Switzerland.
- ISO/IEC/IEEE. (2015). Systems and Software Engineering - System Life Cycle Processes. Geneva, Switzerland.
- Switzerland.
- LE PARLEMENT EUROPÉEN, E. L. C., & DE L'UNION, E. U. R. O. (2009). *DIRECTIVE 2006/42/CE DU PARLEMENT EUROPÉEN ET DU CONSEIL du 17 mai 2006 relative aux machines et modifiant la directive 95/16/CE (refonte)*.
- Marsot, J. (2001). Prévention et innovation. *Hygiène et sécurité du travail*, 69-75. (Brochure INRS).
- Martin, P., Daille-Lefèvre, B., Marsot, J., Godot, X., Abba, G., Siadat, A., & Gomez-Echeverri, M. (2019). *New issues for workers safety in the factory of the future*. In *Advances on Mechanics, Design Engineering and Manufacturing II*, Springer Nature Switzerland AG, 402-411.
- Østergaard, E. (2018). L'industrie 5.0, une alliée de poids pour la compétitivité de la France. [Online] *Technos et innovations*, L'usine nouvelle. Disponible sur : <https://n9.cl/lwqd>.
- Pahl, G., & Beitz, W. (1996). *Engineering design: a systematic approach*, 2nd edition, Springer-Verlag, London, ENGLAND.
- Paynter, H. M. (1961). *Analysis and Design of Engineering Systems*. MIT Press.
- Scaravetti, D. (2004) *Formalisation préalable d'un problème de conception pour l'aide à la décision en conception préliminaire*, (Thèse de doctorat ENSAM).
- Suh, N. P. (1990). *The principles of design*, Oxford University Press, New-York, USA.
- Ullman, D. G. (2003) *The mechanical design process*, 3rd edition, Mc Graw-Hill Higher Education, New-York, USA.

Application du MBSA avec AltaRica 3.0 dans un cadre système de systèmes pour le transport autonome

MICHEL BATTEUX¹ ET ROMAIN KERNEIS¹

¹ IRT SystemX

8, Avenue de la Vauve – CS 90070, 91127 Palaiseau CEDEX, France
prenom.nom@irt-systemx.fr

Résumé – Un système de mobilité autonome et multimodale, classiquement vu comme un système de systèmes (SdS), implique plusieurs systèmes : e.g. des voitures et navettes autonomes et connectés, des équipements d’infrastructures routières connectés, des centres de téléopération, etc. De tels systèmes doivent interagir ensemble afin d’assurer une mission principale de transport de personnes. L’étude de la sûreté de fonctionnement d’un tel SdS est des plus importantes. D’une part car elle implique l’évaluation de la sûreté des véhicules autonomes, qui représente un véritable défi technologique en soit. D’autre part car les interactions entre ces différents systèmes multiplient les potentiels dysfonctionnements. L’utilisation d’une approche dite basée modèles (MBSA), sous-entendues modèles de ‘haut-niveau’ proches des architectures fonctionnelles et organiques, rencontre une attention de plus en plus forte. Le langage AltaRica 3.0 est une telle solution MBSA. Cette publication montre l’utilisation de ce langage pour évaluer la sûreté de fonctionnement de ce SdS. Le modèle AltaRica 3.0 est construit en fonction d’événements redoutés issus d’une analyse préliminaire de risques et des architectures fonctionnelles et organiques du SdS. Ce modèle permet ensuite d’analyser les coupes minimales et leurs taux d’occurrences, en fonction de critères d’évaluation de sûreté spécifiés (ordre, taux d’occurrence, etc.).

Abstract – A multi-modal and autonomous mobility solution for users involves several systems (e.g. autonomous vehicles and shuttles communicating with their environment, traffic signals or GPS, a remote control center, etc.), interacting all together to ensure the main mission of passenger transportation. Such combination of several different systems is classically called a system of systems (SoS). The safety analysis of such a system is one of the major actual challenges. In fact, not only it implies the safety evaluation of the autonomous vehicles, which represents a real technological challenge. But also interactions between the different systems of the SoS increase potential failures. The use of a model based approach (MBSA), meaning ‘high-level’ models closed to functional and physical architectures, is an appropriate solution. The AltaRica 3.0 modeling language is such a MBSA solution. This paper presents the use of this language to assess the safety of these SoS. The AltaRica 3.0 model is designed according to a functional failure analysis and the functional and physical architectures. This model is used to produce minimal cut sets which are then analyzed according to specified safety criteria (order, occurrence rate, etc.).

Mots clés - MBSA, AltaRica, Système de Systèmes, Transport Autonome, ISO 26262, SOTIF.

Keywords – MBSA, AltaRica, System of Systems, Autonomous Transportation, ISO 26262, SOTIF.

1 INTRODUCTION

Un système de mobilité autonome et multimodale implique plusieurs systèmes, par exemple des voitures et navettes autonomes et connectés, des équipements d’infrastructures routières connectés, des centres de téléopération, etc. De tels systèmes doivent interagir ensemble afin d’assurer une mission principale de transport de personnes. Une telle combinaison de différents systèmes est considéré comme un ‘système de systèmes’ (SdS) ([Maier, 2005]), et est conçue suivant un ensemble d’exigences, qui peuvent ne pas être totalement en cohérence les unes des autres. En effet, la conception d’un tel SdS nécessite non seulement de prendre en compte les exigences des différentes parties-prenantes, mais aussi d’utiliser des systèmes, qui sont déjà conçus suivant des exigences pas nécessairement en cohérence avec les exigences du SdS. Dans ce cadre, il devient de plus en plus nécessaire des considérer des approches dites ‘formelles’ de modélisation, c’est-à-dire basées sur des langages de modélisation rigoureusement définis.

Les études de sûreté de fonctionnement dites basée modèles, MBSA pour Model-Based Safety Assessment, sous-entendues modèles de ‘haut-niveau’ rencontrent une attention de plus en

plus forte. Les modèles MBSA sont conçus au moyen de langages de modélisation permettant de les appréhender plus simplement, au sens qu’ils sont proches des architectures fonctionnelles et organiques rencontrés dans les démarches d’architecture système. AltaRica 3.0 est une telle solution de langage de modélisation MBSA et permet de réaliser des études de sûreté de fonctionnement, ainsi que d’assurer un lien étroit avec les descriptions d’architecture système.

Dans cette publication, nous montrons l’utilisation du langage AltaRica 3.0 pour étudier la sûreté de fonctionnement d’un SdS de mobilité autonome et multimodale. Nous montrons comment, en nous basant sur les descriptions d’architecture système, nous utilisons les spécifications de conception de ce SdS pour mener les études de sûreté de fonctionnement.

La suite de cette publication est organisée comme suit. La deuxième partie présente le SdS de mobilité autonome et multimodale considéré, en abordant assez brièvement les descriptions d’architecture système utilisées comme spécifications. La troisième partie présentera le langage AltaRica 3.0, ainsi que la plateforme OpenAltaRica permettant de réaliser des études de sûreté de fonctionnement. La quatrième partie montrera l’utilisation de AltaRica 3.0 pour l’étude de sûreté de fonctionnement du SdS de mobilité autonome et

multimodale. Enfin la cinquième et dernière partie conclura cette publication.

2 SPECIFICATION DU SYSTEME DE SYSTEMES DE MOBILITE AUTONOME ET MULTIMODALE

Nous considérons un SdS de mobilité autonome et multimodale, comme représenté en Figure 1. Cette figure ne représente qu'une certaine 'vue' du SdS. Ce SdS est constitué de différents véhicules autonomes, qu'ils soient des voitures ou des navettes, qui roulent dans un environnement ouvert pour transporter des personnes d'un endroit à un autre. L'environnement est considéré ouvert dans le sens où les voies de circulation sont ouvertes, avec néanmoins certaines restrictions que nous n'aborderons pas dans cette publication. Le SdS consiste également en un ensemble de composants de communication avec les véhicules autonomes : les états des feux tricolores qui sont dits 'connectés', des antennes relais (nommées 'Balise RTK') ou des satellites (nommés 'GNSS') qui fournissent des indications de positions, des informations sur le trafic via des caméras (nommées 'RSU'). Enfin un centre de téléopération est également considéré. Il permet d'une part de coordonner les demandes de transports des personnes, de fournir de nouveaux chemins, ou même de reprendre en main, si nécessaire, les véhicules autonomes en cas de problèmes.

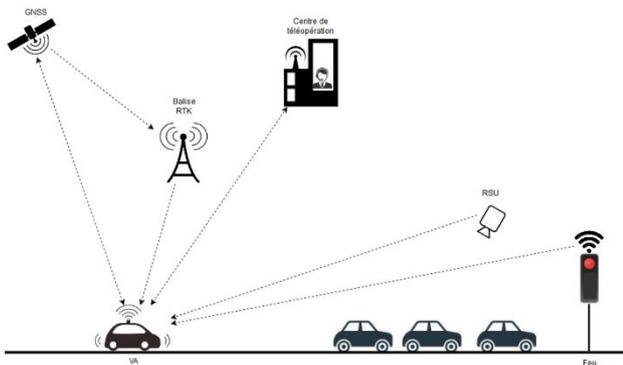


Figure 1 : Système de systèmes de mobilité autonome et multimodale

La spécification de ce SdS de mobilité autonome et multimodale a été réalisée en utilisant des méthodes, plus précisément des cadres d'architecture, utilisant des diagrammes descriptifs de spécification. Les architectures opérationnelle, fonctionnelle et organique ont été réalisées au moyen de diagrammes d'interactions par l'outil Capella. Les dynamiques associées à ces diagrammes fonctionnel et organique ont été spécifiées au moyen de l'outil CPNTools. Enfin ces travaux de spécification ont produit un ensemble d'exigences, dont certaines spécifiant du point de vue de la sûreté de fonctionnement ([Addoui et al., 2020]). Notre étude MBSA est donc réalisée à partir de ces éléments de spécification, notamment issus des niveaux des architectures opérationnelles, fonctionnelles et organiques. En effet, et nous le verrons en partie 4.1, les études de la sûreté de fonctionnement s'initieront par l'analyse des 33 fonctions du SdS, et se déploieront ensuite sur les différents systèmes, voire parties de systèmes, du SdS.

3 ALTARICA 3.0 ET OPENALTARICA

Dans cette partie nous présentons le langage de modélisation AltaRica 3.0, ainsi que les outils associés d'évaluation au sein de la plateforme OpenAltaRica.

3.1 Le Langage AltaRica 3.0

AltaRica 3.0 est un langage de modélisation événementiel et orienté objet dédié, initialement, aux analyses probabilistes du risque de systèmes complexes ([Batteux et al., 2019-1]). Ce langage permet de modéliser les systèmes à plus haut niveau que les formalismes classiques d'analyse du risque (arbres de défaillance, chaînes de Markov, réseaux de Pétri stochastiques, etc.), sans augmenter la complexité des calculs d'indicateurs sous-jacents du risque. AltaRica 3.0 est la combinaison du cadre mathématique GTS, pour Guarded Transitions System ([Batteux et al., 2017]), et du paradigme de structuration S2ML, pour System Structure Modeling Language ([Batteux et al., 2018-1]).

Le cadre mathématique GTS généralise les formalismes probabilistes états/transitions classiques, et permet de rendre compte des aspects comportementaux que l'on souhaite modéliser. Il s'agit de la sémantique d'exécution des modèles. Cette exécution reste similaire aux autres formalismes à événements discrets dans le sens où chaque fois qu'une transition est tirable, elle est planifiée et sera potentiellement tirée après un certain délai ([Cassandras et al., 2008]) et [Zimmermann, 2008]). Ces délais peuvent être déterministes ou stochastiques, et le langage fournit les délais usuels du type Dirac, exponentiel, Weibull, etc., et des délais empiriques.

Le paradigme de structuration S2ML permet de structurer les modèles. Il unifie les deux principaux paradigmes (de structuration) dominants des langages de modélisation : l'orienté objet à classe et l'orienté objet à prototype. S2ML permet de modéliser suivant deux approches combinables. L'approche dite 'top-down', avec une vision au niveau système permettant la réutilisation de schémas de modélisation, et donc utilisant le paradigme orienté objet à prototypes. L'approche dite 'bottom-up', avec une vision au niveau composants permettant la réutilisation de composants, définis en bibliothèques, et donc utilisant le paradigme orienté objet à classes.

La Figure 2 présente un modèle AltaRica 3.0 représentant deux composants 'TrafficLight' et 'GNSS' qui fournissent des informations de communication à un troisième composant 'AVehicle'.

```

class NRComponent
  Boolean _working (init = true);
  parameter Real lambda = 1.0e-6;
  event failure (delay = exponential(lambda));
  transition
    failure: _working -> _working := false;
  Boolean in, out (reset = false);
  assertion
    out := _working and in;
end

block System
  NRComponent AVehicle (lambda = 1.0e-7);
  NRComponent TrafficLight (lambda = 1.0e-5);
  NRComponent GNSS (lambda = 1.0e-8);
  assertion
    TrafficLight.in := true; GNSS.in := true;
    AVehicle.in := TrafficLight.out or GNSS.out;
  observer Boolean ComLost =
    AVehicle.out == false;
end

```

Figure 2 : Exemple de modèle AltaRica 3.0

Les deux principales constructions structurelles du langage, et qui sont utilisées dans ce modèle sont les 'block' et les 'class'. Elles correspondent aux éléments des formalismes orientés objet à classes et à prototypes. Un élément 'class' représente une construction structurelle définissant un composant générique.

Un élément 'class' est d'abord déclaré en dehors de tout autre élément, comme montré avec la déclaration de l'élément 'class' 'NRComponent', puis est ensuite utilisée par héritage ou par instanciation via les instances 'TrafficLight', 'GNSS' et 'AVehicle'. Un élément 'block' représente une construction structurelle ayant une unique occurrence. Un élément 'block' est déclaré et directement utilisé, comme montré avec l'élément 'block' 'System'. Il n'est pas possible d'instancier un élément 'block', mais il est néanmoins possible de le cloner afin d'en faire une copie.

Du point de vue comportemental, l'élément 'class' 'NRComponent' représente un composant qui est en marche et qui peut tomber en panne suite à l'occurrence de l'événement panne. Ce comportement est représenté par la machine à états de la Figure 3. Une variable d'état booléenne '_working' est définie avec pour valeur initiale 'true'. Cette variable représente l'état du composant et passera de cette valeur 'true' à la valeur 'false' par la transition, labélisée par l'événement 'failure' défini avec un délai exponentiel. Les informations de communication sont définies au moyen de deux variables de flux 'in' et 'out'. La valeur de la variable 'out', qui représente la sortie de 'NRComponent', est mise à jour dans l'assertion en fonction de la valeur de variable d'état '_working' et de la valeur de la variable 'in' qui représente l'entrée de 'NRComponent'. Il est bien sûr possible de considérer des comportements beaucoup plus sophistiqués, mettant en œuvre de variables de types énuméré, entier ou réel, ainsi que des transitions modélisant des caractéristiques fonctionnelles, de reconfigurations, de maintenances, etc.

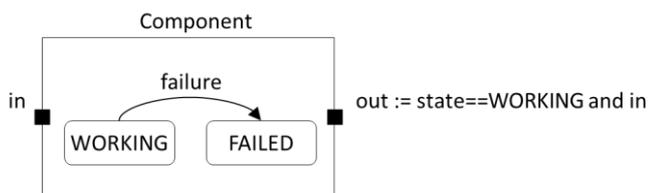


Figure 3 : Comportement d'un objet 'Component'

Au niveau de l'élément 'block' 'System', l'assertion représente les échanges de flux entre les trois instances 'TrafficLight', 'GNSS' et 'AVehicle'.

3.2 La plateforme OpenAltaRica

La plateforme OpenAltaRica implémente le langage AltaRica 3.0 et fournit un ensemble d'outil d'évaluation pour étudier ces modèles, comme représenté en Figure 4. Elle est disponible en téléchargement, sur le site du projet associé, via le lien <https://www.openaltarica.fr/docs-downloads/>. Cette plateforme permet d'éditer les modèles AltaRica 3.0, puis de les évaluer grâce à différents outils, ou chaînes d'outils. L'évaluation de défaillances qui utilise la compilation d'un modèle AltaRica 3.0 en un arbre de défaillances, puis l'évaluation de cet arbre par le moteur de calcul XFTA. La génération de séquences critiques d'événements, qui permet de prendre en compte les aspects dynamiques des modèles AltaRica 3.0, ce qui n'est pas le cas pour les arbres de défaillances. Un simulateur pas-à-pas qui permet de jouer, à la main, des scénarios d'exécutions du modèle. Enfin un simulateur stochastique, outil polyvalent qui permet d'évaluer n'importe quel modèle quel que soit sa taille. Plus d'informations sur cette plateforme OpenAltaRica peuvent être trouvées dans ([Batteux et al., 2018-2]).

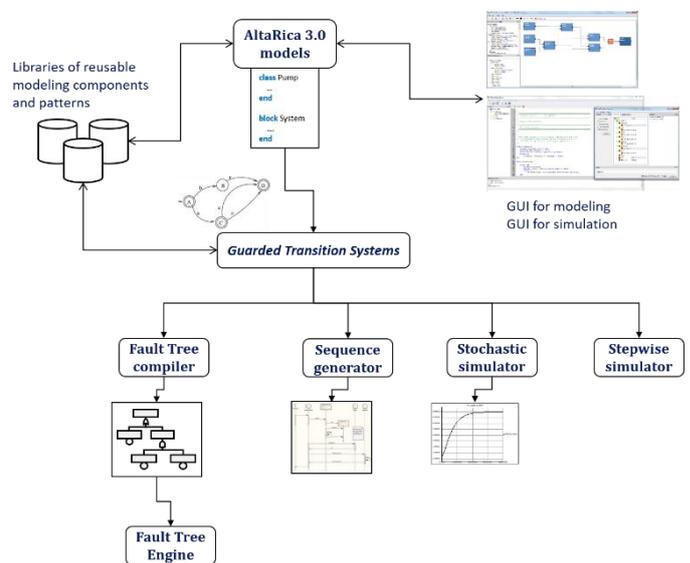


Figure 4 : Vision outils de la plateforme OpenAltaRica

4 ETUDE DE LA SURETE DE FONCTIONNEMENT DU SYSTEME DE SYSTEMES DE MOBILITE AUTONOME ET MULTIMODALE

Dans cette partie, nous décrivons l'étude de sûreté de fonctionnement du SdS de mobilité autonome et multimodale réalisée. Nous avons commencé par produire une analyse préliminaire des risques, basée sur les 33 grandes fonctions spécifiées du SdS. Cette analyse préliminaire des risques nous a produit un ensemble d'événements redoutés à prendre en compte. Ensuite pour ces événements redoutés à prendre en compte, et qui sont les plus critiques, nous avons étudié qualitativement et quantitativement leurs potentielles occurrences au moyen de modèles MBSA en AltaRica 3.0. Dans le cadre de cette publication, nous nous limiterons à l'étude de deux événements redoutés.

4.1 Analyse préliminaire de risques

L'analyse préliminaire des risques, dénoté APR, est une méthode classique permettant d'extraire les événements redoutés du système considéré, le SdS de mobilité autonome et multimodale dans notre cas. Ces événements redoutés sont définis en fonction des défaillances des (grandes) fonctions du système. Nous avons pour cela repris les éléments de l'architecture fonctionnelle, l'ensemble des 33 grandes fonctions du SdS et les avons étudiées suivant les modes associés de défaillance : perte de la fonction, ou absence de la fonction, ou fonctionnement intempestif de la fonction, ou encore fonctionnement dégradé de la fonction. Nous avons ensuite coté ces modes de défaillances des fonctions suivant le principe ASIL de la norme ISO 26262 ([ISO 26262, 2011]) : niveau QM ne demandant pas de mesures "safety", ainsi que les niveaux ASIL A à ASIL D. Ces différents niveaux expriment les risques associés à ces modes de défaillances des fonctions. Un tel risque est donc une mesure qui s'exprime suivant la sévérité du mode de défaillance, son exposition, et sa contrôlabilité.

- La sévérité exprime la gravité des conséquences en cas d'accident, et se décline en 4 classes suivant les préjudices associés (au sens blessures ou pertes humaines) : aucun préjudice, préjudice léger et modéré, préjudice sévère engageant le pronostic vital mais avec survie probable, et préjudice grave ou fatal engageant le pronostic vital avec survie incertaine.

- L'exposition exprime la probabilité de la présence de la situation extérieure, et se décline en 5 classes de taux d'exposition suivant la situation opérationnelle : invraisemblable, très faible avec taux inférieur à 0.1%, faible avec taux entre 0.1% et 1%, moyen avec taux entre 1% et 10%, fort avec taux supérieur à 10%.
- Enfin la contrôlabilité exprime la capacité des différents acteurs de la scène à garder le contrôle en cas de défaillance, et se décline en 4 classes : tout le temps contrôlable, généralement contrôlable (supérieure ou égale à 99%), souvent contrôlable (entre 90% et 99%), et difficilement contrôlable ou incontrôlable (inférieure à 90%).

Nous avons ensuite considéré 4 événements redoutés, correspondant à 4 modes de défaillances de fonctions : « Indisponibilité de la téléopération », « Donnée erronée sur le chemin envoyé par la téléopération », « Perte de la détection des feux connectés », ainsi que « Perte de la détection de bouchons ».

Dans le cadre de cette publication, nous présentons la poursuite de l'étude avec les événements redoutés : perte de la détection des feux connectés et perte de la fonction de détection de bouchons, qui ont une cotation ASIL C. Différents scénarios peuvent être considérés pour ces événements redoutés, suivant la configuration de l'environnement. Pour la perte de la fonction de détection de bouchons, nous considérons un véhicule autonome en approche d'un embouteillage, qui reçoit sa position par satellite et renforcée par un balise. Ces deux données peuvent être perdues que ce soit à cause de l'environnement, par exemple un tunnel, ou soit à cause de l'émetteur. Nous estimons aussi que le bouchon se situe proche d'une caméra capable de le détecter et de communiquer son emplacement au véhicule autonome. Pour la perte de la détection des feux connectés, nous considérons un véhicule autonome en approche d'un feu connecté et qui reçoit l'état du feu par celui-ci.

4.2 Modèle AltaRica 3.0

Le modèle AltaRica 3.0 associé, pour l'étude des deux événements redoutés, est construit par combinaison et constitution des différents composants du SdS ; pour cette communication principalement ceux nécessaires à l'étude de ces événements redoutés. Dans ce modèle, non seulement nous considérons les défaillances des composants, mais aussi les dysfonctionnements de perception en l'absence de défaillances matérielles ou logicielles. Dans ce second cas, ces dysfonctionnements ont pour principales causes l'interprétation des données, par exemple une reconnaissance de forme erronée, ou encore l'interaction avec l'environnement, par exemple l'éblouissement causé par la neige, ou même la combinaison des deux. Ce type de dysfonctionnements est issu de [ISO/WD PAS 21448, 2019], aussi appelé SOTIF (pour "Safety Of The Intended Functionality"). Dans le cadre du transport autonome, il n'existe pas, à l'heure actuelle, de référentiel fournissant des données complètes donnant des informations qualitatives et quantitatives liées aux défaillances des composants. Nous considérons donc qu'une défaillance d'un composant est une panne et qu'elle provoque la perte du composant : c'est-à-dire qu'il ne fonctionne plus. Concernant les valeurs quantitatives, liées à ces défaillances, nous expliciterons les hypothèses prises lors de l'analyse du modèle en partie 4.3.1.

Nous commencerons par présenter la structure globale de ce modèle, puis nous spécifierons les définitions des différents composants. Bien évidemment cette publication ne fait qu'une description du modèle AltaRica 3.0, il ne s'agit pas du modèle

tel que défini dans la plateforme OpenAltaRica. En effet cette plateforme permet de gérer de tels modèles avec une organisation en plusieurs fichiers, suivant les différents constituants du modèle, comme décrit dans [Batteux et al., 2018-2].

4.2.1 Structure globale

La modélisation de la structure globale du SdS est présentée en Figure 5. Cette partie décrit la déclaration des systèmes, les interactions entre eux, et enfin les observateurs permettant de définir les événements redoutés.

```

block SoS
  block GNSS /* behavior */ end
  clones GNSS as RTK;
  block RSU /* behavior */ end
  block TLight /* behavior */ end
  block AVehicule /* behavior */ end
  assertion
    AVehicule.V2X.inRSU := RSU.out;
    AVehicule.V2X.inTLight := TLight.out;
    AVehicule.RTK.in := RTK.out;
    AVehicule.GNSS.inGNSS := GNSS.out;
  observer Boolean TJDetectionLost =
    AVehicule.Decision.outTJ == false;
  observer Boolean TLDetectionLost =
    AVehicule.Decision.outTL == false;
end

```

Figure 5 : Modélisation de la structure globale du SdS

Cette structure globale définit le SdS dans un élément 'block' principal 'SoS'. Cet élément 'block' principal déclare les 5 systèmes constituant du SoS : trois éléments 'block' 'GNSS' pour le satellite GNSS, 'RSU' pour la caméra (externe), 'TLight' pour le feu connecté et 'AVehicule' pour le véhicule autonome. Ces éléments sont définis dans les parties 4.2.2 et 4.2.3 suivantes. Un clone, c'est-à-dire une copie, 'RTK' pour la balise RTK, de l'élément 'block' 'GNSS' est également déclaré. L'assertion définit les interactions entre ces systèmes, principalement comment le véhicule autonome réceptionne les données issues des autres systèmes, ce qui sera explicité en partie 4.2.3.

4.2.2 Composants de communication

La Figure 6 présente la modélisation du système satellite 'GNSS'. Il s'agit d'un 'block' assez simple, qui hérite de la classe 'NRComponent', définie en Figure 2, et permettant d'obtenir le comportement de composants non-réparable. La communication de ce satellite est donnée par la variable 'out' qui peut prendre les deux valeurs 'OK' et 'KO' définies dans le domaine 'WorkingMode'. Nous souhaitons dans ce cas abstraire le fait que le satellite envoie un signal lorsqu'il fonctionne (valeur 'OK') et n'envoie pas de signal lorsqu'il est en panne ; ce signal étant réceptionné par le véhicule autonome.

```

domain WorkingMode {OK, KO}

block GNSS
  extends NRComponent;
  WorkingMode out (reset = OK);
  assertion
    out := if _working then OK else KO;
end

```

Figure 6 : Modélisation du système 'GNSS'

La Figure 7 présente la modélisation du système caméra 'RSU'. Une telle caméra est un élément situé dans la voirie et qui permet d'informer des véhicules autonomes sur l'état de la circulation. Ici aussi il s'agit d'un 'block' qui hérite de la classe 'NRComponent'. Une telle caméra peut aussi envoyer une

information erronée, et pour cela une variable d'état '_erroneous', un événement 'erroneous', et une transition associée, sont définis. Enfin la communication est donnée par la variable 'out' qui peut prendre les trois valeurs 'VEHICLES', 'NOVEHICLE' et 'KO' définies dans le domaine 'RSUMessage'. Dans ce cas la caméra envoie la présence de véhicules dans le cas où elle fonctionne bien, elle envoie l'absence de véhicules dans le cas où elle est en mode erroné, enfin elle n'envoie pas de message dans le cas où elle est en panne. Nous remarquons bien dans ce composant le fait qu'un comportement erroné soit pris en compte, en plus du comportement de panne dérivant de la classe 'NRComponent'; ce qui retranscrit les éléments normatifs du SOTIF. Il s'agit donc de comportements complémentaires, dits fonctionnels, du composant qui sont pris en compte.

```

domain RSUMessage {VEHICLES, NOVEHICLE, KO}

block RSU
  extends NRComponent;
  Boolean _erroneous (init = false);
  event erroneous (delay = exponential(1.0e-4));
  transition
    erroneous: not _erroneous and _working
      -> _erroneous := true;
  RSUMessage out (reset = VEHICLES);
  assertion
    out := if _working then
      if not _erroneous then VEHICLES
      else NOVEHICLE
    else KO;
end

```

Figure 7 : Modélisation du système 'RSU'

La Figure 8 présente la modélisation du système feu connecté 'TLight'. Au renommage près des valeurs de la variable 'out', le comportement de ce système est équivalent à celui du système caméra 'RSU' précédemment présenté.

```

domain TLMesssage {CORRECT, ERRONEOUS, KO}

block TrafficLight
  extends NRComponent;
  Boolean _erroneous (init = false);
  event erroneous (delay = exponential(1.0e-4));
  transition
    erroneous: not _erroneous and _working
      -> _erroneous := true;
  TLMesssage out (reset = CORRECT);
  assertion
    out := if _working then
      if not _erroneous then CORRECT
      else ERRONEOUS
    else KO;
end

```

Figure 8 : Modélisation du système 'TrafficLight'

4.2.3 Le véhicule autonome

Le modèle du véhicule autonome ne contient que les composants de perception de l'environnement. Nous ne considérons pas les éléments liés au contrôle du véhicule. De plus le modèle montré et discuté dans cette publication n'est pas le modèle global ayant servi à l'étude complète ; nous n'avons considéré que les éléments nécessaires à la publication. Par exemple le composant 'V2X Component' reçoit des informations, au sens communications, de beaucoup de systèmes externes au véhicules : caméra externes, feu connecté, etc.. Nous n'avons considéré que les communications, avec ces systèmes externes, nécessaires à l'étude des deux événements redoutés.

La Figure 9 ci-dessous représente l'ensemble de ces composants, ainsi que les liens entre ceux-ci. Il est à noter que

les trois composants 'V2X Component', 'RTK Component' et 'GNSS Component' sont les composants qui réceptionnent les données issues des systèmes externes définie précédemment : 'GNSS' pour le satellite GNSS, 'RSU' pour la caméra, 'TLight' pour le feu connecté et 'RTK' pour la balise RTK.

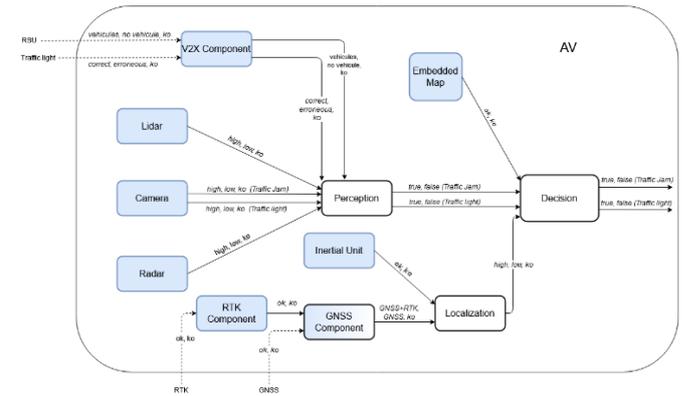


Figure 9 : Composants internes d'un véhicule autonome

La modélisation de la structure globale du véhicule autonome est présentée en Figure 10. Cette partie du modèle correspond au schéma descriptif de la Figure 9, et décrit la déclaration des composants de perception de l'environnement et leurs interactions. Dans la suite nous ne décrirons que quelques composants, ceux ayant les comportements les plus compliqués.

Les composants 'Map', 'Odometer', 'RTK', 'V2X' ne sont donc pas décrits car ils ont des comportements simples. Ils héritent de la classe 'NRComponent' décrite en Figure 2. 'Map' et 'Odometer' n'ont qu'une sortie, ils envoient donc un signal 'OK' s'ils fonctionnent bien et 'KO' sinon. 'RTK' et 'V2X' ont en plus des entrées : une entrée issue de la balise RTK pour 'RTK', et deux entrées issues de la caméra externe et du feu connecté pour 'V2X'. Ils renvoient ces entrées s'ils fonctionnent bien et 'KO' sinon.

```

block AVehicle
  block Map /* behavior */ end
  block Lidar /* behavior */ end
  clones Lidar as Camera;
  clones Lidar as Radar;
  block V2X /* behavior */ end
  block Perception /* behavior */ end
  block Odometer /* behavior */ end
  block RTK /* behavior */ end
  block GNSS /* behavior */ end
  block Localization /* behavior */ end
  block Decision /* behavior */ end
  assertion
    Perception.inLidar := Lidar.out;
    Perception.inV2XRSU := V2X.outRSU;
    Perception.inV2XTLight := V2X.outTLight;
    Perception.inRadar := Radar.out;
    Perception.inCameraTJam := Camera.outTJam;
    Perception.inCameraTLight := Camera.outTLight;
    GNSS.inRTK := RTK.out;
    Localization.inGNSS := GNSS.out;
    Localization.inOdo := Odometer.out;
    Decision.inMap := Map.out;
    Decision.inPerception := Perception.out;
    Decision.inLocalization :=
      Localization.out;
end

```

Figure 10 : Modélisation de la structure globale du véhicule autonome

La Figure 11 présente la modélisation des composants de type capteur : le 'Lidar', la 'Camera' et le 'Radar'. Comme montré

en Figure 10, seul le composant ‘Lidar’ est défini dans un élément ‘block’; les composants ‘Camera’ et ‘Radar’ sont clonés de ce composant ‘Lidar’. Les capteurs prennent en compte la confiance, plus ou moins bonne, de la perception d’un objet, comme définie dans SOTIF. Il s’agit donc de comportements complémentaires, dits fonctionnels, du composant qui sont pris en compte. La classe ‘Sensor’ hérite de la classe ‘NRComponent’, définie en Figure 2, et ajoute les éléments comportementaux permettant d’appréhender les perceptions plus ou moins fiables. Le composant ‘Lidar’ hérite donc de cette classe ‘Sensor’ et y ajoute le comportement externe, via la variable ‘out’, suivant les états internes de marche ou de panne, et de perception.

Concernant le composant ‘Camera’, nous faisons l’hypothèse que la prise en compte d’un bouchon ou d’un feu se réalise de la même manière. Cette simplification est possible car les événements redoutés sont ensuite étudiés indépendamment les uns des autres.

```

domain Confident {HIGH, LOW, KO}

class Sensor
  extends NRComponent;
  Confident _perception (init = HIGH);
  event high (delay = exponential(1.0e-4));
  event low (delay = exponential(1.0e-4));
  transition
    high: _working and _perception != HIGH
      -> _perception := HIGH;
    low: _working and _perception != LOW
      -> _perception := LOW;
end

block Lidar
  extends Sensor;
  Confident out (reset = HIGH);
  assertion
    out := if _working then _perception else KO;
end

```

Figure 11 : Modélisation des composants de type capteur

La Figure 12 présente la modélisation du composant ‘GNSS’. Ce composant hérite de la classe ‘NRComponent’. Il prend deux entrées, liées au satellite ‘GNSS’ et le composant ‘RTK’, et fournit en sortie les bonnes réceptions de ces entrées combinées avec son état.

```

domain GNSSMessage {GNSS_RTK, GNSS, KO}

block GNSS
  extends NRComponent;
  WorkingMode inRTK, inGNSS (reset = OK);
  GNSSMessage out (reset = GNSS_RTK);
  assertion
    out := if _working then
      if ((inRTK == OK) and (inGNSS == OK))
        then GNSS_RTK
      else
        if (inGNSS == OK) then GNSS
        else KO;
      end
    end
end

```

Figure 12 : Modélisation du composant GNSS

Les composants ‘Localization’, ‘Decision’ et ‘Perception’, présenté en Figure 13, sont des composants logiques. Le composant ‘Localization’ fournit en sortie la confiance liée aux deux entrées issues des composants ‘Odometer’ et ‘GNSS’. Le composant ‘Decision’ agrège les entrées des composants ‘Perception’, ‘Localization’ et ‘Map’. Enfin le composant ‘Perception’ agrège les entrées des

composants ‘Lidar’, ‘Camera’, ‘Radar’, et ‘V2X’, afin de fournir en sortie les informations sur la présence ou non de véhicules devant, ainsi que le feu connecté. Ces agrégations suivent les tables de vérités données en Table 1 et Table 2. La Table 1 caractérise un vote majoritaire à 3, avec un choix le plus conservatif en cas d’égalité à 2.

```

block Localization
  GNSSMessage inGNSS (reset = GNSS_RTK);
  WorkingMode inOdo (reset = OK);
  Confident out (reset = HIGH);
  assertion
    out := if ((inGNSS == GNSS_RTK)
      and (inOdo == OK)) then HIGH
    else
      if ((inGNSS == KO) and (inOdo == KO))
        then KO else LOW;
    end
end

block Decision
  WorkingMode inMap (reset = OK);
  Boolean inPerception (reset = true);
  Confident inLocalization (reset = HIGH);
  Boolean out (reset = true);
  assertion
    out := if ((inMap == OK) and
      (inPerception == true) and
      ((inLocalization == HIGH) or
      (inLocalization == LOW)))
    then true else false;
end

block Perception
  Confident inLidar, inCamera, inRadar
    (reset = HIGH);
  RSUMessage inV2XRSU (reset = VEHICLES);
  TLMMessage inV2XTLight (reset = CORRECT);
  Boolean outTJam, outTLight
    (reset = true);
  assertion
    outTrafficJam := // done by Table 1
    outTrafficLight := // done by Table 2
end

```

Figure 13 : Modélisation des composants Localization et Decision et Perception

Table 1 : Extrait de la table de vérités de la perception pour la sortie détection de bouchons

inLidar	inCamera	inRadar	inV2XRSU	outTJam
HIGH	HIGH	HIGH	VEHICLES	true
HIGH	HIGH	HIGH	NOVEHICLE	true
HIGH	HIGH	HIGH	KO	true
HIGH	HIGH	LOW	VEHICLES	true
HIGH	HIGH	LOW	NOVEHICLE	false
HIGH	HIGH	LOW	KO	true
HIGH	HIGH	KO	VEHICLES	true
HIGH	HIGH	KO	NOVEHICLE	true
HIGH	HIGH	KO	KO	true
...

Table 2 : Table de vérités de la perception pour la sortie détection de feux connectés

inCamera	inV2XTLight	outTLight
HIGH	CORRECT	true
LOW	CORRECT	false
KO	CORRECT	true
HIGH	ERRONEOUS	true
LOW	ERRONEOUS	false
KO	ERRONEOUS	false
HIGH	KO	true
LOW	KO	false
KO	KO	false

4.3 Evaluation du modèle AltaRica 3.0

Dans cette partie, nous allons évaluer le modèle AltaRica 3.0 conçu pour notre étude et lié aux événements redoutés de perte

de la fonction de détection de bouchons et de perte de la fonction de détection de feu. Nous allons d'abord proposer des données quantitatives justifiées sur les exigences de ces événements redoutés, et sur les défaillances des composants. Nous mènerons ensuite l'étude via la génération d'un arbre de défaillances.

4.3.1 Données quantitatives

Comme déjà indiqué il n'existe pas, à l'heure actuelle, de référentiel fournissant des données quantitatives liées aux défaillances des composants considérés pour notre SdS de mobilité autonome et multimodale. Nous avons donc définis ceux-ci suivant des critères empiriques.

En premier lieu, nous considérons le fonctionnement de notre SdS sur 5 heures. En effet cela traduit une borne maximum de durée de fonctionnement pour un usage normal du SdS par un seul groupe d'utilisateurs. Dans ce cas, nous souhaitons que les événements redoutés de perte de la fonction de détection de bouchons, soit au moins à un taux d'occurrence égale à ce qui se passe actuellement pour un accident avec un véhicule léger : taux d'occurrence de $8.46e-7$ (bilan de la sécurité routière de 2018). Nous définissons donc le taux d'occurrence de l'événement redouté à $8.46e-7$.

Ensuite concernant les défaillances des composants, nous estimons que les défaillances et erronés des composants suivent une loi exponentielle. Les défaillances sont exprimées en 1 défaillance toute les X heures de fonctionnement (1/Xh). Ces taux ont été définis de façon empirique, et en fonction de la nouveauté et/ou de la connaissance des technologies utilisées. Nous avons choisi trois catégories progressives de taux de défaillances :

- Composants nouveaux relatifs à l'expérimentation, avec un taux de 1/10 000 ;
- Composants relatifs aux véhicules autonomes, avec un taux de 1/20 000 ;
- Composants maîtrisés avant les véhicules autonomes, avec un taux de 1/30 000.

La Table 3 suivante associe ces taux aux éléments (systèmes et composants) du SdS.

Table 3 : Association des taux aux éléments du SdS

		RSU	
		1/10 000	RTK
Composant VA MAP			
Composant VA RTK			
1/20 000	Composant VA Caméra		
	Composant VA Lidar		
	Composant VA Radar		
1/30 000	GNSS		
	Composant VA Odomètre		
	Composant VA GNSS		
Erroné	1/20 000	RSU	
		Composant VA Caméra	
		Composant VA Lidar	
		Composant VA Radar	

4.3.2 Génération et analyses par arbres de défaillances

Pour évaluer ce modèle AltaRica 3.0, nous avons utilisé la chaîne d'outils, de la plateforme OpenAltaRica, liée à l'analyse par arbres de défaillances. Pour cela le modèle est d'abord mis à plat, c'est-à-dire transformé en un format intermédiaire qui permet d'être utilisé par tous les moteurs de calculs de la plateforme. À la suite de cette mise à plat, le compilateur vers les arbres de défaillances a été utilisé afin de transformer le modèle en arbres de défaillances. Dans le modèle AltaRica 3.0, les événements redoutés, qui représentent les pertes des fonctions de détection de bouchons et de feux connectés, sont exprimés par les observateurs Booléen (voir Figure 5) qui

considèrent les sorties du composant 'Decision' du véhicule autonome évaluées à la valeur 'false'. Ils se retrouvent dans les arbres de défaillance par les événements sommets. Cette génération d'arbres de défaillances nous permet ensuite d'utiliser le moteur de calcul XFTA ([Rauzy, 2012]) qui nous fournit des résultats qualitatifs et quantitatifs associés : l'ensemble des coupes minimales avec leurs taux d'occurrences, exprimés par 5 heures de fonctionnement du SdS.

La Table 4 suivante fournit ces coupes et les taux associés pour la perte de la fonction de détection de bouchons. 'AV' est utilisé pour le composant 'AVehicle' du modèle, et 'Odo' pour le composant 'Odometer'. 'F' est utilisé pour l'événement 'failure' de l'élément 'class' 'NRComponent' dans le modèle, 'L' pour l'événement 'low' de l'élément 'class' 'Sensor', et 'E' pour l'événement 'erroneous' du composant 'Camera'. Nous remarquons qu'il y a une coupe d'ordre 1 'AV.Map.F', ce qui signifie que la défaillance de ce composant 'Map' de 'AVehicle' mène directement à l'événement redouté de perte de la fonction de détection de bouchons. De plus le taux d'occurrence de cette coupe est de $4.99e-4$, ce qui est bien supérieur au taux d'occurrence de l'événement redouté donné à $8.46e-7$. Ce cas est donc à traiter obligatoirement, par exemple en redondant un tel composant. Concernant les autres coupes, celles d'ordre 2 sont à étudier plus finement, notamment celles impliquant le satellite, afin d'avoir les vrais taux des occurrences des défaillances. Celles d'ordre 3 ont des taux d'occurrence bien inférieurs au taux d'occurrence de l'événement redouté, elles peuvent donc être acceptées.

Table 4 : Coupes minimales et taux d'occurrences pour la perte de la fonction de détection de bouchons

0.000499875	AV.Map.F		
6.24844e-08	AV.Camera.L	AV.Radar.L	
6.24844e-08	AV.Camera.L	AV.Lidar.L	
6.24844e-08	AV.Lidar.L	AV.Radar.L	
2.77176e-08	AV.GNSS.F	AV.Odo.F	
2.77176e-08	AV.Odo.F	GNSS.F	
3.12344e-11	AV.Camera.F	AV.Lidar.L	Camera.E
3.12344e-11	AV.Camera.F	AV.Radar.L	Camera.E
3.12344e-11	AV.Lidar.F	AV.Radar.L	Camera.E
3.12344e-11	AV.Camera.L	AV.Lidar.F	Camera.E
3.12344e-11	AV.Camera.L	AV.Radar.F	Camera.E
3.12344e-11	AV.Lidar.L	AV.Radar.F	Camera.E
1.56191e-11	AV.Camera.F	AV.Lidar.F	AV.Radar.F
1.56191e-11	AV.Camera.F	AV.Lidar.F	AV.Radar.L
1.56191e-11	AV.Camera.F	AV.Lidar.L	AV.Radar.F
1.56191e-11	AV.Camera.L	AV.Lidar.F	AV.Radar.F

La Table 5 suivante fournit les coupes et les taux associés pour la perte de la fonction de détection de feux connectés. Il y a ici deux coupes d'ordre 1, dont la première est la même que la coupe d'ordre 1 de l'événement redouté précédent. L'autre coupe d'ordre 1 est liée à la valeur 'Low' de la caméra avec un taux d'occurrence de $2.49e-4$. Dans le cadre de cette publication, cette seconde coupe d'ordre 1 ne sera pas considérée. Les autres coupes sont d'ordre 2 et sont, ici aussi, à étudier plus finement afin d'avoir les vrais taux des occurrences des défaillances.

Table 5 : Coupes minimales et taux d'occurrences pour la perte de la fonction de feux connectés

0.000499875	AV.Map.F	
0.000249969	AV.Camera.L	
1.24953e-07	AV.Camera.F	AV.V2X.F
6.24844e-08	AV.Camera.F	TrafficLight.F
6.24844e-08	AV.Camera.F	TrafficLight.E
2.77176e-08	AV.GNSS.F	AV.Odo.F
2.77176e-08	AV.Odo.F	GNSS.F

4.4 Recommandations pour la conception

L'étude menée sur le modèle AltaRica 3.0 du SdS de mobilité autonome et multimodale a fait apparaître au moins un élément à prendre en compte : la défaillance du composant 'Map' de 'AVehicle' mène directement aux événements redoutés de pertes des fonctions de détection de bouchons et de détection de feux connectés. Nous avons donc proposé une modification au niveau de la conception du système : la duplication de ce composant 'Map'. À la suite de cette modification, le modèle a été mis à jour puis de nouveau évalué. Cette mise à jour a été simple à réaliser en seulement 3 actions. D'abord l'ajout d'un clone de ce composant 'Map', ainsi qu'une entrée spécifique dans le composant 'Decision', comme montré en Figure 14. Enfin au sein de ce composant 'Decision', la prise en compte de ce second composant 'RedMap' dans le calcul de la sortie.

```
block AVehicle
  block Map /* behavior */ end
  clones Map as RedMap;
  /...
  assertion
    /...
    Decision.inMap := Map.out;
    Decision.inRedMap:= RedMap.out;
    /...
end
```

Figure 14 : Modélisation de la redondance du composant Map

Une nouvelle étude a été réalisée, que nous ne montrons pas dans cette publication. Elle a montré que l'ajout de ce nouveau composant a supprimé la coupe d'ordre 1 et diminué le taux d'occurrence associé.

5 CONCLUSION ET PERSPECTIVES

Dans cette publication, nous avons montré l'utilisation du langage AltaRica 3.0 pour étudier la sûreté de fonctionnement d'un système de systèmes de mobilité autonome et multimodale. En nous basant sur les descriptions d'architecture système, nous avons mené une analyse préliminaire des risques. Cette analyse nous a permis d'extraire l'ensemble des événements redoutés. Ces événements redoutés ont été cotés, puis nous avons pris en compte les plus critiques dans la suite des études. Nous avons ensuite présenté la modélisation et l'évaluation de deux événements redoutés spécifiques : les pertes des fonctions de détection de bouchons et de faux connectés. La modélisation a été réalisée via le langage AltaRica 3.0, puis l'étude associée avec les moteurs de calculs de la plateforme OpenAltaRica. L'étude a consisté à évaluer les deux arbres de défaillances, issus du modèle AltaRica 3.0, et de cette évaluation est sortie la non-teneur du taux d'occurrence de l'évènement redouté. Nous avons montré comment la recommandation d'apport d'un composant redondant a permis de résoudre ce point.

6 REMERCIEMENTS

Ces travaux ont été réalisés dans le cadre du projet de recherche EVA (<https://www.irt-systemx.fr/projets/eva/>) de l'IRT SystemX.

7 REFERENCES

Addoui, I., Batteux, M., Dubois, A., & Kerneis, R. (2020). Formal System and Safety Design of a System of Systems: a proof of concepts. In *Proceedings of European Safety and Reliability Conference, ESREL 2020. Venice, Italy*.

- Aupetit, B., Batteux, M., Rauzy, A. & Roussel, J. M. (2015). Improving performances of the AltaRica 3.0 stochastic simulator. In *Proceedings of European Safety and Reliability Conference (ESREL), Zurich, Switzerland. pp 1815-1823. September*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2017). AltaRica 3.0 assertions: the why and the wherefore. *Journal of Risk and Reliability. Professional Engineering Publishing. September*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2018-1). From Models of Structures to Structures of Models. In *IEEE International Symposium on Systems Engineering (ISSE 2018). Roma, Italy. October*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2018-2). AltaRica Wizard: an integrated modeling and simulation environment for AltaRica 3.0. In *Actes du congrès Lambda-Mu 21 (actes électroniques). Reims, France*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2019-1). AltaRica 3.0 in 10 Modeling Patterns. In *International Journal of Critical Computer-Based Systems. Inderscience Publishers. Vol. 9, Num. 1–2, pp 133–165*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2019-2). Modeling patterns for the assessment of maintenance policies with AltaRica 3.0. In *Yiannis Papadopoulos and Koorosh Aslansefat and Panagiotis Katsaros and Marco Bozzano ed., Model-Based Safety and Assessment. Thessaloniki, Greece. Vol. 11842, pp 32–46*.
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2020). Modélisation de combinaisons de maintenances en AltaRica 3.0. In *Actes du Congrès Lambda-Mu 22 (actes électroniques)*.
- Cassandras, C. G., & Lafortune, S. (2008). Introduction to Discrete Event Systems. *New-York, NY, USA: Springer*.
- ISO 26262 (2011). Road Vehicle – Functional Safety, Working Group TC22 SC32. *International Organization for Standardization, Standard*.
- ISO/WD PAS 21448 (2019). Road vehicles – Safety of the intended functionality. *International Organization for Standardization*.
- Maier M. W. (2005). Research Challenges for Systems-of-Systems. *IEEE International Conference on Systems, Man and Cybernetics, 2005, vol. 4, pp. 3149-3154*.
- Rauzy A. (2012). Anatomy of an Efficient Fault Tree Assessment Engine. In *Reino Virolainen ed., Proceedings of International Joint Conference PSAM'11/ESREL'12. Helsinki, Finland. pp 3333–3343, June*.
- Zimmermann, A. (2008). Stochastic Discrete Event Systems. *Springer-Verlag Berlin Heidelberg*.
- Zio, E. (2013). The Monte Carlo Simulation Method for System Reliability and Risk Analysis. *Springer. ISBN 978-1-4471-4588-2*

Session 2.5 : Sûreté de Fonctionnement & Maintenance

Chairman : Daoud AIT-KADI - Jean-Marie Flaus

Optimisation de stratégies de maintenances par simulation stochastique AltaRica 3.0

MICHEL BATTEUX¹ ET SELMA KHEBBACHE¹

¹ IRT SystemX

8, Avenue de la Vauve – CS 90070, 91127 Palaiseau CEDEX, France
prenom.nom@irt-systemx.fr

Résumé – La maintenance des systèmes industriels de production est un des enjeux actuels majeurs. Dans ce cadre la gestion de leurs maintenances est un facteur important de compétitivité. Une bonne stratégie de maintenance permet en effet une disponibilité élevée du système, tout en minimisant les coûts liés aux interventions. Il est pour cela possible d'utiliser différents types de maintenances: les maintenances correctives qui réparent le système après les occurrences de défaillances, et les maintenances préventives qui interviennent sur le système avant les occurrences des défaillances. Cet article propose une méthode de couplage entre de la simulation stochastique de modèles à événements discrets AltaRica 3.0, et d'algorithmes d'optimisation, dans un objectif d'optimisation de la stratégie des maintenances d'un système. Le modèle AltaRica 3.0 représente l'ensemble des composants du système, avec les défaillances et les maintenances associées. L'idée est de diriger la simulation du modèle suivant des paramètres obtenus par un algorithme d'optimisation, afin d'obtenir une, ou un ensemble de stratégies de maintenances optimales suivant des critères technico-économiques donnés (la disponibilité, les coûts de production ou d'exploitation, etc.).

Abstract – The maintenance of industrial production systems is one of the major actual challenges. In this case, the management of their maintenances is an important competitiveness factor. In fact a suitable maintenance strategy normally increases the system availability, and decreases costs due to interventions. Different kinds of maintenance policies can be used: corrective maintenances to repair the system after the occurrences of failures, preventive maintenances to maintain the system just before occurrences of failures. This work proposes a method for the coupling of the stochastic simulation of discrete event models AltaRica 3.0, and optimization algorithms, with a purpose of optimizing the maintenance strategy of the system. The AltaRica 3.0 model represents the components of the system, with their failures and associated maintenances. The idea is to drive the simulation according to parameters coming from an optimization algorithm, in order to obtain one, or a set of optimal maintenance strategies according to technico-economic criteria (the availability, the production or exploitation costs, etc.).

Mots clés - Maintenance, Optimisation, Simulation Stochastique, AltaRica.

Keywords – Maintenance, Optimization, Stochastic Simulation, AltaRica.

1 INTRODUCTION

La complexité des systèmes industriels de production est en constante augmentation. La raison principale tient dans la transformation numérique de ces systèmes, avec l'introduction de mécanismes de contrôle omniprésents et de composants combinés intensivement avec des parties logicielles. Un des enjeux actuels majeurs est d'améliorer les méthodes et outils utilisés pour le maintien en condition opérationnelle de tels systèmes de production. Ceci afin que ces méthodes et outils soient capables de prendre en compte ces nouvelles sources de complexité.

Dans ce cadre, la gestion de la maintenance de tels systèmes industriels est un facteur important de compétitivité. En effet, une bonne stratégie de maintenance permet une disponibilité élevée du système, tout en minimisant les coûts liés aux interventions. Différents types de maintenances existent. Les maintenances correctives qui consistent à intervenir uniquement après l'occurrence d'une défaillance du système. Bien que ces maintenances soient généralement peu coûteuses, elles peuvent cependant engendrer des coûts importants de non-fiabilité et d'indisponibilité du système. Les maintenances préventives qui consistent à intervenir par anticipation avant l'occurrence d'une défaillance du système. Certains types de telles maintenances préventives, comme les maintenances conditionnelles ou prévisionnelles, sont généralement plus coûteuses dans leur

mise en place car elles nécessitent d'instrumenter un suivi du système.

Dans cet article, nous proposons une méthode de couplage entre la simulation stochastique de modèles à événements discrets AltaRica 3.0, et un algorithme d'optimisation, dans un objectif d'optimisation de la stratégie des maintenances du système. Un tel modèle AltaRica 3.0 représente le système avec ses différents composants en interaction, les défaillances de ces composants, ainsi que les maintenances associées. L'idée est de diriger la simulation du modèle suivant des paramètres obtenus par un algorithme d'optimisation, afin d'obtenir une, ou un ensemble de stratégies de maintenance optimales suivant des critères technico-économiques donnés (par exemple la disponibilité globale du système, les coûts de production ou d'exploitation, etc.). Nous montrons ce couplage au travers d'un petit cas d'étude illustratif.

La suite de ce travail est organisée comme suit. La deuxième partie introduit le cas d'étude applicatif, qui nous servira de fil conducteur durant toute la publication. La troisième partie montrera conceptuellement le couplage entre l'optimisation et la simulation. La quatrième partie introduira, brièvement, le formalisme AltaRica 3.0, utilisé pour la modélisation des systèmes à événements discrets. Cette partie présentera aussi la simulation stochastique de tels modèles. La cinquième partie présentera les critères technico-économiques abordés pour l'optimisation de stratégies de maintenance, ainsi que les

algorithmes associés d'optimisation. La sixième partie montrera des résultats d'expérimentation de ce couplage sur le cas d'étude applicatif. Enfin la septième et dernière partie conclura cette publication, en ouvrant également des perspectives futures.

2 CAS D'ETUDE APPLICATIF

Pour exemplifier la méthode de couplage proposée, nous nous basons sur un cas d'étude applicatif défini par un petit système, nommé '5 composants', comme montré en Figure 1.

Ce système est constitué de cinq composants reliés entre eux suivant une structure combinant série/parallèle. Il transforme un flux (quelconque) via les différents composants, venant d'une source 'S' vers une cible 'T'. Ces composants sont divisés en deux lignes complémentaires 'L1' et 'L2'. Cela signifie que la cible 'T' est alimentée par le flux si et seulement si les deux lignes 'L1' et 'L2' sont opérationnelles. De plus, les deux composants 'C4' et 'C5' sont en redondance chaude, au sens où il suffit que l'un des deux soit fonctionnel pour qu'il y ait un flux en sortie. Enfin nous considérons des flux Booléens ; c'est-à-dire que nous ne modéliserons que la présence ou l'absence du flux.

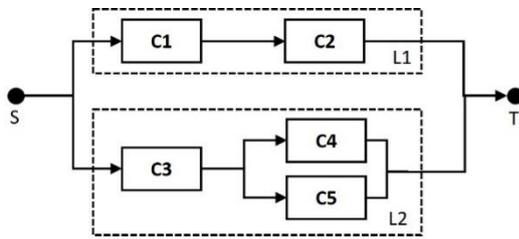


Figure 1 : Système 5 composants

Le comportement des composants est identique et décrit par la machine à état de la Figure 2. Chacun des composants peut donc être dans 3 états distincts : en marche, en panne ou en maintenance. Un composant est initialement en marche et il peut tomber en panne suivant une loi exponentielle. Il passe en état maintenance lorsqu'une maintenance du composant est mise en place, et indépendamment qu'il soit en marche ou en panne. À la suite d'une maintenance il revient en état en marche. Les paramètres des lois exponentielles de panne des composants sont respectivement ' $1.0e-3$ ' pour 'C1', 'C2' et 'C3', et ' $2.0e-3$ ' pour 'C4' et 'C5'. Les maintenances des composants ne sont pas contrôlées par les composants eux-mêmes, mais au niveau du système, par un composant de supervision, que nous n'avons pas représenté en Figure 1. Nous considérons de plus uniquement des maintenances planifiées à intervalles fixes. L'enjeu de l'optimisation sera de définir, pour chaque composant, ces intervalles des maintenances planifiées les plus optimaux par rapport à un critère de disponibilité du système.

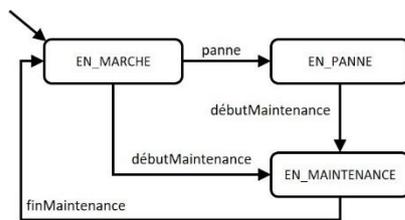


Figure 2 : Machine à états du comportement des composants

Il ne s'agit pour cette étude que d'un exemple simple pour montrer la preuve de concept. Il est bien sûr possible de modéliser des comportements plus complexes combinant

plusieurs modes de fonctionnements nominaux, des modes dégradés, des reconfigurations, des maintenances conditionnelles, etc. ; permis par le pouvoir d'expression du langage AltaRica 3.0 (voir la partie 4.1).

3 COUPLAGE ENTRE OPTIMISATION ET SIMULATION ALTARICA 3.0

Dans un objectif d'optimisation de stratégies de maintenances, la méthode proposée de couplage se réalise entre d'une part la simulation stochastique de modèles à événements discrets AltaRica 3.0, et d'autre part des algorithmes d'optimisation. La Figure 3 représente le principe de ce couplage. Concrètement il s'agit de contrôler la simulation par l'optimisation via des itérations successives. À une itération N, la partie optimisation fournit donc un ensemble de valeurs de paramètres à la partie simulation : les valeurs des intervalles des maintenances planifiées pour chaque composant. Ensuite la partie simulation utilise ces valeurs des paramètres lors des simulations stochastiques pour produire des résultats, plus précisément des valeurs statistiques. Ces résultats statistiques sont alors réutilisés par la partie optimisation à l'itération N+1 afin de produire de nouvelles valeurs de paramètres affinées en fonction d'une part des critères d'optimisation, et d'autre part des résultats statistiques.

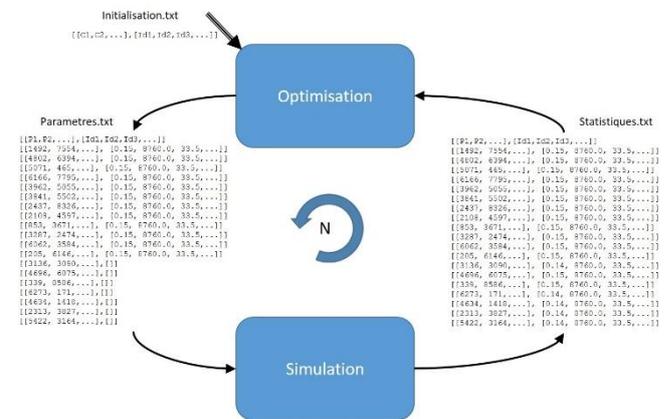


Figure 3 : Couplage optimisation et simulation

Concrètement dans le cadre de notre cas d'étude applicatif, nous souhaitons optimiser, pour chaque composant, les valeurs des intervalles de temps entre deux planifications de maintenances. Cette optimisation se réalisant en fonction du critère de disponibilité globale au système.

4 MODELISATION ET SIMULATION STOCHASTIQUE EN ALTARICA 3.0

Dans cette partie, nous décrivons le langage de modélisation AltaRica 3.0, utilisé pour la modélisation des systèmes à événements discrets. Cette partie présentera aussi la simulation stochastique de tels modèles.

4.1 Le langage AltaRica 3.0

AltaRica 3.0 est un langage de modélisation événementiel et orienté objet, dédié, initialement, aux analyses probabilistes du risque de systèmes complexes ([Batteux et al., 2019-1]). Ce langage permet de modéliser les systèmes à plus haut niveau que les formalismes classiques d'analyse du risque (arbres de défaillance, chaînes de Markov, réseaux de Petri stochastiques, etc.), sans augmenter la complexité des calculs d'indicateurs du risque. AltaRica 3.0 est la combinaison du cadre mathématique GTS, pour Guarded Transitions System ([Batteux et al., 2017]),

et du paradigme de structuration S2ML, pour System Structure Modeling Language ([Batteux et al., 2018]).

Le cadre mathématique GTS permet de rendre compte des aspects comportementaux que l'on souhaite modéliser, c'est-à-dire la sémantique d'exécution. Cette exécution reste similaire aux autres formalismes à événements discrets dans le sens où chaque fois qu'une transition est tirable, elle est planifiée et sera potentiellement tirée après un certain délai ([Cassandras et al., 2008]) et [Zimmermann, 2008]). Ces délais peuvent être déterministes ou stochastiques ; et AltaRica 3.0 fournit les délais usuels (du type exponentiel, Weibull, etc.) et des délais empiriques.

Un exemple de modèle AltaRica 3.0, décrivant un composant tel que présenté en Figure 2, est donné par la Figure 4. Un domaine de valeurs 'CState' est déclaré. Il servira à typer la variable d'état du composant. La déclaration du composant est ensuite donnée, via le mot-clé 'block'. Ce composant définit d'abord le comportement interne du composant. Pour cela il déclare différents éléments internes, puis définit les transitions associées à ces éléments. Sont donc d'abord déclarés une variable d'état 'vs' de type 'CState', trois paramètres réels, ainsi que trois événements. Les paramètres permettent de définir les valeurs des paramètres des lois des délais associés aux événements : par exemple l'événement 'failure' a pour délai une loi exponentielle de paramètre 'lambda' qui vaut la valeur réelle '1.0e-3'. La partie transition permet de définir le passage d'un état à un autre suivant l'occurrence d'un événement. Par exemple la transition, labélisée par l'événement 'startMaintenance', d'une part est conditionnée par le fait que la variable d'état 'vs' ne doit pas avoir la valeur 'MAINTENANCE', et d'autre part met à jour la valeur de cette variable d'état 'vs'. Enfin la dernière partie exprime le comportement externe du composant. Elle déclare deux variables booléennes, 'in' et 'out', dites de flux. Elle définit ensuite l'assertion qui exprime la relation de propagation entre ces deux variables de flux et en fonction de la variable d'état 'vs'.

```

domain CStates {WORKING, FAILED, MAINTENANCE}

block Component
  CStates vs (init = WORKING);
  parameter Real lambda = 1.0e-3;
  parameter Real mu = 100.0;
  parameter Real nu = 10.0;
  event failure (delay = exponential(lambda));
  event startMaintenance (delay = Dirac(mu));
  event finishMaintenance (delay = Dirac(nu));
  transition
    failure: vs == WORKING -> vs := FAILED;
    startMaintenance: not vs == MAINTENANCE
      -> vs := MAINTENANCE;
    finishMaintenance: vs == MAINTENANCE
      -> vs := WORKING;
  Boolean in, out (reset = false);
  assertion
    out := vs == WORKING and in;
end

```

Figure 4 : Modèle AltaRica 3.0 d'un composant

Comme déjà indiqué, le pouvoir d'expression du langage AltaRica 3.0 permet d'appréhender des comportements plus complexes non seulement au niveau des composants, mais aussi au niveau du système. Par exemple combinant plusieurs modes de fonctionnements nominaux, des modes dégradés, des reconfigurations, des maintenances conditionnelles, etc.

4.2 Simulation Stochastique de modèles AltaRica 3.0

La simulation stochastique consiste à simuler un grand nombre de fois, et de façon pseudo-aléatoire, l'évolution du système modélisé, aussi nommée histoire ; et d'observer pour chacune de ces histoires, la valeurs de différents indicateurs, puis de faire des statistiques sur ces valeurs observées ([Zio, 2013]).

Dans le cadre AltaRica 3.0, il s'agit d'exécuter le modèle via les tirages des transitions, dépendant elles-mêmes des délais des événements associés. La caractéristique pseudo-aléatoire est, dans ce cadre, donnée d'une part par les choix des dates des tirages, lorsque les délais sont stochastiques ; d'autre part par les choix à faire lorsque deux transitions tirables sont en concurrence, c'est-à-dire des transitions dont les délais sont déterministes, et qui peuvent être tirées en même temps ([Aupetit et al., 2015]).

En positionnant des observateurs sur le modèle (voir dans la partie suivante), il est possible de définir différents indicateurs associés permettant de réaliser des statistiques. Les indicateurs sont considérés comme des variables aléatoires réelles, qui fournissent, à des dates spécifiques des histoires, les valeurs prises par les observateurs. Ces dates des histoires sont typiquement à la fin, ou des instants spécifiés dans le temps. Enfin les statistiques réalisées sur ces indicateurs sont principalement des moyennes, des écarts-types ou des intervalles de confiance.

4.3 Modélisation et simulation stochastique AltaRica 3.0 du cas d'étude applicatif

De premiers travaux de modélisation de maintenances en AltaRica 3.0 ont déjà été réalisés ([Batteux et al., 2020] ou [Batteux et al., 2019-2]). Néanmoins ces travaux n'ont pas étudié les possibilités d'optimisation de différentes stratégies de maintenances d'un système.

Pour modéliser notre cas d'étude applicatif en AltaRica 3.0, nous reprenons le composant défini en Figure 4, sauf que nous ne le considérons pas comme un objet 'block' mais comme un objet 'class', comme montré en Figure 5. Nous pourrions ainsi instancier plusieurs fois cette classe 'Component' dans le modèle du système.

```

class Component
  // elements of a component
end

```

Figure 5 : Classe AltaRica 3.0 d'un composant

La Figure 6 représente le modèle AltaRica 3.0 du système 5 composants. Ce système est défini par un objet 'block' AltaRica 3.0. Les cinq composants 'C1' à 'C5' sont d'abord déclarés, ce sont des instances de la classe 'Component' donnée en Figure 5. Les paramètres 'lambda' et 'mu' de ces cinq instances sont surchargés respectivement par les valeurs données en partie 2 pour les 'lambda', et par des paramètres 'muC1' à 'muC5' qui seront définis ailleurs. Ce seront ces valeurs des paramètres qui seront fournies par la partie optimisation et qui seront modifiées à chaque itération de la boucle de couplage entre la simulation et l'optimisation. Deux variables booléennes sont définies afin de représenter la source et la cible de ce système 5 composants. Ensuite l'assertion spécifie les connexions entre les variables de flux des cinq composants, suivant ce qui est représenté en Figure 1. Enfin un observateur Booléen 'available' est défini et positionné sur la variable Booléenne 'Target' lorsqu'elle prend la valeur 'true'.

Pour la simulation stochastique de ce modèle, nous considérons un temps de mission de 1 an, c'est-à-dire 8760 heures. Nous définissons un indicateur de type temps-de-séjour positionné sur

l'observateur 'available' à la valeur 'true'. Cela s'exprime par la variable aléatoire 'sojour-time(true,8760.0)' et représente la durée totale des intervalles de temps durant lesquels l'observateur 'available' prend la valeur 'true' sur la période de temps '[0.0,8760.0]'. Cela nous fournit le temps de disponibilité, ou de fonctionnement, moyen du système durant le temps de mission. Il s'agit de cet indicateur qui est utilisé pour fournir la statistique de temps de disponibilité, à chaque itération, à la partie optimisation.

```

block System
Component C1 (lambda = 1.0e-3 , mu = muC1);
Component C2 (lambda = 1.0e-3 , mu = muC2);
Component C3 (lambda = 1.0e-3 , mu = muC3);
Component C4 (lambda = 2.0e-3 , mu = muC4);
Component C5 (lambda = 2.0e-3 , mu = muC5);
Boolean Source, Target (reset = false);
assertion
Source := true;
C1.in := Source;
C3.in := Source;
C2.in := C1.out;
C4.in := C3.out;
C5.in := C3.out;
Target := C2.out and (C4.out or C5.out);
observer Boolean available = Target == true;
end

```

Figure 6 : Modèle AltaRica 3.0 du système 5 composants

Pour tester ce modèle en simulation stochastique nous avons défini différentes valeurs des paramètres 'muC1' à 'muC5'. Le Tableau 1 montre les résultats de différentes simulations stochastiques de ce modèle AltaRica 3.0 du système 5 composants suivant ces différents jeux de paramètres. Nous avons procédé à 100000 histoires par simulation stochastique

Tableau 1 : Simulations stochastiques du système 5 composants

muC1 = muC2 = muC3 = 2000.0	muC4 = muC5 = 1000.0	
Fired transitions	M	84.2358
	Min	70
	Max	90
	M	SD
sojour-time(true,8760.0)	1255.34	539.282
number-of-occurrences(false,8760.0)	5.63601	1.03424
mean-time-between-occurrences(false,8760.0)	1753.14	392.3
muC1 = muC2 = muC3 = 1000.0	muC4 = muC5 = 500.0	
Fired transitions	M	155.487
	Min	139
	Max	171
	M	SD
sojour-time(true,8760.0)	2350.76	716.645
number-of-occurrences(false,8760.0)	14.0766	2.00837
mean-time-between-occurrences(false,8760.0)	630.761	104.228
muC1 = muC2 = muC3 = 500.0	muC4 = muC5 = 250.0	
Fired transitions	M	281.907
	Min	258
	Max	305
	M	SD
sojour-time(true,8760.0)	4037.85	682.327
number-of-occurrences(false,8760.0)	33.7144	3.18675
mean-time-between-occurrences(false,8760.0)	257.007	25.5432

Ce tableau est décomposé en trois parties correspondantes aux différentes valeurs des paramètres 'muC1' à 'muC5'. Dans chacune des parties, d'une part sont données les statistiques sur le nombre de transitions tirées par simulation ('M' représentant la moyenne), et d'autre part les résultats des trois indicateurs sont donnés : la moyenne 'M' de l'indicateur et son écart type 'SD' associé. Outre l'indicateur 'sojour-time(true,8760.0)', nous avons considéré deux autres indicateurs : 'number-of-occurrences(false,8760.0)' qui représente le nombre de dates auxquelles l'observateur prend la valeur 'false' sur la période de temps '[0.0,8760.0]', et 'mean-time-between-occurrences(false,8760.0)' qui représente la durée moyenne entre deux dates successives auxquelles l'observateur prend la valeur 'false' sur la période de temps '[0.0,8760.0]'. Nous observons que plus les délais des maintenances sont rapprochés, plus le temps de disponibilité (moyen) du système est grand. Bien évidemment cela implique aussi que le nombre d'occurrences d'indisponibilité du système est plus grand, et que le temps moyen entre deux indisponibilités du système est aussi plus grand. Enfin cela implique aussi qu'il y a plus de transitions tirées, ce qui s'explique par le fait que la réduction du temps entre maintenance provoque des tirages de transitions plus rapprochés.

5 OPTIMISATION DE STRATEGIES DE MAINTENANCE

Nous proposons dans cette section de présenter une approche de décision visant à ordonnancer les opérations de maintenance, afin d'avoir la meilleure disponibilité possible du système. Il s'agit donc de définir les durées entre les maintenances pour chaque composant, afin de permettre au système d'assurer sa mission dans des conditions nominales et de ce fait, de maximiser la disponibilité des équipements. L'optimisation de la stratégie de maintenance se réalise donc en fonction du critère de disponibilité du système.

Dans cet article, nous développons une approche basée sur l'algorithme génétique couplé avec le simulateur afin de maximiser le temps de disponibilité du système à 5 composants décrit en partie 2.

L'Algorithme Génétique (AG) développé par [Holland, 1975] fait partie de la famille des algorithmes évolutionnistes de type métaheuristiques. C'est une méthode approchée, très flexible et efficace qui peut résoudre des problèmes très complexes. Nous présentons dans ce qui suit les étapes de cette méthode basée AG.

5.1 Chromosomes et évaluation

Chaque Chromosome est une séquence de gènes, de taille 5 pour notre application. Ces gènes représentent la durée entre les maintenances pour chaque composant du système. Ils sont générés dans un intervalle [LB, UB] suivant une distribution uniforme (LB est la limite inférieure, UB est la limite supérieure) comme l'illustre la Figure 7. LB indique la durée minimale entre deux maintenances pour chaque composant et UB est la durée de mission de la simulation.

Un chromosome n'est autre qu'un vecteur dont chaque élément est une période d'intervention possible pour chaque composant. Les temps générés sont liés aux contraintes du problème.

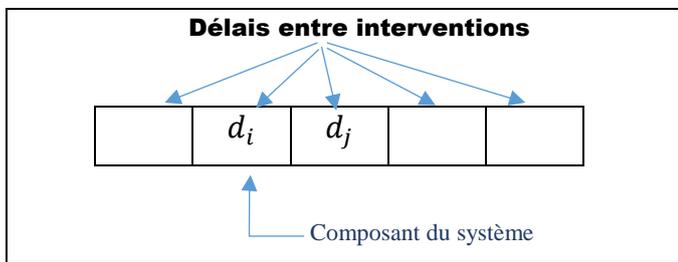


Figure 7 : Chromosome

5.2 Croisement

Dans ce travail, nous avons utilisé trois techniques de croisement :

1. **Croisement de type 1X** : On choisit aléatoirement deux chromosomes dans la population, un point est ensuite positionné de manière aléatoire dans les chromosomes parents et on aboutit par une succession de permutations à deux chromosomes dit enfants.
2. **Croisement de type OX** : proposé par [Davis, 1985] est un croisement entre deux points. Les deux parents produiront deux enfants. Premièrement, les segments entre points de coupures sont copiés. Ensuite, il suffit de partir du point de coupure d'un des parents, les lots de l'autre parent sont copiés dans le même ordre, omettant les gènes déjà présents. Lorsque l'on atteint la fin du vecteur, on continue depuis le premier emplacement de la chaîne. Cette séquence est placée dans le premier enfant. Le point de départ est la seconde coupure. De la même manière, on obtient le deuxième enfant.
3. **Croisement de type LOX** : est également un croisement en deux points. Proposé par [Falkenauer et Bouffoux 1991], ce croisement est une variante du croisement OX. Cette modification s'effectue lors de la construction des enfants. Au lieu de considérer comme point de départ la seconde coupure, on considère le début du chromosome parent en le balayant en entier.

5.3 Mutation et Population

La mutation consiste à prendre des gènes aléatoires dans le chromosome enfant et à les altérer. Dans notre approche, nous permutons entre deux gènes de manière aléatoire. La mutation est appliquée avec une faible probabilité.

La population contient un nombre fixe de chromosomes triés par ordre croissant de la fonction objectif. Un chromosome ou une solution est formé de 5 gènes, chaque gène est un composant du système contenant le délai d_i qui est la période entre les interventions du composant i . Chaque chromosome doit respecter la contrainte que deux maintenances pour deux composants i et j ne peuvent pas être effectuées en même temps ($d_i \neq d_j$) et pour considérer le temps de maintenance on suppose qu'au moins une heure doit-être respectée entre les périodes des différents composants aussi les périodes ne dépassent pas l'horizon de simulation fixé.

5.4 Approche basée Algorithme Génétique

L'approche développée dans cet article repose sur un couplage entre un module de simulation via l'outil présenté dans la partie 4 et une méthode d'optimisation basée sur l'algorithme génétique, dans le but d'ordonner de manière optimale (ou sous-optimale) les opérations de maintenance afin de maximiser le temps de disponibilité du système étudié. Peu d'études dans la littérature utilisent les approches du couplage entre simulation

et optimisation des opérations de maintenance. Un article de [Alradghi *et al.*, 2016] synthétise des outils et techniques de modélisation et simulation des systèmes complexes. Les mêmes auteurs [Alradghi *et al.*, 2017] ont proposé un couplage entre un outil de simulation et l'algorithme du recuit simulé afin d'optimiser l'ordonnement des opérations de maintenance préventive en prenant en compte la disponibilité des pièces de rechange afin de minimiser les coûts totaux de la maintenance. [Khebbache-Hadji *et al.*, 2012] ont proposé un modèle de simulation sous Arena couplé avec un algorithme génétique pour l'ordonnement des tâches de maintenance d'une entreprise de fabrication des pièces de transmission.

Nous proposons dans ce travail, d'optimiser la planification des tâches de maintenance en couplant l'outil de simulation stochastique AltaRica 3.0 avec un algorithme génétique. Il existe de nombreux travaux portant sur de tels couplages, et notre objectif dans cette publication n'est pas ni de les répertorier, ni d'apporter des éléments de comparaison, mais uniquement de montrer la preuve de concept : la faisabilité de couplage entre cet outil spécifique de simulation stochastique AltaRica 3.0 avec un algorithme génétique.

La structure globale de cette approche est présentée dans la Figure 3 et détaillée dans le pseudo-code suivant :

Algorithme : Couplage simulation et optimisation

Initialisation : population initiale de taille N

Entrés : $Iterations_{number}$, P_c et P_m

Etape 1 : lancer le module de simulation qui renvoi les temps des disponibilités du système pour N solutions de la population initiale.

Etape 2 : Trier la population par ordre décroissant de la fonction fitness (le temps de disponibilité).

Tant que : $i \leq (Iterations_{number} - 1)$

Etape 3 : Appliquer les techniques de croisements avec une probabilité P_c

Etape 4 : Si r (valeur aléatoire) $< P_m$ alors appliquer la mutation.

Etape 5 : le module de simulation est appelé à nouveau pour calculer les temps des disponibilités du système pour les chromosomes enfants.

Etape 6 : Trier les solutions, les meilleurs chromosomes en termes de temps de disponibilité sont gardés, les mauvaises solutions sont supprimées de la population.

Etape 7 : Aller à l'Etape 3.

Fin Tant que

L'algorithme génère la population initiale, les solutions de cette dernière sont envoyées au module de simulation qui fournit le temps de disponibilité du système pour chaque chromosome.

Ensuite, la population est triée selon ce temps de disponibilité, par ordre décroissant comme première étape de l'algorithme génétique. Le tri élitiste est important pour considérer les meilleurs individus. Le croisement à 50% de la population est appliqué, dont 40% subissent le croisement LOX et si les chromosomes enfants ne vérifient pas les contraintes, le croisement OX est appliqué, et 60% subissent le croisement OX et les chromosomes qui ne sont pas conformes subissent le croisement de type LOX. Enfin si les enfants sont toujours non compatibles, l'itération est répétée. Après croisement, un chromosome peut subir une mutation avec une probabilité de $P_m = 0.005$. Les chromosomes enfants sont ajoutés à la population pour récupérer ensuite leurs temps des disponibilités

par simulation. Ainsi à chaque itération, la population est modifiée.

6 EXPERIMENTATIONS

Le couplage entre la simulation AltaRica 3.0 et l'optimisation, comme montré en Figure 3, se réalise par l'utilisation de différents outils. Le simulateur stochastique AltaRica 3.0 est utilisé tel quel. Nous avons implémenté un script Python permettant d'une part d'appeler ce simulateur stochastique suivant les paramètres fournis par l'optimisation via le fichier d'échange, et d'autres part de produire le fichier d'échange contenant les valeurs des statistiques des simulations. L'algorithme d'optimisation est implémenté par des scripts Python. Enfin le couplage entre l'optimisation et la simulation est lui aussi implémenté par un script Python. Ce dernier script permet d'appeler itérativement les scripts d'optimisation et de simulation. Hormis le code du simulateur stochastique AltaRica 3.0, qui est en C++, l'ensemble de ces scripts python représente environ 2400 lignes de code Python.

Nous avons exécuté le couplage en paramétrant le nombre d'itération à 50 et le nombre de chromosome à 100. Cela signifie qu'à la première itération 100 chromosomes sont produits par l'optimisation, c'est-à-dire 100 séquences de valeurs représentant les dates des maintenances des 5 composants et la simulation réalise 100 simulations stochastiques pour calculer les temps de disponibilité correspondants à ces 100 chromosomes. Ensuite à chaque nouvelle itération, l'optimisation réalise les différentes étapes de l'algorithme génétique : production de 50 nouveaux chromosomes, suivant les croisements, en fonction des résultats issus de la simulation sur les chromosomes restant de la précédente itération, puis simulation de ces 50 nouveaux chromosomes.

La Figure 8 représente la convergence du temps de disponibilité en fonction du nombre d'itérations. L'axe des abscisses représente les itérations et l'axe des ordonnées représente le temps de disponibilité du système. Nous observons sur le graphique que la convergence apparaît vers la 23ème itération. À noter qu'entre la 13ème et la 22ème itération, la valeur du temps de disponibilité est de 4948,93, et qu'à partir de la 23ème itération elle passe à 4949,17 ; ce qui ne se voit pas sur le graphique. Ceci montre bien l'efficacité de cet algorithme pour ce genre de problèmes avec une convergence très rapide. Cette représentation ne montre pas les différents gènes des chromosomes, c'est-à-dire les durées entre les maintenances pour chaque composant du système. En effet, il faudrait pour cela montrer, pour chaque itération (en abscisse), les 5 gènes du chromosome produisant la meilleure disponibilité à cette itération.

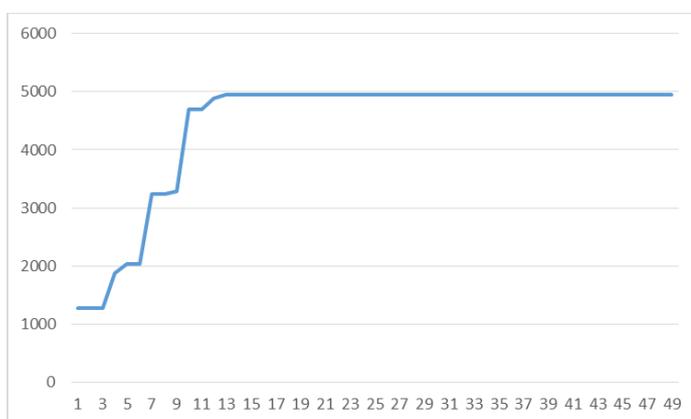


Figure 8 : Convergence de la disponibilité

La Figure 9 représente l'évolution du temps de disponibilité en variant les taux de défaillance des composants. L'axe des abscisses correspond au temps de disponibilité du système et l'axe des ordonnées représente les différents taux de pannes. On remarque que plus le taux de défaillance est élevé, et plus le temps de disponibilité est faible. Avec un taux de défaillance de 3×10^{-3} le temps de disponibilité dépasse 5500. Selon le graphique, en considérant un taux de 10^{-1} le temps de disponibilité est de 3.5455. Plus les pannes sont rapprochés, et plus le temps d'indisponibilité du système augmente.

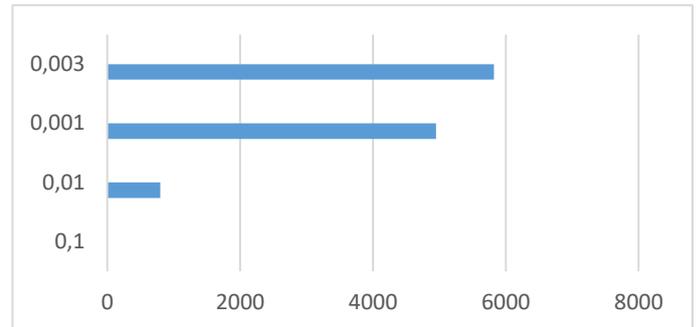


Figure 9 : Evolution de la disponibilité en fonction des taux de pannes

Nous n'avons pas considéré les temps d'exécution de ce couplage, car d'une part le modèle considéré est petit, et d'autre part des travaux sont en cours de réalisation sur le simulateur stochastique AltaRica 3.0. Ces travaux permettront d'éviter l'étape de compilation C++ pour la production du simulateur ; étape qui doit se réaliser pour chaque nouvel ensemble de paramètres issus d'un chromosome.

7 CONCLUSION

Dans cet article, nous avons présenté une proposition de méthode de couplage entre de la simulation stochastique de modèles à événements discrets AltaRica 3.0, et d'algorithmes d'optimisation. Ce couplage se réalisant dans un objectif d'optimisation de stratégie des maintenances de systèmes industriels de production. Ceci se justifiant par le fait qu'une bonne gestion de la maintenance de tels systèmes est un facteur important de compétitivité. Au travers d'un petit cas d'étude applicatif, un système 5 Composants, nous avons montré comment se réalise ce couplage, et notamment l'idée sous-jacente de diriger la simulation de modèles suivant des paramètres obtenus par un algorithme d'optimisation, afin d'obtenir une, ou un ensemble de stratégies de maintenance optimales suivant des critères technico-économiques donnés (la disponibilité, les coûts de production ou d'exploitation, etc.). Nous avons donc présenté d'une part le formalisme AltaRica 3.0 permettant de modéliser et simuler de tels systèmes industriels de production, intégrant les éléments du système, leurs interactions fonctionnelles, leurs défaillances, les maintenances associées, etc. Nous avons présenté d'autre part l'algorithme d'optimisation associé à ce que nous souhaitons évaluer pour notre cas d'étude applicatif 5 Composants : l'Algorithme Génétique. Enfin nous avons présenté l'expérimentation associée du couplage et montré la convergence du temps de disponibilité du système en fonction du nombre d'itérations.

S'agissant d'une première expérimentation permettant d'apporter une preuve de concept de diriger la simulation stochastique AltaRica 3.0 par des algorithmes d'optimisation, nos perspectives s'orientent donc suivant deux principaux axes. D'une part sur l'aspect modélisation et simulation, nous allons

considérer d'autres exemples de système plus importants qui combineront différents comportements en fonctionnement nominal, en fonctionnements dégradés, en défaillances, etc., ce que permet le pouvoir d'expression du langage. Nous intégrerons aussi différentes possibilités de maintenances (préventives ou correctives), combinés ou à choisir sur les différents composants. Nous intégrerons de plus l'évolution du simulateur stochastique AltaRica 3.0, et réaliseront des tests de performances afin d'évaluer notre approche. D'autre part sur l'aspect optimisation, nous allons étudier d'autres algorithmes d'optimisation suivant les critères technico-économiques à considérer, mais aussi en fonction des paramètres différents possibles du modèle. Par exemple faire un choix entre une maintenance corrective ou une maintenance préventive. D'autres algorithmes d'optimisation notamment des approches exactes seront explorés afin d'apporter une comparaison avec l'algorithme présenté dans cet article.

8 REFERENCES

- Alrabghi, A., Tiwari, A. (2016). A novel approach for modelling complex maintenance systems using discrete event simulation, *Reliability Engineering & System Safety*, Vol 154, pp 160-170.
- Alrabghi, A. & Tiwari, A. (2017). Simulation-based optimisation of maintenance systems: Industrial case studies. *Journal of Manufacturing Systems*, Journal of Manufacturing Systems, vol 44, pp191-206.
- Aupetit, B., Batteux, M., Rauzy, A. & Roussel, J. M. (2015). Improving performances of the AltaRica 3.0 stochastic simulator. *In Proceedings of European Safety and Reliability Conference (ESREL), Zurich, Switzerland. pp 1815-1823. September.*
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2017). Altarica 3.0 assertions: the why and the wherefore. *Journal of Risk and Reliability. Professional Engineering Publishing. September.*
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2018). From Models of Structures to Structures of Models. *In IEEE International Symposium on Systems Engineering (ISSE 2018). Roma, Italy. October.*
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2019-1). AltaRica 3.0 in 10 Modeling Patterns. *In International Journal of Critical Computer-Based Systems. Inderscience Publishers. Vol. 9, Num. 1-2, pp 133-165.*
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2019-2). Modeling patterns for the assessment of maintenance policies with AltaRica 3.0. *In Yiannis Papadopoulos and Koorosh Aslansefat and Panagiotis Katsaros and Marco Bozzano ed., Model-Based Safety and Assessment. Thessaloniki, Greece. Vol. 11842, pp 32-46.*
- Batteux, M., Prosvirnova, T., & Rauzy, A. (2020). Modélisation de combinaisons de maintenances en AltaRica 3.0. *In Actes du Congrès Lambda-Mu 22 (actes électroniques).*
- Cassandras, C. G., & Lafortune, S. (2008). Introduction to Discrete Event Systems. *New-York, NY, USA: Springer.*
- Holland, J. H. (1975). Adaptation in natural and artificial systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan press.
- Davis, L. (1985). Applying adaptive algorithms to epistatic domains. *In Proceedings of the 9th international joint*

- conference on Artificial intelligence - Volume 1 (IJCAI'85). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 162-164.
- Falkenauer, E & Bouffouix, S. (1991). A Genetic algorithm for Job Shop. *Proceedings of 1991 IEEE International Conference on Robotics and Automation, Sacramento, CA, USA, 1991, pp. 824-829 vol.1.*
- Khebbache-Hadji, S. Hani, Y. Lahiani N. & El Mhamedi, A. (2012). Genetic algorithm used in simulation model: Application in a maintenance process. *14th IFAC Symposium on Information Control Problems in Manufacturing. Buchrest, Roumania, May.*
- Zimmermann, A. (2008). Stochastic Discrete Event Systems. *Springer-Verlag Berlin Heidelberg.*
- Zio, E. (2013). The Monte Carlo Simulation Method for System Reliability and Risk Analysis. *Springer. ISBN 978-1-4471-4588-2*

Stratégie intégrée de maintenance préventive périodique, de production et de contrôle qualité dans les systèmes manufacturiers

NADIA BAHRIA¹, IMEN HARBAOUI DRIDI², ANIS CHELBI³, HANEN BOUCHRIHA²

¹ Université de Tunis El Manar, OASIS
École Nationale d'Ingénieurs de Tunis (ENIT)
BP 37, le Belvédère 1002 Tunis - Tunisie
École Polytechnique Internationale Privée de Tunis, PI-LAB
Rue du Lac d'Annecy, Tunis 1053 - Tunisie
nadia.bahria@enit.utm.tn

² Université de Tunis El Manar, LACS
École Nationale d'Ingénieurs de Tunis (ENIT)
BP 37, le Belvédère 1002 Tunis - Tunisie
imen.harbaoui@issatkr.rnu.tn, hanen.bouchriha@enit.rnu.tn

³ Université de Tunis, CEREP
École Nationale Supérieure d'Ingénieurs de Tunis (ENSIT)
5 Av. Taha Hussein, 1008 Tunis – Tunisie
anis.chelbi@planet.tn

Résumé – Ce papier présente une stratégie d'optimisation conjointe du stock, de la maintenance et de la qualité considérant une politique de maintenance préventive périodique selon un échéancier. Nous considérons une seule machine soumise à un processus de dégradation qui affecte directement la qualité des produits. Le but principal de ce travail est de développer un modèle qui renforce le lien sous-jacent entre les politiques de maintenance préventive et corrective, le stock tampon et les paramètres d'une carte de contrôle 'x-bar'. La stratégie de maintenance préventive consiste à effectuer des actions de maintenance parfaites périodiques (la période étant un multiple de l'intervalle d'échantillonnage). Une action de maintenance corrective est entreprise suite au dépassement des limites de contrôle de la carte. L'objectif est de déterminer les valeurs optimales de la période de maintenance préventive, de la taille du stock tampon, de la taille de l'échantillon, de l'intervalle d'échantillonnage et les limites de la carte de contrôle, qui minimisent le coût total moyen par unité de temps.

Abstract – This paper presents a joint optimization strategy of inventory, maintenance and quality considering a periodic preventive maintenance policy according to a schedule. We consider a single machine subjected to a degradation process which directly affects the quality of the products. The main goal of this work is to develop a model that strengthens the underlying link between preventive and corrective maintenance policies, buffer stock and parameters of an 'x-bar' control chart. The preventive maintenance strategy consists in carrying out periodic perfect maintenance actions (the period being a multiple of the sampling interval). Corrective maintenance action is taken when the control chart limits are exceeded. The goal is to determine the optimal values of preventive maintenance period, buffer stock size, sample size, sampling interval and control chart limits, which minimize the average total cost per time unit.

Mots clés – Stratégie intégrée, carte de contrôle, optimisation, stock tampon, maintenance préventive et corrective.

Keywords – Integrated strategy, control chart, optimization, preventive and corrective maintenance.

1 INTRODUCTION

La gestion des opérations est devenue une tâche complexe qui vise simultanément, à accroître l'efficacité des moyens de production, à réduire la durée et la fréquence d'occurrence des pannes, à fabriquer des produits de bonne qualité, à satisfaire les clients et à minimiser les coûts opérationnels. Dans le secteur manufacturier, la gestion des opérations repose essentiellement sur trois principales fonctions : production, maintenance et qualité. La conception et l'optimisation conjointe de ces trois fonctions permettent d'améliorer

l'efficacité des unités de production et de réduire considérablement les coûts. Les unités de production sont généralement sujettes à plusieurs processus de dégradation qui sont à l'origine de la perturbation des activités de production et affectent la qualité des produits. Toutefois, ces trois fonctions sont souvent gérées séparément dans la pratique bien qu'elles soient étroitement liées.

Cette problématique a fait l'objet de plusieurs travaux de recherche. Parmi ceux-ci, [Ben-Daya, 2002] a proposé un modèle intégré afin d'optimiser la quantité économique de

production (QEP) et la période de maintenance préventive (MP) pour un processus imparfait ayant une distribution de détérioration générale avec un taux de panne croissant. Dans le concept de maintenance imparfaite, les activités de MP sont effectuées avec les inspections de contrôle de la qualité afin de réduire les coûts liés au contrôle de la qualité. La variable de décision considérée est le nombre d'inspections m minimisant le coût total moyen par unité de temps. [Suliman et Jawad, 2012] ont étendu le modèle dans [Ben-Daya, 2002] en supposant que la défaillance du système peut survenir à n'importe quel instant après qu'il passe à l'état 'hors contrôle'. Les auteurs ont également considéré le coût de la pénurie et une durée de MP non négligeable.

[Chakraborty et al., 2009] ont cherché à optimiser la QEP et le nombre d'inspections en considérant un système de production imparfait soumis à des pannes aléatoires. [Chakraborty et Giri, 2012] ont pris en compte la possibilité de dégradation du processus pendant la production après une durée aléatoire. La stratégie de maintenance proposée consiste à effectuer une action de MP à la fin de chaque cycle de production et une action de MC en cas de panne. Les auteurs ont suggéré un algorithme de résolution pour déterminer les valeurs optimales du stock de sécurité et de la longueur du cycle de production. [Chelbi et al., 2008] et [Dhouib et al., 2012] ont proposé une approche intégrée de contrôle de la production et de la maintenance pour un système de production imparfait constitué d'un seul équipement et qui doit satisfaire une demande constante et continue. L'unité manufacturière peut passer à l'état 'hors contrôle' après une période aléatoire de production où elle commence à produire des pièces non conformes. Les non-conformités sont principalement causées par la dégradation de la machine. Les auteurs ont établi une stratégie de MP basée sur l'âge afin de réduire le risque d'occurrence du passage à l'état 'hors contrôle'. Au cas où cela se produirait, une action de MC est programmée pour restaurer la machine à l'état 'As Good As New'. Le modèle développé permet de déterminer la taille optimale du stock tampon et l'âge auquel une action de MP doit être effectuée. [Radhoui et al., 2010a] ont proposé un modèle intégré pour un système de production imparfait constitué d'un seul équipement et qui doit satisfaire une demande constante et continue. Chaque lot produit est systématiquement soumis à un contrôle de qualité automatisé de coût et de durée négligeable afin de déterminer le taux d'unités non conformes. En fonction du taux d'unités non conformes détecté, il est décidé d'entreprendre ou non des actions de maintenance. Les auteurs ont développé un modèle mathématique afin de minimiser le coût total moyen par unité de temps. Trois variables de décision caractérisent la stratégie proposée : le taux des unités non-conformes à partir duquel des actions de maintenance devraient être entreprises, la période d'inspection selon laquelle le taux de non-conformité doit être mesuré, et la taille du stock tampon à construire pour pallier les perturbations causées par les arrêts de la production pour effectuer les actions de maintenance. [Radhoui et al., 2010b] ont repris le même principe du modèle établi par [Radhoui et al., 2010a] et ont développé une approche de modélisation basée sur la simulation et les plans d'expérience et en considérant une maintenance imparfaite. Les variables de décision considérées sont le taux d'unités non conformes sur la base duquel des actions de maintenance préventive doivent être exécutées, et la taille du stock tampon. [Hajji et al., 2011] considèrent conjointement le contrôle de la production et la conception économique des produits dans un système de fabrication. Le système est sujet à des défaillances aléatoires de durées aléatoires. Ils ont supposé que les articles produits

sont inspectés et qu'une proportion d'items non conformes est produite. Les variables de décision prises en compte sont le taux de production et les spécifications techniques du produit.

[Chen, 2013] a traité un problème intégré de production-inventaire, de MP et d'inspection pour un processus de production imparfait. Des inspections périodiques sont effectuées dans le but de surveiller l'état du système de production. Si le processus est à l'état de contrôle, une action de MP imparfaite d'une durée négligeable est effectuée. Lorsque le processus est à un état hors contrôle, une action de restauration est mise en place. Cette étude suppose qu'une certaine proportion de produits non conformes peut être retravaillée. Le modèle mathématique développé consiste à déterminer la QEP, le nombre optimal d'inspections, les périodes d'inspection optimales et la politique de MP optimale minimisant le coût total moyen. [Rivera-Gómez et al., 2013] ont étendu le modèle précédent et ont déterminé également une stratégie de MP optimale. [Pal et al., 2013] ont développé un modèle mathématique pour déterminer la QEP pour une demande stochastique. Après une période de production régulière, les articles produits de mauvaise qualité sont retravaillés et des actions de MP périodiques sont effectuées. Les auteurs ont cherché les valeurs optimales du stock tampon, du taux de production et du temps de production. [Liao, 2013] a considéré les mêmes hypothèses de [Pal et al., 2013] pour déterminer le temps de production optimal pour une unité de fabrication imparfaite avec une demande déterministe. Rivera-[Beheshti Fakher et al., 2017] ont étudié un système composé de machines parallèles qui se détériorent avec le temps et pouvant passer à un état dégradé avec un taux de défaut et de défaillance élevé. Dans ce système, la planification de la maintenance correspond à une maintenance imparfaite de type âge. Les auteurs ont proposé un modèle intégré pour la détermination des décisions concernant la taille des lots, l'âge auquel une action de MP est entreprise et les intervalles d'inspection minimisant le coût total pour un système multi-produits. Ce modèle est basé sur la recherche Tabou et les algorithmes génétiques. [Beheshti Fakher et al., 2018] ont repris le modèle précédent et ont maximisé le profit attendu en considérant une stratégie de MP périodique. Une action de restauration est effectuée si l'inspection révèle que le système a produit des articles non conformes ou en cas de panne. Les travaux susmentionnés sont basés sur le passage aléatoire du processus à un état 'hors-contrôle' après une durée aléatoire. Ils ne considèrent pas un outil de contrôle de la qualité et ses paramètres.

[Pandey et al., 2011] ont établi un modèle analytique pour l'optimisation conjointe de l'ordonnancement de la production, de la période de MP et des paramètres de la carte de contrôle pour un système manufacturier imparfait. Une action de MC est réalisée lorsque la machine tombe en panne ou si la moyenne d'un échantillon dépasse les limites de contrôle de la carte. Le modèle proposé ne tient pas compte du stock. Ainsi, la demande reçue pendant les interventions de maintenance n'est pas satisfaite. [Bouslah et al., 2013] ont formulé un modèle dynamique stochastique pour les systèmes de production imparfaits, où le contrôle de la qualité des lots produits est effectué à l'aide d'un plan d'échantillonnage par acceptation. Les auteurs supposent qu'une action de restauration est réalisée lorsqu'une défaillance survient au cours du cycle de production. L'objectif est d'obtenir simultanément le dimensionnement du lot et le taux de production minimisant le coût total. La proportion d'articles défectueux, les pannes et les délais de réparation suivent les distributions générales. Le modèle développé ne prend pas en

considération la MP. [Bouslah et al., 2014] ont repris le modèle précédent en considérant un stock tampon du produit fini pour éviter les pénuries pendant les actions de MC. L'objectif est de déterminer simultanément la QEP, la taille optimale du stock tampon et le plan d'échantillonnage optimal. Un modèle mathématique stochastique est développé en utilisant une approche d'optimisation basée sur la simulation. Plus récemment, [Bouslah et al., 2016] ont étendu le modèle précédent en intégrant les politiques de MP avec le contrôle qualité par les plans d'échantillonnage. Lorsque la proportion d'articles défectueux dans un lot atteint ou dépasse un certain seuil, une action de MC doit être entreprise. L'objectif de ce travail est d'optimiser le stock tampon, la période de MP et les paramètres du plan d'échantillonnage, en tenant compte d'une contrainte qui est la limite moyenne de la qualité sortante (AOQL). [Bahria et al., 2018] ont proposé une stratégie intégrée de contrôle de la maintenance et des cartes de contrôle qui vise à minimiser le coût total moyen par unité de temps. Les variables de décision considérées dans ce travail sont : la taille de l'échantillon, l'intervalle d'échantillonnage et les limites de la carte de contrôle.

Dans cet article, nous proposons une nouvelle stratégie conjointe de production-qualité-maintenance en considérant une politique de maintenance préventive périodique. Ce modèle permet d'intégrer et d'optimiser le stock tampon, la période de MP et les paramètres de la carte de contrôle simultanément, à savoir : la taille de l'échantillon, la période d'échantillonnage et les limites de contrôle. Ainsi, ce modèle vise à assurer la satisfaction continue de la demande tout en compensant la consommation moyenne durant les périodes de maintenance, ainsi que les unités non-conformes rejetées au cours de la période qui suit le passage à l'état 'hors contrôle'. De plus, la taille du lot de production Q peut être déduite à partir de la valeur optimale du stock tampon. En effet, ce modèle consiste à suivre l'influence des actions de maintenance sur l'évolution du stock. Ainsi, la contribution majeure de ce travail consiste à considérer et optimiser la taille du stock tampon avec une politique de MP périodique et les cartes de contrôle.

Le reste de ce papier est organisé comme suit : la section 2 présente la politique intégrée proposée, les hypothèses de travail considérées et les notations que nous allons utiliser pour modéliser le problème étudié. Dans la section 3, nous présentons le modèle mathématique développé exprimant le coût total moyen par unité de temps. La section 4 est dédiée à un exemple numérique. Une analyse de sensibilité est illustrée dans la section 5. Le papier se termine par une conclusion dans la section 6.

2 STRATEGIE INTEGREE, HYPOTHESES DE TRAVAIL ET NOTATIONS

2.1 Stratégie intégrée

Nous considérons un système de fabrication composé d'une seule machine et produisant un seul type de produit et devant satisfaire une demande constante et continue D . Ce système est sujet à un processus de dégradation avec un taux de panne croissant, ce qui est à l'origine de la production d'unités non conformes. Chaque lot produit est soumis à un contrôle qualité en utilisant une carte de contrôle moyenne-écart type. Nous supposons que l'état de la dégradation de la machine présente la cause principale et dominante de non-maîtrise du processus, et par la suite du dépassement des limites de contrôle de la

carte. Ainsi, l'état de l'unité manufacturière peut être évalué en mesurant l'indicateur de qualité x du produit fabriqué.

Pendant que la machine produise, un échantillon de taille n est prélevé et contrôlé à chaque I_s unités de temps pour suivre l'évolution de l'indicateur x . Les résultats de l'échantillonnage sont enregistrés dans une carte de contrôle 'x-bar'. Un cycle de production commence par un système supposé à l'état 'sous-contrôle', produisant des articles de qualité acceptable au taux de production maximal P qui est supérieur à la demande D . A chaque fois que le processus atteint un niveau de stock maximal h , le taux de production est ramené égal à D . On suppose que la moyenne, μ , et l'écart type, σ , du processus sont connus ou bien estimés. Lorsque le processus est sous contrôle, la moyenne des mesures de l'échantillon de l'indicateur de qualité est à sa valeur cible entre la limite de contrôle inférieure LCL et la limite de contrôle supérieure UCL .

Les limites de la carte de contrôle 'x-bar' sont données par :

$$LCL = \mu - \frac{k}{\sqrt{n}} \times \sigma \quad (1)$$

$$UCL = \mu + \frac{k}{\sqrt{n}} \times \sigma \quad (2)$$

Afin de réduire la probabilité de passer à l'état 'hors-contrôle', des actions de MP de durées négligeables sont planifiées pendant le temps de fonctionnement dans l'état 'sous-contrôle'. Nous supposons que les actions de MP sont effectuées à tous les $\alpha \times I_s$ unités de temps, où α est un nombre entier différent de zéro et I_s est l'intervalle d'échantillonnage (Figure 1). La moyenne du processus peut passer à l'état 'hors-contrôle', en raison de la dégradation de la machine, de μ à $\mu + \delta \times \sigma$ pour un échantillon J . Dans ce cas, la moyenne de l'échantillon des mesures d'indicateur de qualité est située au-delà des limites de contrôle ($\bar{x}_e < LCL$ ou $\bar{x}_e > UCL$). Par conséquent, la production est interrompue et une action de maintenance corrective de durée moyenne μ_c doit être effectuée (Figure 2).

De plus, tous les éléments produits pendant le dernier intervalle d'échantillonnage avant la transition vers l'état hors contrôle (entre les échantillons $(J-1)$ et J) sont rejetés. De même, afin de pallier les perturbations causées par l'interruption de la production lors des actions de restauration, un stock tampon h est constitué au début de chaque cycle de production pour assurer la continuité de la chaîne de production. Nous considérons que le niveau h du stock est atteint avant le passage à l'état 'hors-contrôle'. Le stock tampon est renouvelé chaque fois qu'il est totalement consommé.

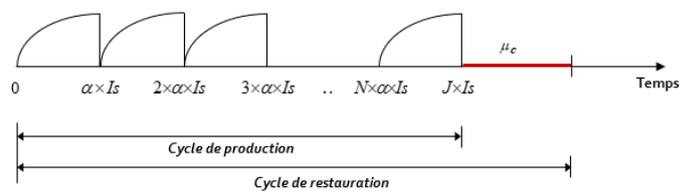


Figure 1. La stratégie de maintenance du système

La Figure 2 illustre la politique intégrée de contrôle qualité et de maintenance dans le cas d'une seule action de MP pendant un cycle de production.

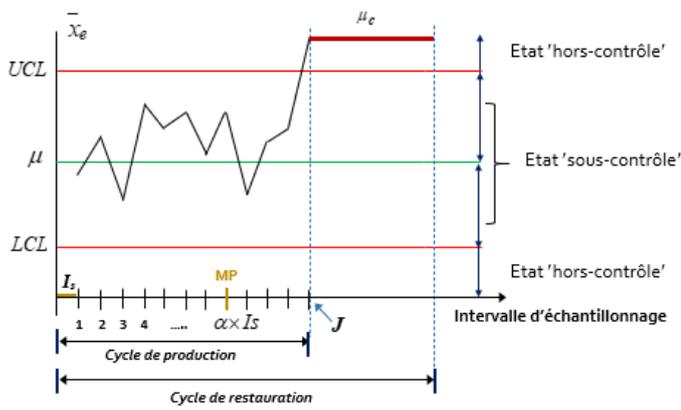


Figure 2. Illustration de la stratégie intégrée de contrôle qualité et de maintenance

2.2 Hypothèses

- (H1) Les articles produits sont périssables avec le temps.
- (H2) L'indicateur qualité contrôlé est distribué suivant la loi Normale.
- (H3) Les demandes qui ne peuvent être satisfaites sont perdues.
- (H4) La non-conformité des produits est uniquement due à la dégradation de la machine.
- (H5) Les articles non conformes ne sont pas réinsérés dans le processus de production.
- (H6) Chaque action de maintenance renvoie l'unité de production à la condition « As Good As New ».
- (H7) Les coûts unitaires liés à l'inventaire, à la maintenance et à la qualité sont connus et constants.
- (H8) Les ressources pour effectuer les actions de maintenance sont toujours disponibles.

2.3 Notations

- Q Taille du lot de production
- $q(t)$ Niveau du lot à l'instant t
- $bs(t)$ Niveau du stock à l'instant t
- h Taille du stock tampon
- P Taux de production maximal
- D Taux de demande
- x_i Mesure de l'indicateur de qualité à contrôler de la pièce i prélevée
- \bar{x}_e Moyenne des mesures de l'indicateur de qualité pour l'échantillon e
- n Taille de l'échantillon
- μ Moyenne du processus
- μ_1 La moyenne qui correspond à la variabilité du processus suite au passage à l'état 'hors-contrôle'
- σ L'écart type du processus
- δ Amplitude moyenne de la dérive de l'indicateur de qualité en cas de dépassement des limites de contrôle
- I_s Intervalle d'échantillonnage
- $q(I_s)$ Niveau du lot à l'instant I_s
- m Nombre d'échantillons contrôlés lors de la production d'un lot
- UCL Limite de contrôle supérieure
- LCL Limite de contrôle inférieure

- k Nombre d'écart-types de la distribution de l'échantillon entre la ligne centrale de la carte de contrôle et les limites de contrôle
- J Nombre moyen d'échantillons pour détecter l'état 'hors-contrôle'
- $F(\cdot)$ Fonction de distribution de la loi Normale centrée réduite
- μ_c Durée moyenne d'une action de maintenance corrective
- p_{sc} Probabilité que le système soit dans l'état 'sous-contrôle'
- p_{hc} Probabilité que le système soit dans l'état 'hors-contrôle'
- β Probabilité de non-détection de l'état 'hors-contrôle'
- C_s Coût de détention d'un item par unité de temps
- C_p Coût de pénurie d'une pièce perdue
- C_i Coût de l'inspection d'un produit
- C_r Coût d'un produit défectueux
- C_{MP} Coût d'une action de MP
- C_{MC} Coût d'une action de MC
- Γ_{setup} Coût moyen de setup
- Γ_{S1} Coût moyen d'inventaire pour le scénario où le stock tampon est atteint avant le passage à l'état hors contrôle de la machine (scénario 1)
- Γ_{S2} Coût moyen d'inventaire pour le scénario où le stock tampon n'a pas été complètement bâti avant le passage à l'état hors contrôle (scénario 2)
- Γ_{stot} Coût total moyen d'inventaire
- Γ_M Coût total moyen de maintenance
- Γ_Q Coût total moyen de qualité
- Γ_{ech} Coût total moyen d'échantillonnage
- Γ_{NC} Coût total moyen de non-conformité des produits
- Γ_{FA} Coût total moyen de fausses alarmes
- Γ_{tot} Coût total moyen par unité de temps

3 MODELE MATHEMATIQUE

Notre objectif est de déterminer les paramètres optimaux de la carte de contrôle qui sont : la taille de l'échantillon n , l'intervalle d'échantillonnage I_s , ainsi que le coefficient des limites de contrôle, k . D'autre part, nous proposons d'optimiser la taille du stock tampon h et l'indice de l'échantillon α auquel l'action de MP doit être entreprise. L'objectif est de minimiser le coût total moyen par unité de temps pendant un cycle de restauration. Le coût total moyen correspond à la somme du coût de setup, du coût total moyen de l'inventaire, du coût total moyen de maintenance et du coût total moyen de la qualité, y compris les coûts d'échantillonnage, de fausses alarmes et de rejet. La politique de production consiste à constituer au début de chaque cycle de production un stock tampon de taille h à un taux $(P - D)$. Le cycle de production se termine lorsque les limites de contrôle de la carte de contrôle sont dépassées. Par conséquent, le cycle de production se termine après $(J \times I_s)$ unités de temps, où J est le nombre moyen d'échantillons pour détecter l'état 'hors contrôle'.

3.1 Durée moyenne du cycle de restauration

La durée moyenne du cycle de restauration RCD est donnée par :

$$RCD = J \times Is + \mu_c \quad (3)$$

Avec J est le nombre moyen d'échantillons pour détecter le passage à l'état 'hors contrôle'. Il est exprimé comme suit [Montgomery, 2004].

$$J = \frac{1}{1-\beta} \quad (4)$$

β étant la probabilité de non-détection de l'état 'hors contrôle'. Dans ce cas, la moyenne de l'indicateur de qualité ne dépasse pas les limites de contrôle ($LCL \leq \bar{x}_e \leq UCL$) alors que le processus est hors de contrôle.

$$\beta = \left\{ \text{Probabilité}(LCL \leq \bar{x}_e \leq UCL) \mid \mu_1 = \mu + \delta \times \sigma \right\} \quad (5)$$

$$\beta = F\left(\frac{UCL - \mu_1}{\sigma/\sqrt{n}}\right) - F\left(\frac{LCL - \mu_1}{\sigma/\sqrt{n}}\right) \quad (6)$$

$$\beta = F(k - \delta \times \sqrt{n}) - F(-k - \delta \times \sqrt{n}) \quad (7)$$

Et à partir des équations (4) et (7), nous avons :

$$J = \frac{1}{1 - F(k - \delta \times \sqrt{n}) + F(-k - \delta \times \sqrt{n})} \quad (8)$$

En combinant les équations (3) et (8), nous obtenons :

$$RCD = \frac{Is}{1 - F(k - \delta \times \sqrt{n}) + F(-k - \delta \times \sqrt{n})} + \mu_c \quad (9)$$

3.2 Probabilités des deux états du système

- p_{sc} est la probabilité que le système soit dans l'état 'sous-contrôle'. Elle est formulée comme suit :

$$p_{sc} = \text{Probabilité}(LCL \leq \bar{x}_e \leq UCL) \quad (10)$$

$$p_{sc} = F\left(\mu + k \times \frac{\sigma}{\sqrt{n}}\right) - F\left(\mu - k \times \frac{\sigma}{\sqrt{n}}\right) \quad (11)$$

- p_{hc} est la probabilité que le système soit dans l'état 'hors-contrôle'. Elle est donnée par :

$$p_{hc} = \left\{ \begin{array}{l} \text{Probabilité}(\bar{x}_e \leq LCL) \\ + \text{Probabilité}(\bar{x}_e \geq UCL) \end{array} \right\} \quad (12)$$

$$p_{hc} = F\left(\mu - k \times \frac{\sigma}{\sqrt{n}}\right) + 1 - F\left(\mu + k \times \frac{\sigma}{\sqrt{n}}\right) \quad (13)$$

3.3 Coût moyen d'inventaire

Deux cas sont possibles pour l'évolution du stock : soit que le temps de consommation du stock est supérieur à la durée d'une action de maintenance corrective μ_c (cas sans pénurie), soit, la durée de consommation du stock est inférieure à celle de l'arrêt de la machine (cas avec pénurie).

▪ Scénario 1 : Cas sans pénurie

La figure 3 illustre l'évolution de l'état du stock au cours d'un cycle de restauration pour le cas sans pénurie.

Afin de trouver le coût de possession, le diagramme d'inventaire est divisé en zones d'inventaire Z_i désignant la quantité d'inventaire moyenne retenue pendant chaque période.

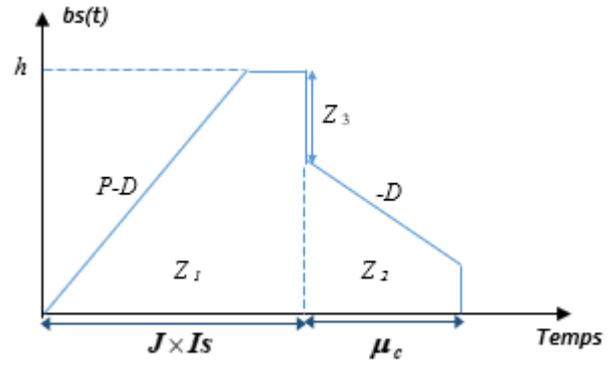


Figure 3. Evolution du niveau du stock : cas sans pénurie

Le coût moyen de possession de stock Γ_{s1} pour le scénario 1 et dans le cas sans rupture est donné par :

$$\Gamma_{s1} = C_s \times \left\{ (p_{sc} \times Z_1) + (p_{hc} \times Z_2) \right\} \quad (14)$$

D'autre part :

$$Z_1 = \frac{h^2}{2 \times (P-D)} + h \times \left(J \times Is - \frac{h}{(P-D)} \right) \quad (15)$$

$$Z_2 = \mu_c \times (h - Z_3) - \frac{\mu_c^2 \times D}{2} \quad (16)$$

Avec Z_3 représente la quantité produite durant le dernier intervalle d'échantillonnage (Is) avant de dépasser les limites de contrôle.

$$Z_3 = D \times Is \quad (17)$$

D'où Z_2 :

$$Z_2 = \mu_c \times (h - D \times Is) - \frac{\mu_c^2 \times D}{2} \quad (18)$$

Ainsi, en considérant ensemble les zones 1, 2 et 3, le coût total moyen de détention du stock est donné par l'expression suivante :

$$\Gamma_{s1} = C_s \times \left\{ \begin{array}{l} p_{sc} \times \left(\frac{h^2}{2 \times (P-D)} + h \times \left(J \times Is - \frac{h}{(P-D)} \right) \right) \\ + p_{hc} \times \left(\mu_c \times (h - D \times Is) - \frac{\mu_c^2 \times D}{2} \right) \end{array} \right\} \quad (19)$$

▪ Scénario 2 : Cas avec pénurie

La figure 4 montre l'évolution du niveau du stock pour un cycle de restauration avec pénurie.

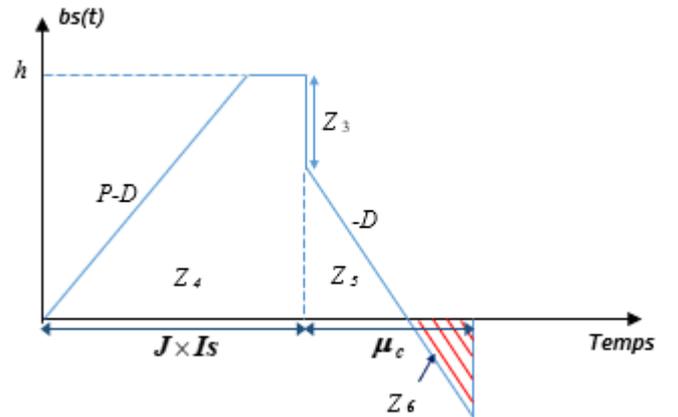


Figure 4. Evolution du niveau du stock : cas avec pénurie

Dans ce cas, le coût moyen d'inventaire Γ_{s2} est exprimé en faisant la somme du coût de possession et celui de rupture du stock :

$$\Gamma_{S2} = \left\{ C_S \times \left((p_{sc} \times Z_4) + (p_{hc} \times Z_5) \right) + C_p \times (p_{hc} \times Z_6) \right\} \quad (20)$$

Avec :

$$Z_4 = \frac{h^2}{2 \times (P - D)} + h \times \left(J \times Is - \frac{h}{(P - D)} \right) \quad (21)$$

La quantité d'inventaire restante à la fin de la période ($J \times Is$) est égale à $(h - Z_3)$. Ainsi :

$$Z_5 = \frac{(h - Z_3)^2}{2 \times D} \quad (22)$$

En combinant les équations (17) et (22), nous obtenons :

$$Z_5 = \frac{(h - D \times Is)^2}{2 \times D} \quad (23)$$

$$Z_6 = \left(\mu_c - \frac{(h - Z_3)}{D} \right) \times D \quad (24)$$

Donc :

$$Z_6 = \left(\mu_c - \frac{(h - D \times Is)}{D} \right) \times D \quad (25)$$

Ce qui permet d'écrire :

$$\Gamma_{S2} = \left\{ \begin{array}{l} C_S \times \left(\left(p_{sc} \times \left(\frac{h^2}{2 \times (P - D)} + h \times \left(J \times Is - \frac{h}{(P - D)} \right) \right) \right) \right) \\ + \left(p_{hc} \times \frac{(h - D \times Is)^2}{2 \times D} \right) \\ + C_p \times \left(p_{hc} \times \left(\mu_c - \frac{(h - D \times Is)}{D} \right) \times D \right) \end{array} \right\} \quad (26)$$

Soit φ_1 l'indicateur de pénurie.

$$\varphi_1 = \text{indicateur}(\mu_c > \frac{h - D \times Is}{D}) = \begin{cases} 1 & \text{si } \mu_c > \frac{h - D \times Is}{D} \\ 0 & \text{sinon} \end{cases} \quad (27)$$

Le coût total moyen de l'inventaire Γ_S tenant compte des deux cas (sans et avec pénurie) est exprimé par :

$$\Gamma_{Stor} = \Gamma_{S1} \times (1 - \varphi_1) + \Gamma_{S2} \times \varphi_1 \quad (28)$$

La taille du lot de production Q présente la quantité produite pendant le temps de production réel pendant un cycle de restauration. En se référant à la figure 3, nous pouvons soulever que la machine produit à son taux de production maximal P avant d'atteindre la taille maximale d'inventaire h , c'est-à-dire pendant la période $\left(\frac{h}{P - D} \right)$. Lorsque le niveau de

stock maximal h est bâti, la machine produit seulement pour satisfaire la demande D avant de passer à l'état 'hors-contrôle', c'est-à-dire pendant la période $\left(J \times Is - \frac{h}{P - D} \right)$.

Par la suite, la taille du lot de production Q est exprimée par :

$$Q = P \times \frac{h}{P - D} + D \times \left(J \times Is - \frac{h}{P - D} \right) \quad (29)$$

3.4 Coût total moyen de maintenance

L'expression du coût total moyen de maintenance est donnée comme suit pour tous les scénarios :

$$\Gamma_M = C_{MP} \times \left(\frac{J}{\alpha} \right) + C_{MC} \quad (30)$$

Où $(.)$ représente la fonction de partie entière.

3.5 Coût total moyen de qualité

Le coût moyen total de la qualité correspond à la somme du coût relatif aux fausses alarmes, du coût d'échantillonnage et du coût des unités rejetés pendant un cycle de production.

$$\Gamma_Q = \Gamma_{FA} + \Gamma_{éch} + \Gamma_{NC} \quad (31)$$

3.4.1 Coût moyen de fausses alarmes

La notion de fausse alarme est introduite dans le cas de détection d'un dépassement des limites de la carte bien que le processus soit sous contrôle. Dans ce cas, un coût supplémentaire appelé coût d'une fausse alarme est intégré dans le modèle. Il est donné par [Bahria et al., 2018] :

$$\Gamma_{FA} = 2 \times C_F \times \left(\sum_{i=0}^m i \times (p_{sc})^i \times (1 - p_{sc}) \right) \times F(-k) \quad (32)$$

3.4.2 Coût moyen d'échantillonnage

Le coût total moyen de l'échantillonnage et de l'inspection est exprimé par :

$$\Gamma_{éch} = C_i \times n \times J \quad (33)$$

3.4.3 Coût moyen de rejet

En se référant à la figure 3, la quantité produite durant le dernier intervalle d'échantillonnage avant de passer à l'état 'hors-contrôle' est égale à Z_3 . Ainsi, d'après l'équation 17, le coût moyen de non-conformité associé au premier scénario est donné par :

$$\Gamma_{NC} = C_r \times D \times Is \quad (34)$$

Par la suite, le coût moyen de la qualité est donné par :

$$\Gamma_Q = \left\{ \begin{array}{l} 2 \times C_F \times \left(\sum_{i=0}^m i \times (p_{sc})^i \times (1 - p_{sc}) \right) \times F(-k) \\ + C_i \times n \times J + C_r \times D \times Is \end{array} \right\} \quad (35)$$

3.4.4 Coût total moyen par unité du temps

En combinant les équations (28), (30) et (35), et en divisant par la durée moyenne d'un cycle de restauration donnée par l'équation (3), nous obtenons le coût total moyen par unité de temps, comme suit :

$$\Gamma_{tot} = \left\{ \begin{array}{l} \frac{1}{(J \times h + \mu_c)} \\ \left(C_S \times \left(p_{sc} \times \left(\frac{h^2}{2 \times (P - D)} + h \times \left(J \times Is - \frac{h}{(P - D)} \right) \right) \right) \right) \\ + C_S \times \left(p_{hc} \times \left(\frac{\mu_c \times (h - D \times Is)}{2} \right) \right) \times (1 - \varphi_1) \\ \left(C_S \times \left(p_{hc} \times \frac{(h - D \times Is)^2}{2 \times D} \right) \right) \\ + C_p \times \left(p_{hc} \times \left(\mu_c - \frac{(h - D \times Is)}{D} \right) \times D \right) \\ + C_{PM} \times \left(\frac{J}{\alpha} \right) + C_{CM} \\ + 2 \times C_F \times \left(\sum_{i=0}^m i \times (p_{sc})^i \times (1 - p_{sc}) \right) \times F(-k) \\ + C_i \times (n \times J) + C_r \times (D \times h) \end{array} \right\} \times \varphi_1 \quad (36)$$

4 EXEMPLE NUMERIQUE

Afin d'illustrer notre approche, nous avons simulé un exemple de simulation dont les paramètres sont choisis arbitrairement tout en considérant des cadres réalistes : $\mu = 5$, $\sigma = 1.5$, $P = 30$ unités/heure, $D = 20$ unités/heure, $\mu_c = 2$ heures, $C_s = 0.1$ \$/unité/heure, $C_p = 5$ \$/unité/heure, $C_{MP} = 200$ \$, $C_{MC} = 1000$ \$, $C_i = 15$ \$/produit, $C_r = 70$ \$/unité défectueuse, $\delta = 0.8$. Une procédure d'optimisation numérique itérative a été développée et codée avec le logiciel MATLAB®.

La meilleure stratégie obtenue préconise de prélever un échantillon de taille 62 toutes les 5,8 heures. En outre, en ce qui concerne la conception de la carte de contrôle, le nombre optimal d'écart-types entre la ligne centrale et les limites de contrôle est 2,6 (Figure 5). Une action de maintenance préventive est effectuée en moyenne à tous les 14 échantillons, c'est-à-dire à toutes les 81,2 heures. De plus, le passage à l'état 'hors-contrôle' survient en moyenne après $J = 59$ échantillons. Par conséquent, le temps prévu pour passer à l'état 'hors-contrôle' est égal à 342,2 heures. Pendant cette période, 3 actions de maintenance préventive en moyenne sont entreprises durant un cycle de production. Le cycle de restauration moyen RCD correspondant à la stratégie optimale est de 344,2 heures. Il sera également nécessaire de bâtir un stock tampon de 274 unités afin de continuer à satisfaire la demande lorsque la production est arrêtée pour effectuer une action de restauration (Figure 6).

La taille du lot à produire entre deux actions de restauration consécutives est égale à 3970 unités. Ce faisant, il en coûterait en moyenne 186,04 (\$ / h), ce qui correspond au coût moyen total minimum par unité de temps.

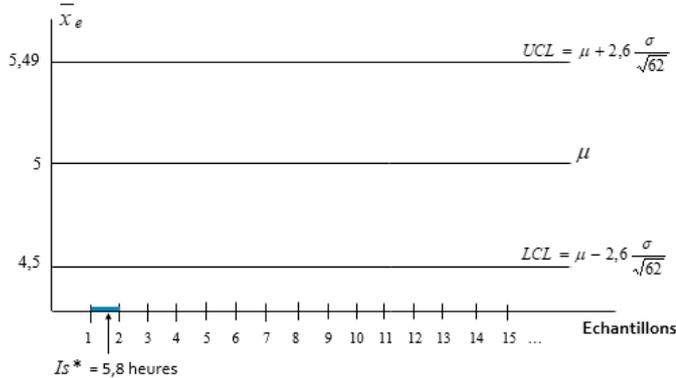


Figure 5. Paramètres de conception optimale de la carte de contrôle

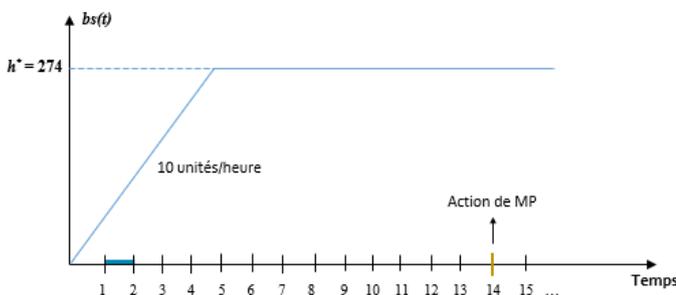


Figure 6. Evolution du niveau du stock pour la stratégie optimale

5 ANALYSE DE SENSIBILITE

Des expériences additionnelles ont été menées afin d'analyser et valider le comportement du système. Le tableau de l'annexe 1 présente les cas de variation des paramètres du modèle (coûts

unitaires et paramètres de maintenance et de la qualité). Les effets de ces variations sur la solution optimale de base et sur le coût total moyen optimal sont expliqués par la suite.

5.1 Variation du coût de détention d'une pièce par unité de temps

Lorsque le coût de détention unitaire augmente, le stock tampon h^* diminue dans le but de réduire le coût total moyen de stockage. De même, la taille optimale du lot de production Q diminue. Parce que la durée moyenne de stockage diminue proportionnellement à la diminution de la quantité d'inventaire, la fréquence d'échantillonnage I_s^* et le nombre d'échantillons prélevés J^* durant un cycle de restauration diminuent, réduisant ainsi la durée moyenne de stockage. De ce fait, le temps moyen pour effectuer une action de MC diminue. Cependant, une fréquence d'échantillonnage plus faible implique que le dépassement des limites de contrôle de la carte est détecté plus rapidement, justifiant la sévérité de la carte de contrôle par la réduction du nombre d'écart types entre la ligne centrale et les limites de la carte k^* . En outre, la taille de l'échantillon augmente de manière à obtenir plus d'informations sur le processus.

5.2 Variation du coût d'inspection unitaire

Lorsque le coût d'inspection d'une pièce augmente, le nombre d'items inspectés n^* diminue afin de minimiser le coût total d'échantillonnage. De plus, l'intervalle d'échantillonnage I_s^* augmente afin d'effectuer les inspections moins fréquemment. D'autre part, la carte de contrôle devient moins serrée et les limites de la carte s'éloignent (augmentation de k^*). Les actions de maintenance préventive sont entreprises moins fréquemment (augmentation de α), et le temps moyen pour effectuer une action de MC augmente (augmentation de J) c'est pourquoi la taille du stock tampon h^* nécessaire diminue.

5.3 Variation du coût de rejet d'une unité non conforme

Quand le coût de rejet d'une unité non conforme augmente, l'intervalle d'échantillonnage diminue de façon à réduire la quantité des pièces non conformes $D \times I_s^*$. De même, la taille de l'échantillon et la fréquence d'échantillonnage augmentent afin de contrôler plus le processus. De plus, la carte de contrôle devient plus serrée et on a plus de tendance à dépasser les limites de la carte. De ce fait, le temps moyen pour entreprendre des actions de MC augmente (augmentation de J) accroissant la taille du stock tampon. De plus, la période de MP diminue augmentant la fréquence d'effectuer des actions de MP.

5.4 Variation du coût d'une action de MP

Lorsque le coût d'une action de MP augmente, le paramètre α augmente réduisant la fréquence d'entreprendre des actions de MP. En outre, la fréquence d'effectuer des actions de MC diminue conduisant à la réduction de la taille du stock tampon. De plus, le coefficient k augmente et les limites de contrôle de la carte s'éloignent. Ainsi, le processus est moins sévère, ce qui est à l'origine de l'augmentation de la taille de l'échantillon et de la diminution de l'intervalle d'échantillonnage.

5.5 Variation de la durée moyenne d'une action de MC

Lorsque la durée d'une action de MC augmente, le système réagit de façon à réduire le temps moyen pour effectuer des actions de MC en augmentant le paramètre J . De plus, la variable k^* augmente étendant la marge entre les limites de la

carte de contrôle. Ainsi, le processus est plus serré ce qui implique l'augmentation de la taille des échantillons prélevés et de la fréquence d'échantillonnage.

La taille du stock tampon diminue vu que la durée moyenne pour effectuer des actions de MC est réduite. De même, la variable α augmente effectuant ainsi des actions de MP moins fréquemment.

5.6 Variation de l'amplitude la dérive suite au passage à un état de hors de contrôle par rapport à la ligne centrale

L'augmentation du paramètre δ signifie que la moyenne des mesures de l'indicateur qualité va s'éloigner de la ligne centrale de la carte de contrôle. Par la suite, le processus est plus dégradé. De ce fait, la carte de contrôle devient plus serrée (diminution de la variable k^*). Ainsi, nous avons tendance à dépasser les limites de contrôle de la carte ce qui explique la diminution du paramètre J . Par conséquent, le temps moyen pour effectuer des actions de MC augmente ainsi que la taille optimale du stock tampon pour avantager la protection contre la pénurie. En outre, les actions de MP sont entreprises plus fréquemment (diminution de α).

6 CONCLUSION

Ce papier présente une stratégie conjointe de contrôle de la qualité, de la maintenance et de la production pour un processus de production imparfait produisant des unités conformes et non conformes. Le but de cette stratégie combinée est de révéler l'effet synergique entre la dégradation de l'équipement, les paramètres de la carte de contrôle et la politique de gestion des stocks. Cinq variables de décision caractérisent la stratégie proposée : la taille de l'échantillon, l'intervalle d'échantillonnage, le coefficient des limites de contrôle, la taille du stock tampon et la période de maintenance préventive. Un modèle mathématique a été développé afin de trouver la stratégie optimale qui minimise le coût total moyen par unité de temps. Un exemple numérique a été proposé et une analyse de sensibilité du modèle est présentée afin d'étudier l'effet de la variation des différents paramètres sur le comportement du système. De nouvelles extensions consistent à travailler avec des systèmes produisant différents types de produits avec une demande variable. Nous pouvons aussi intégrer un système de reprise des items défectueux. Une autre perspective consiste à étudier une stratégie de maintenance préventive non périodique.

6 REFERENCES

- Bahria, N., Chelbi, A., Dridi, I.H. and Bouchriha, H. (2018). Maintenance and quality control integrated strategy for manufacturing systems, *European Journal of Industrial Engineering*, vol. 12(3), pp. 307-331.
- Beheshti Fakher, H., Nourelfath, M., & Gendreau, M. (2018). Integrating production, maintenance and quality: A multi-period multi-product profit-maximization model. *Reliability Engineering & System Safety*, vol. 170, pp. 191-201.
- Beheshti Fakher, H., Nourelfath, M., & Gendreau, M. (2017). A cost minimisation model for joint production and maintenance planning under quality constraints. *International Journal of Production Research*, vol. 55(8), pp. 2163-2176.
- Ben-Daya, M. (2002). The economic production lot-sizing problem with imperfect production processes and imperfect maintenance. *International Journal for Production Economics*, vol. 76, pp. 257-264.
- Bouslah, B., Gharbi, A., & Pellerin, R. (2014). Joint production and quality control of unreliable batch manufacturing systems with rectifying inspection. *International Journal of Production Research*, vol. 52(14), pp. 4103-4117.
- Bouslah, B., Gharbi, A., & Pellerin, R. (2016). Integrated production, sampling quality control and maintenance of deteriorating production systems with AOQL constraint. *Omega*, vol. 61, pp. 110-126.
- Bouslah, B., Gharbi, A., Pellerin, R. (2013). Joint optimal lot sizing and production control policy in an unreliable and imperfect manufacturing system. *International Journal of Production Economics*, vol. 144, pp. 143-156.
- Chakraborty, T., Giri, B., & Chaudhuri, K. (2009). Production lot sizing with process deterioration and machine breakdown under inspection schedule. *Omega*, vol. 37(2), pp. 257-271.
- Chakraborty, T., Giri, B.C., (2012). Joint determination of optimal safety stocks and production policy for an imperfect production system. *Applied Mathematical Modeling*, vol. 36, pp. 712-722.
- Chelbi, A., Rezg, N., & Radhoui M. (2008). Simultaneous determination of production lot size and preventive maintenance schedule for unreliable production system. *Journal of Quality in Maintenance Engineering*, vol. 14(2), pp. 161-176.
- Chen, Y. C. (2013). An optimal production and inspection strategy with preventive maintenance error and rework. *Journal of Manufacturing Systems*, vol. 32(1), pp. 99-106.
- Dhouib, K., Gharbi, A., Ben Aziza, M.N., (2012). Joint optimal production control/preventive maintenance policy for imperfect process manufacturing cell. *International Journal of Production Economics*, vol. 137, pp. 126-136.
- Hajji, A., Mhada, F., Gharbi, A., Pellerin, R., & Malhame, R. (2011). Integrated product specifications and productivity decision making in unreliable manufacturing systems. *International Journal for Production Economics*, vol. 129, pp. 32-42.
- Liao G-L. (2013). Optimal economic production quantity policy for randomly failing process with minimal repair, backorder and preventive maintenance. *International Journal of Systems Science*, vol. 44(9), pp. 1602-1612.
- Montgomery, D. C. (2004) Introduction to statistical quality control, John Wiley & Sons, New York, NY.
- Pal, B., Sana, S. S., & Chaudhuri, K. (2013). A mathematical model on EPQ for stochastic demand in an imperfect production system. *Journal of Manufacturing Systems*, vol. 32(1), pp. 260-270.
- Pandey, D., Kulkarni, M.S., Vrat, P. (2011). A methodology for joint optimization for maintenance planning, process quality and production scheduling. *Computers and Industrial Engineering*, vol. 61, pp. 1098-1106.
- Radhoui, M., Rezg, N., Chelbi, A., (2010a). Joint quality control and preventive maintenance strategy for imperfect production processes. *Journal of Intelligent Manufacturing*, vol. 21, pp. 205-212.
- Radhoui, M., Rezg, N., Chelbi, A. (2010b). Integrated maintenance policy based on quality. *Computers and Industrial Engineering*, vol. 58, pp. 443-451.
- Rivera-Gómez, H., Gharbi, A., & Kenné, J. P. (2013). Joint production and major maintenance planning policy of a manufacturing system with deteriorating quality. *International Journal of Production Economics*, vol. 146(2), pp. 575-587.
- Suliman, S. M., & Jawad, S. H. (2012). Optimization of preventive maintenance schedule and production lot size. *International Journal of Production Economics*, vol. 137(1), pp. 19-28.

Annexe 1 : Analyse de sensibilité des paramètres

Cas	Paramètre	Variation	n^*	\bar{b}^*	k^*	h^*	α^*	J	Q	RCD	Γ_{Stot} (\$)	Γ_M (\$)	Γ_{NC} (\$)	Γ_{FA} (\$)	Γ_{ich} (\$)	Γ_{tot} (\$/h)
Base			62	5,8	2,6	274	14	59	3970	344,2	36025,6	1600	8120	0,082	18290	186,04
1	C_s	-0,2	56	7,2	2,9	358	18	67	10006	484,4	42421,7	1600	10080	0,14	18760	150,42
2		-0,1	58	6,5	2,8	312	17	62	8372	405	40150,4	1600	9100	0,12	17980	169,95
3		+0,1	63	5,4	2,5	234	12	54	6066	293,6	29896,9	1800	7560	0,073	17010	191,64
4		+0,2	67	4,9	2,2	219	11	48	4923	237,2	25133,7	1800	6860	0,068	16080	210,26
5	C_f	-2	75	4,3	2,2	405	9	44	4189	191,2	24377,5	1800	6020	0,064	9900	220,17
6		-1	68	5,1	2,4	346	12	51	5548	262,1	25829,4	1800	7140	0,069	13872	185,58
7		+1	55	6,4	2,7	258	16	66	8706	424,4	44391,1	1800	8960	0,16	21780	181,27
8		+2	50	7,7	2,8	224	20	71	11158	548,7	52784,8	1600	10780	0,24	24850	164,05
9	C_T	-20	47	10,4	2,9	126	23	68	14270	709,2	28373,3	1400	10400	0,31	15980	79,17
10		-10	53	8,1	2,7	203	18	64	10571	520,4	38419,4	1600	9720	0,22	16960	128,16
11		+10	76	4,9	2,3	394	11	56	5882	276,4	42191,9	2000	7840	0,063	21280	265,23
12		+20	83	4	2,2	478	7	51	4558	206	39510,1	2400	7200	0,051	21165	341,14
13	C_{MP}	-100	90	2,5	2	549	7	38	2449	97	12220,2	1500	3500	0,062	17100	353,82
14		-50	78	4,9	2,4	353	10	46	4861	227,4	25738,1	1200	6860	0,074	17940	227,52
15		+50	58	6,4	2,7	209	18	67	8785	430,8	36295,9	1150	8960	0,16	19430	152,82
16		+100	55	6,8	2,8	160	21	74	10322	512,6	30424,7	1300	9520	0,42	20350	120,16
17	μ_c	-1	84	3,1	1,9	385	9	41	2927	129,1	12954,6	1400	4340	0,033	17220	278,19
18		-0,5	76	4,5	2,3	316	12	46	4548	209	19248,9	1300	6300	0,068	17480	212,1
19		+0,5	57	7,1	2,8	242	17	70	10182	499	49883,9	1400	9940	0,79	19950	162,67
20		+1	49	8,3	3	214	22	84	14158	699,2	64644,2	1300	11620	1,34	20580	140,36
21	δ	-0,2	38	7,6	3	312	21	68	10648	518,8	57725,3	1300	10640	1,65	12920	159,18
22		-0,1	46	6,3	2,9	293	17	64	8485	405,2	50677,2	1300	8820	1,22	14720	186,37
23		+0,1	65	4,7	2,4	238	12	56	5502	265,2	21268,3	1500	6580	0,058	18200	179,29
24		+0,2	74	4	2,1	195	9	51	4275	206	12532,2	1500	5600	0,034	18870	186,9

Estimating ignition probability of gas pipeline rupture using artificial neural networks modeling

KARIM YOUNSI ^{1,2}, MERZAK NEDJAR ², ABDELNACER SMATI ²

¹ LABORATOIRE DE FIABILITE DES EQUIPEMENTS PETROLIERS, FACULTE DES HYDROCARBURES ET DE LACHIMIE, UNIVERSITE DE BOUMERDES
Avenue de l'indépendance, Boumerdes, 35000, Algérie
karim_zd@yahoo.fr

² PEGAZ ENGINEERING
Rue Coop el Bina N° 80 Dely Ibrahim 16000, Algérie
merzakedjar@yahoo.fr, a_smati@yahoo.com

Résumé – L'estimation de la probabilité d'inflammation est un paramètre clé de la méthodologie d'évaluation quantitative des risques (QRA). En termes de sécurité humaine et de dommages matériels, les conséquences associées aux ruptures enflammées de gazoducs peuvent être extrêmement graves. Cependant, il existe une grande variation dans les valeurs de probabilité d'inflammation données par les bases de données et les modèles disponibles. Dans cet article, l'objectif de l'étude est de développer un nouveau modèle de probabilité d'inflammation après rupture des gazoducs qui surmonte les inconvénients des modèles existants en utilisant l'intelligence artificielle et la modélisation par réseaux de neurones.

Abstract – estimating the probability of ignition turns a key parameter of the quantitative risk assessment methodology (QRA). In terms of human safety and property damage, consequences associated with ignited ruptures of gas pipelines can be extremely severe. However, there is a wide variation in the values of ignition probability given by the available data sources and models. In this paper, the objective of the study is to develop a new ignition probability model for rupture of transmission gas pipelines which overcomes the drawbacks of existing models by using artificial intelligence and neural networks modeling.

Mots clés - gazoduc, rupture, probabilité d'inflammation, réseaux de neurones artificiels.

Keywords – gas pipeline, rupture, probability of ignition, artificial neural networks

1 INTRODUCTION

Algeria has one of the largest gas pipeline transmission networks in the world. Regulatory authorities are moving away from qualitative approaches for keeping natural gas pipelines safe. As an alternative, risk management based on a quantitative assessment (QRA) is being considered to improve the level of safety. In general, the QRA of an onshore gas transmission pipeline involves four steps (Figure.1) where estimating the probability of ignition (*POI*) turns a key parameter of the methodology

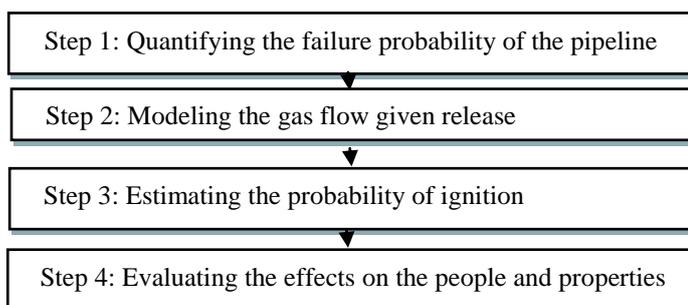


Figure 1. General Framework of QRA Analysis

In terms of human safety and property damage, consequences associated with ignited leaks of gas pipelines are in general low, but consequences of ignited ruptures can be extremely severe. However, there is a wide variation in the values of ignition probabilities given by the various data sources (Table.1).

2 PREVIOUS WORKS

Acton et Baldwin. (2008) reviewed the gas pipeline incident data obtained from the PIPESAFE Group (1970-2004) and the database administered by PHMSA of the United States Department of Transportation (2002-2007).

They investigated the relationship between *POI* and the product of the internal pressure at the time of the incident (p) and square of the outside diameter (d^2) of the pipeline. The authors grouped pipelines into different ranges of pd^2 values and calculated *POI* within a given range of pd^2 values as the proportion of the number of ignited incidents to the total number of incidents within the range (Table.2). The choice of pd^2 as input parameter is justified by the proportionality relation which links it to the initial flow rate of a leak.

Table 1. Overall Probability of Ignition of Pipeline Failures [Jones et al. 1986, Townsend et Fearnough. 1986, 7th Report of the European Gas Pipeline Incident Data Group, 2008]

Data Source	Failure Type	Probability of Ignition
U.K Gas (Townsend.1986)	Leaks:	0.1
	Ruptures:	0.5
U.S Gas (Jones. 1986)	All sizes:	0.16
	Ruptures:	0.26
European Gas (European Gas Pipeline Incident Data Group. 2008)	Pinholes/cracks:	0.04
	Holes:	0.02
	Ruptures \leq 16 in:	0.10
	Ruptures $>$ 16 in :	0.33

Table 2. Variation of Overall Ignition Probability From Acton (2008)

pd^2 range (bar m ²)	Mean [pd^2] (bar m ²)	Number of incidents	Number of ignited incidents	Ignition probability
0-5	1.3	183	11	0.06
5-15	9.8	51	11	0.22
15-30	19.7	52	16	0.31
30-45	35.7	18	10	0.56
45-80	57.0	21	17	0.81

The authors therefore proposed a simple *POI* model for rupture as follows:

$$POI = A + B pd^2 \leq POI_{max} \quad (1)$$

Where *A* and *B* are the model parameters evaluated by the least squares-based linear regression analysis and POI_{max} is an upper bound probability of ignition.

In [Lam et Zhou, 2015], Lam and Zhou proposed a log-logistic probability of ignition (*POI*) model for rupture of onshore natural gas transmission pipelines:

$$POI = \frac{1}{1 + e^{-\beta_1 - \beta_2 \ln(pd^2)}} \quad (2)$$

The parameters of the proposed *POI* model are evaluated using a maximum likelihood method based on a total of 188 rupture incidents that occurred on onshore gas transmission pipelines in the United States between 2002 and 2014 as recorded in the pipeline incident database administered PHMSA of the US Department of Transportation (table.3). The product of the pipe internal pressure at the time of rupture and outside diameter squared is also adopted as the predictor in the *POI* model and the maximum likelihood method is employed to evaluate the model parameters.

There are several drawbacks associated with the *POI* model given by Eq. (1) and Eq. (2). The results obtained by the two models for the same values of pd^2 are quite distant (Figure.2). This could be explained by the fact that the input data are from different databases and are taken over different periods of time.

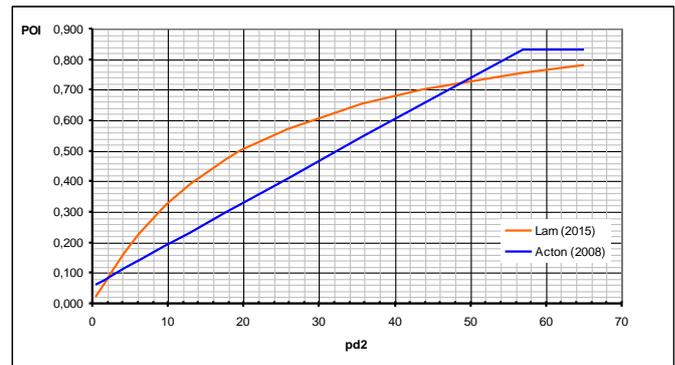


Figure 2. POI Calculated By Acton and Lam Models

3 ANN MODEL FOR POI ESTIMATION

In this paper, the objective of the study is to develop a new probability of ignition model for rupture of transmission gas pipelines which overcomes the drawbacks of existing *POI* models by using artificial intelligence and neural network modeling. Artificial neural networks (ANNs) is a modeling technique based on imitating the structure and mechanisms of the human brain and being used in more and more engineering applications where classical approaches fail or they are too complicated to be used [Anderson,1997, Nielsen, 2015]. ANNs are applied to estimate desired output parameters when enough experimental data is provided. Therefore, ANNs allow the modeling of physical phenomena without requiring explicit mathematical representations. The ability of neural networks to learn by examples makes ANN more flexible and powerful than parametric approaches.

The development of an ANN model involves three basic steps:

- compilation of available data required for training,

- training of ANN networks,
- evaluation and validation of the ANN model.

Many software programs exist on the market such as Neural Designer, GMDH Shell, DarkNet, etc. It should be noticed that these tools allow a very easy implementation of the model. In this work, a toolbox *mntool* of the software package MATLAB was employed.

3.1 Compilation of data required for training

Data issued from Lam and Zhou, (2015) and Acton and Baldwin, (2008) are grouped and used as data required for training (table.4)

3.2 Training the ANN

Following the definition of network structure and components, several ANN models were tested using the training data set described in the previous section. The performance of an ANN is affected by the characteristics of the network such as the number of hidden layers and the number of nodes in each hidden layer. By trial and error with different ANN configurations, the network selected consists of two hidden layers with 8 neurons. The input layer has one neuron that corresponded to pd^2 , while the output layer has also one neuron representing POI (figure.3). The activation function in the hidden layers was chosen as log-Sigmoid function. Consequently, the prediction of the output parameter was performed using a three-layer feed forward ANN. In addition, several parameters including the learning rule. The learning coefficient ratio, the error minimization algorithm and the number of learning cycles had to be specified.

Table 3. Observed POI from Lam and Zhou. (2015)

Group No.	pd^2 Range ($\times 10^5$ Mpa-mm ²)	Mean pd^2 for Incidents ($\times 10^5$ Mpa-mm ²)	No. of Incidents	No. of ignited Incidents	POI
1	≤ 0.8	0.5	19	1	5.3 %
2	0.8-2.3	1.6	19	0	0.0 %
3	2.3-3.2	2.8	19	3	15.8 %
4	3.2-4.9	4.3	19	2	10.5 %
5	4.9-8.1	6.2	19	6	31.6 %
6	8.1-10.6	9.6	18	5	27.8 %
7	10.6-15.5	12.9	19	8	42.1 %
8	15.5-20.1	17.8	19	8	42.1 %
9	20.1-35.8	25.8	19	11	57.9 %
10	> 35.8	43.8	18	13	72.2 %

Table 4. ANN Training Data

Observation No.	Sources	pd^2	POI
1	Lam and Zhou, (2015)	0.5	0.053
2		1.6	0
3		2.8	0.158
4		4.3	0.105
5		6.2	0.316
6		9.6	0.278
7		12.9	0.421
8		17.8	0.421
9		25.8	0.579
10		43.8	0.722
11	Acton and Baldwin, (2008)	1.3	0.06
12		9.8	0.22
13		19.7	0.31
14		35.7	0.56
15		57.0	0.81

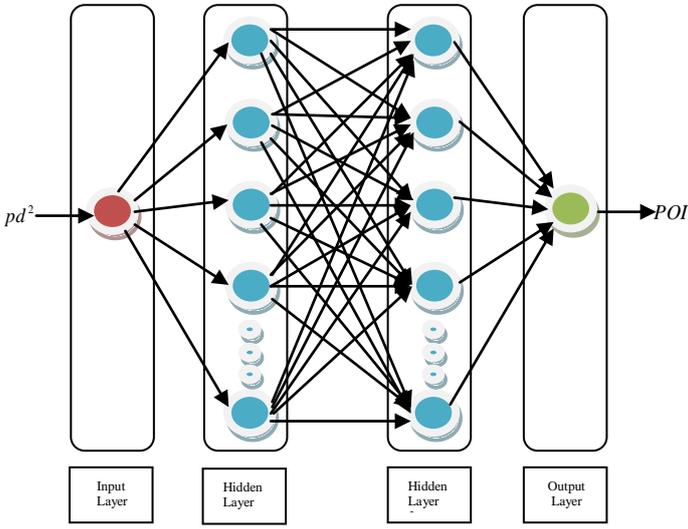


Figure 3. Structure of the ANN

3.3 Evaluation of the ANN model

Evaluation of the ANN model was based upon minimizing the difference among the neural network predicted values and the desired outputs. The performance of the ANN configurations was assessed based on calculating the mean absolute error (*MAE*).

$$MAE = \frac{1}{N} \sum_{i=1}^N \Delta y_A \quad (3)$$

The standard deviations of the absolute error (*STD_A*)

$$STD_A = \sqrt{\frac{\sum_{i=1}^N (\Delta y_A - \overline{\Delta y_A})^2}{N-1}} \quad (4)$$

Where: $\Delta y_A = |y_p - y_D|$

y_p, y_D : are respectively predicted and desired values

The mean relative error (*MRE*)

$$MRE = \frac{1}{N} \sum_{i=1}^N \Delta y_R \quad (6)$$

The standard deviations of the relative error (*STD_R*)

$$STD_R = \sqrt{\frac{\sum_{i=1}^N (\Delta y_R - \overline{\Delta y_R})^2}{N-1}} \quad (7)$$

Where:

$$\Delta y_R = \left| \frac{y_p - y_D}{y_D} \right| \quad (8)$$

The root mean square error (*RMSE*):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N \Delta y_A^2}{N}} \quad (9)$$

4 RESULTS AND DISCUSSION

Validation of the ANN model is carried out by comparison with the models of Acton and Lam using criteria described in the previous section. As seen from table.5, results of the ANN model are better for all comparison criteria (figure.6). From figs.4 and 5, it appears clearly that the Acton model underestimate *POI* where Lam model overestimate it. This fact is confirmed by figure.7.

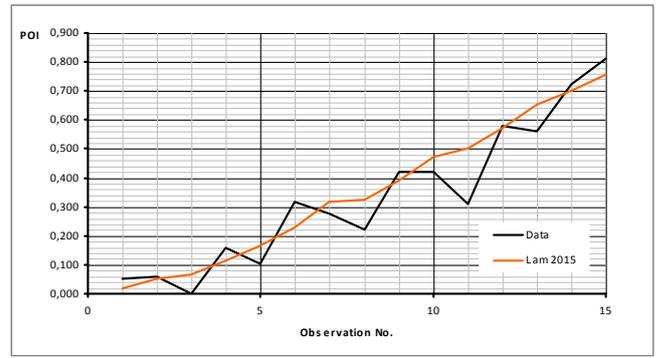


Figure 4. Observed Data and Lam (2015) Model Results

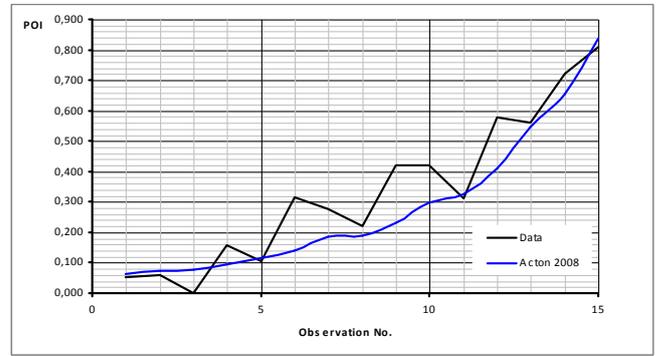


Figure 5. Observed Data and Acton (2008) Model Results

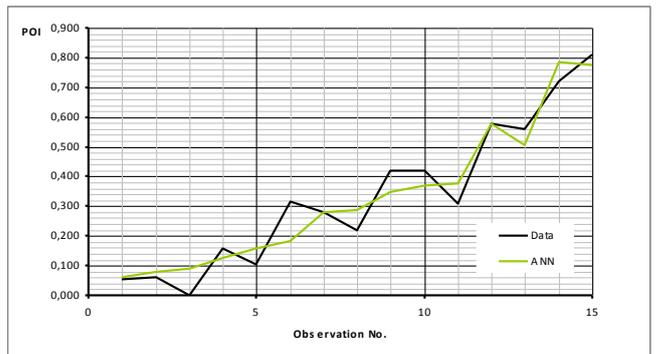


Figure 6. Observed Data and ANN Model Results

Table 5. Results of Various Models

Observation No.	Data		Models		
	pd^2	POI	ANNs	Lam	Acton
1	0.5	0.053	0.0599	0.0200	0.0624
2	1.6	0	0.0893	0.0655	0.0774
3	2.8	0.158	0.1254	0.1126	0.0939
4	4.3	0.105	0.1568	0.1666	0.1144
5	6.2	0.316	0.1824	0.2275	0.1404
6	9.6	0.278	0.2814	0.3189	0.1870
7	12.9	0.421	0.3483	0.3904	0.2322
8	17.8	0.421	0.3690	0.4739	0.2994
9	25.8	0.579	0.5783	0.5718	0.4090
10	43.8	0.722	0.7859	0.7006	0.6556
11	1.3	0.06	0.0804	0.0532	0.0733
12	9.8	0.22	0.2885	0.3237	0.1898
13	19.7	0.31	0.3772	0.5008	0.3254
14	35.7	0.56	0.5055	0.6533	0.5446
15	57.0	0.81	0.7740	0.7557	0.8364
	MAE		0.0502	0.0597	0.0716
	STDA		0.0356	0.0465	0.0646
	MRE		0.1834	0.2442	0.215
	STDR		0.1487	0.2223	0.1657
	RMSE		0.0609	0.0916	0.1164

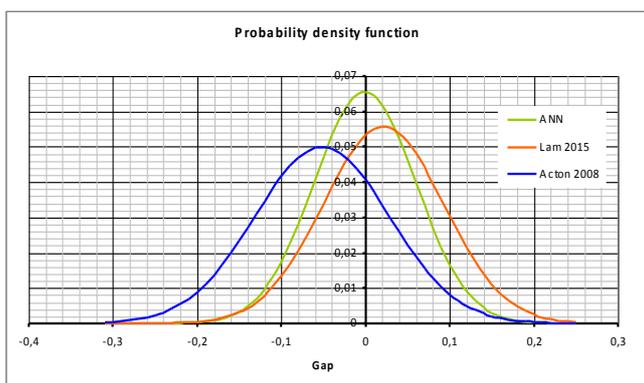


Figure 7. Statistical Gap Distribution of Different Models Compared to the Input Data

5 CONCLUSION

Estimating the probability of ignition turns a key parameter of the Quantitative risk assessment of transmission natural gas pipelines. Results obtained from available data sources and models show a wide variation in the values of ignition probabilities. The new probability of ignition model for rupture of gas pipelines developed in this work using neural network modeling overcomes the drawbacks of existing models and gives more accurate results. Availability of many software programs, where some of which are free, allows a very easy implementation of the model.

6 RÉFÉRENCES

Acton, M. R., & Baldwin, P.J. (2008). Ignition probability for high pressure gas transmission pipelines. *7th International pipeline Conference. American Society of Mechanical Engineers.* 331-339.

Jones, D.J., Kramer, G.S., Gideon, D.N., & Eiber R.J. (1986). An Analysis of Reportable Incidents for Natural Gas Transportation and Gathering lines- NG18 report N° 158. Pipeline research committee of the American Gas Association

Townsend, N.A., & Fearnough. G.D (1986). Control of Risk from UK Gas Transmission Pipelines. British Gas plc. Research and Development Division. Presented at the 7th Symposium on Line Pipe Research. American Gas Association. Houston. USA.

7th Report of the European Gas Pipeline Incident Data Group. (2008) Doc. Number EGIG 08.TV-B.0502.

Lam. C., & Zhou, W (2015). Development of Probability of Ignition Model for Ruptures of Onshore Natural Gas Transmission Pipelines. *Journal of Pressure Vessel Technology.*

Anderson, J. A. (1997). *An introduction to neural networks.* A Bradford book. The MIT Press. Cambridge. Massachusetts. London. England.

Nielsen, M.A. (2015). *Neural Networks and Deep Learning.* Determination Press.

(This paper is not included in this version, please contact authors for more information)

Stratégie intégrée de maintenance préventive conditionnelle, de production et de contrôle qualité dans les systèmes manufacturiers

NADIA BAHRIA¹, ANIS CHELBI², IMEN HARBAOUI DRIDI³, HANEN BOUCHRIHA³

¹ Université de Tunis El Manar, OASIS
École Nationale d'Ingénieurs de Tunis (ENIT)
BP 37, le Belvédère 1002 Tunis - Tunisie
École Polytechnique Internationale Privée de Tunis, PI-LAB
Rue du Lac d'Annecy, Tunis 1053 - Tunisie
nadia.bahria@enit.utm.tn

² Université de Tunis, CEREP
École Nationale Supérieure d'Ingénieurs de Tunis (ENSIT)
5 Av. Taha Hussein, 1008 Tunis – Tunisie
anis.chelbi@planet.tn

³ Université de Tunis El Manar, LACS
École Nationale d'Ingénieurs de Tunis (ENIT)
BP 37, le Belvédère 1002 Tunis - Tunisie
imen.harbaoui@issatkr.rnu.tn, hanen.bouchriha@enit.rnu.tn

Résumé – L'objectif de ce papier est de développer une approche intégrée pour le contrôle conjoint de la production, de la maintenance et de la qualité pour les systèmes de fabrication par lots. Nous considérons de tels systèmes sujets à des dégradations qui sont à l'origine de la production d'unités défectueuses. Le contrôle de la qualité des lots produits est effectué à l'aide d'une carte de contrôle '*x-bar*'. Cet outil graphique permettra d'estimer la qualité du lot produit et éventuellement d'effectuer des actions de maintenance préventive ou corrective parfaite sur le système de production. Un stock tampon est construit pour maintenir la continuité de la chaîne de production pendant les actions de maintenance. Le coût total encouru comprend le coût de réglage, les coûts d'inventaire, les coûts d'entretien et les coûts de qualité. Les variables de décision comprennent la taille du stock tampon, la taille de l'échantillon, l'intervalle d'échantillonnage, les limites de surveillance et de contrôle de la carte de contrôle. Des expériences numériques sont fournies pour évaluer l'efficacité de la politique de contrôle proposée et la robustesse de l'approche de résolution.

Abstract – The goal of this paper is to develop an integrated approach for the joint control of production, maintenance and quality for batch manufacturing systems. We consider such systems subject to degradations which are at the origin of the production of defective units. The quality control of the batches produced is carried out using an '*x-bar*' control chart. This graphic tool will make it possible to estimate the quality of the produced batch and possibly to carry out perfect preventive or corrective maintenance actions on the production system. A buffer stock is built to maintain the continuity of the production chain during maintenance actions. The total cost incurred includes the cost of réglage, inventory costs, maintenance costs and quality costs. Decision variables include buffer stock size, sample size, sampling interval, monitoring and control chart limits. Numerical experiments are provided to assess the effectiveness of the proposed control policy and the robustness of the resolution approach.

Mots clés – Stratégie intégrée, carte de contrôle, optimisation, stock tampon, maintenance préventive et corrective.

Keywords – Integrated strategy, control chart, optimization, preventive and corrective maintenance.

1 INTRODUCTION

La performance des systèmes de fabrication dépend en bonne partie de l'intégration des politiques de maintenance, de production et de contrôle qualité. Toutefois, ces trois fonctions sont souvent gérées séparément dans la pratique bien qu'elles soient étroitement liées. En effet, il est difficile d'entreprendre des actions de maintenance sans la considération de l'activité de production. La révision du plan de production en cas

d'urgence peut engendrer des coûts excessifs. Ainsi, la mise en place des stratégies de maintenance préventive (MP) devient cruciale afin de réduire la probabilité de défaillance et les pertes de production correspondantes. Par ailleurs, un équipement mal entretenu peut produire des produits défectueux, ce qui entraîne des coûts de réparation et de rebut considérables. Dans d'autres cas, les fonctions production et qualité sont intégrées sans étudier l'aspect maintenance. Dans ce contexte, plusieurs auteurs ont travaillé à la définition et à

l'optimisation de politiques intégrées impliquant simultanément paramètres de production, de maintenance et de qualité. Par exemple, [Ben Daya, 2002] a traité un modèle intégré pour l'articulation détermination de l'EPQ et du niveau de PM périodique pour un système qui passe à un état «hors contrôle» après un temps aléatoire. Ils illustrent l'effet des PM sur le coût de non-qualité. [Chelbi et al., 2008] et [Dhouib et al., 2012] ont proposé une approche intégrée de contrôle de la production-inventaire et d'une politique de PM basée sur l'âge. Le système peut passer à l'état «hors-contrôle» après une période de production aléatoire, et commencer à produire des articles non conformes. Dans le cas où cela se produit, une action de maintenance corrective est prévue pour restaurer la machine à l'état neuf. Le modèle développé permet de déterminer la taille optimale du stock tampon et l'âge auquel le PM doit être effectué. [Radhoui et al., 2010a] ont proposé un modèle intégré de qualité PM pour un système de production imparfait.

Deux variables de décision caractérisent ce modèle: le taux d'unités défectueuses pour lesquelles les actions MP doivent être entreprises et la taille du stock tampon. [Radhoui et al., 2010b] ont proposé une autre politique conjointe de contrôle de la qualité, de PM et de production. En fonction du taux d'unités non conformes détectées, il est décidé s'il faut ou non entreprendre des actions de maintenance. Les variables de décision comprennent le taux optimal d'unités défectueuses sur la base duquel l'action de maintenance préventive doit être entreprise, la taille optimale du stock tampon et le niveau seuil du taux de rejet qui correspond à l'instant auquel le stock tampon doit être constitué. [Chakraborty et Giri, 2012] considèrent un processus de production, qui passe aléatoirement à l'état «hors de contrôle». Les actions PM sont effectuées à la fin de chaque cycle de production et l'action CM est entreprise après chaque panne. Les auteurs ont proposé un algorithme de calcul afin de déterminer les valeurs optimales du stock de sécurité et le temps d'exécution de la production. [Pandey et al., 2011] ont proposé un modèle mathématique afin d'optimiser la période PM et les paramètres de la carte de contrôle. De plus, l'intervalle PM optimal obtenu est intégré avec le programme de production afin de déterminer la séquence optimale des lots. [Suliman et Jawad, 2012] ont proposé un modèle pour optimiser l'âge auquel la MP doit être effectuée et la taille du lot de production pour un système de production qui peut passer à un état «hors-contrôle» entraînant la production d'articles non conformes. [Chen, 2013] considéré comme imparfait système de production avec processus de reprise imparfait et PM imparfait. Les auteurs déterminent les valeurs optimales de l'EPQ, le nombre d'inspections, les périodes d'inspection et la période PM.

[Bouslah et al., 2016] ont intégré la production, la maintenance et le contrôle de la qualité à l'aide d'un plan d'échantillonnage. Les variables de décision considérées sont la taille du lot de production, le stock de sécurité, le plan d'échantillonnage et la planification des révisions avec une contrainte de qualité sortante. [Bahria et al., 2018] ont optimisé la taille de l'échantillon, l'intervalle d'échantillonnage et les limites de la carte de contrôle minimisant le coût total moyen par unité de temps dans un système de production imparfait. [Salmasnia et al., 2019] ont proposé un modèle intégré pour la détermination conjointe de la longueur de la production, des paramètres de la carte de contrôle adaptative et de la politique de maintenance. Lorsque le processus est sous contrôle, la MP aura lieu après k intervalles d'inspection. L'action CM est exécutée après le passage à l'état «hors contrôle». [Rivera-

Gomez al., 2020] ont développé une politique intégrée de production, de maintenance et de contrôle qualité pour un système de fabrication monobloc sujet à détérioration. Un plan d'échantillonnage est mis en place pour garantir une certaine moyenne de limite de qualité sortante.

En fonction de la proportion de défectueux détectés lors de l'inspection, le décideur peut décider de lancer immédiatement la MP. En cas de panne, une réparation minimale est effectuée pour restaurer la machine dans un état aussi mauvais que l'ancien. L'objectif de cette étude est de déterminer conjointement le taux de production, la fraction de production inspectée et le taux d'entretien majeur qui minimisent le coût total encouru.

Pour des raisons de simplification, les modèles traités dans la littérature sont basés sur certaines hypothèses simplificatrices et irréalistes pour modéliser la dégradation de la qualité des produits. Ces modèles négligent en effet souvent la corrélation entre les dégradations de la qualité et de la fiabilité. Par ailleurs, les cartes de contrôle représentent un outil important de la Maîtrise Statistique des Processus. Cet outil permet d'éviter le coût excessif du contrôle à 100% et de suivre dans le temps la qualité des produits. Toutefois, bien que les cartes de contrôle soient largement utilisées dans l'industrie, leur intégration avec les politiques de production et de maintenance n'a pas été traitée encore dans la littérature.

Le présent travail fournit une nouvelle stratégie qui traite simultanément les paramètres décisionnels reliés à la fois à la production (taille des stocks tampons et taille des lots de production), à la maintenance (les instants pour effectuer la maintenance préventive et corrective), et à la qualité (les paramètres optimaux de la carte de contrôle). Ce travail explore la dynamique et les propriétés statistiques des cartes de contrôle afin de montrer comment on peut profiter des informations fournies par ces cartes pour surveiller la qualité des produits et planifier des interventions de maintenance tout en assurant la continuité de la chaîne de production.

Nous proposons ainsi une nouvelle stratégie intégrant la production, la maintenance et la qualité pour les systèmes de production imparfaits et pouvant produire des unités défectueuses. Le contrôle de la qualité est effectué à l'aide d'une carte de contrôle ' x -bar'. Selon les résultats de la carte de contrôle, on décide d'entreprendre ou non une action de maintenance pouvant être préventive ou corrective. Cette stratégie est caractérisée par cinq variables de décision : la taille de l'échantillon, l'intervalle d'échantillonnage, les limites de contrôle de la carte de contrôle, la taille du stock tampon réduisant le risque de pénurie lors des périodes d'arrêt de la machine, et la période de maintenance préventive.

Le reste de ce papier est organisé comme suit : la section 2 présente la politique intégrée proposée, les hypothèses de travail considérées et les notations que nous allons utiliser pour modéliser le problème étudié. Dans la section 3, nous présentons le modèle mathématique développé exprimant le coût total moyen par unité de temps. La section 4 est dédiée à un exemple numérique. Le papier se termine par une conclusion dans la section 5.

2 STRATEGIE INTEGREE, HYPOTHESES DE TRAVAIL ET NOTATIONS

2.1 Stratégie intégrée

Nous considérons un système de fabrication composé d'une machine produisant un seul produit par lots pour satisfaire une demande constante et continue au taux D . Un cycle de production commence par un système à l'état neuf (As Good As New), est supposé à l'état 'sous contrôle' produisant des

articles conformes. Initialement, l'unité de production peut fonctionner à la capacité maximale au débit P supérieur à la demande D . Chaque fois qu'un niveau de stock maximum h est atteint, le taux de production est égal à D . Ce stock tampon est construit pour satisfaire la demande pendant les interruptions de service pour les actions de maintenance préventive ou corrective. Le processus est surveillé en prenant des échantillons de taille n toutes les unités de temps I_s pour vérifier l'état du processus. Les résultats d'échantillonnage sont enregistrés à l'aide d'une carte de contrôle 'x-bar'. L'indicateur de qualité x suit une distribution Normale de moyenne μ et d'écart-type σ .

Les paramètres de la carte de contrôle sont :

(I_s) l'intervalle d'échantillonnage (en unités de temps), (n) la taille de l'échantillon, (k_p) le nombre d'écart-types entre la ligne centrale de la carte de contrôle et les limites de surveillance, et (k_c) le nombre d'écart-types de la distribution d'échantillon entre la ligne centrale de la carte de contrôle et les limites de contrôle. Les variables de décision liées à la carte de contrôle sont n , I_s , k_p et k_c .

Selon l'évolution de l'indicateur de qualité, on décide d'entreprendre ou non des actions de maintenance et quel type de maintenance à effectuer. Comme le montre la Figure 1, si la moyenne de l'échantillon de l'indicateur de qualité est trouvée entre les limites de surveillance et de contrôle ($USL < \bar{x}_e < UCL$ ou $LCL < \bar{x}_e < LSL$) pour un échantillon, le processus est considéré comme étant passé à un état hors de contrôle que nous appellerons un état de 'surveillance'. Par conséquent, la machine est arrêtée et une action de MP de durée moyenne μ_p est effectuée. Dans ce cas, une proportion d'unités non conformes est détectée et rejetée. Nous supposons que le nombre d'articles défectueux est proportionnel à la quantité produite le ($j_p - 1$)^{ème} et le j_p ^{ème} intervalle d'échantillonnage. Le nombre d'inspections m est donc égal à j_p .

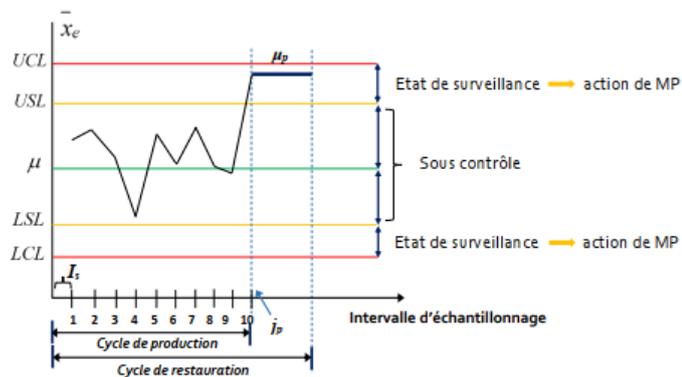


Figure 1. Cas d'une action de MP déclenchée par la carte de contrôle

Si la moyenne d'échantillon de l'indicateur de qualité pour un échantillon est supérieure à la limite de contrôle supérieure UCL ou inférieure à la limite de contrôle inférieure LCL (Figure 2), la machine est considérée dans un état hors de contrôle appelé 'état critique'. Par conséquent, une action de MC de durée moyenne μ_c est planifiée et tous les éléments

produits entre les intervalles d'échantillonnage ($j_c - 1$)^{ème} et j_c ^{ème} sont rejetés. Le nombre d'inspections est donc égal à j_c .

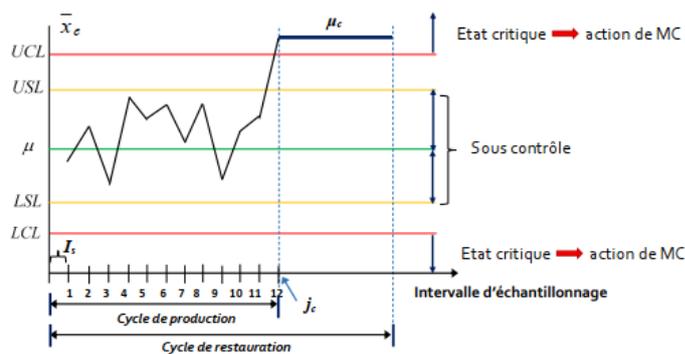


Figure 2. Cas d'une action de MC déclenchée par la carte de contrôle

Tant que la moyenne de l'échantillon est située entre les limites de surveillance, le processus est à l'état 'sous contrôle'. Dans ce cas, la production continue et aucune action de maintenance n'est entreprise.

Ainsi, la stratégie intégrée proposée est résumée dans la Figure 3 qui illustre les trois états possibles du processus générés par la carte de contrôle et sur lesquels les décisions concernant la maintenance et le rejet des articles défectueux sont prises.

L'objectif de cette étude est de minimiser le coût total attendu par unité de temps en trouvant les valeurs optimales du stock tampon " h ", de la taille de l'échantillon " n ", de l'intervalle d'échantillonnage " I_s ", du nombre d'écart-types entre la ligne centrale de la carte de contrôle et les limites de surveillance " k_p ", et du nombre d'écart-types entre la ligne centrale et les limites de contrôle " k_c ".

Un modèle de coût sera élaboré en tenant compte des coûts associés aux éléments suivants : le réglage, la détention du stock et la pénurie, la maintenance préventive et corrective, les fausses alarmes, l'échantillonnage et la non-conformité des produits. Les expressions de ces éléments de coûts seront détaillées dans la Section 3.

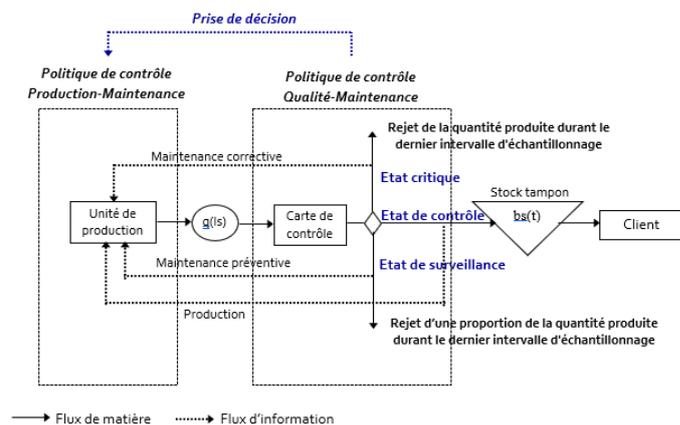


Figure 3. La politique intégrée de production, de maintenance et de contrôle qualité proposée pour une unité de production imparfaite

2.2 Hypothèses

Les hypothèses suivantes sont considérées :

- (H1) La fonction de distribution cumulative associée aux durées de vie de l'unité de production est inconnue.
- (H2) Les demandes qui ne peuvent être satisfaites sont perdues.

- (H3) Les articles produits sont périssables.
 (H4) L'indicateur qualité contrôlé est distribué suivant la loi Normale.
 (H5) Les éléments non conformes ne sont pas réintégrés dans le processus de fabrication (pas de retouche).
 (H6) La non-conformité des produits est due uniquement à la dégradation de la machine.
 (H7) Les actions de maintenance restaurent l'état de l'unité de production « As Good As New ».
 (H8) L'unité de production ne tombe jamais en panne avant la prochaine action de maintenance.
 (H9) Le stock inutilisé est épuisé à zéro après chaque cycle de restauration. La période d'épuisement est considérée comme négligeable.
 (H10) Les ressources nécessaires pour effectuer des actions de maintenance sont disponibles.
 (H11) Les coûts unitaires liés à l'inventaire, à la maintenance et à la qualité sont connus et constants.

2.3 Notations

Q	Taille du lot de production
$q(t)$	Niveau du lot à l'instant t
$bs(t)$	Niveau du stock à l'instant t
h	Taille du stock tampon
P	Taux de production maximal
D	Taux de demande
x_i	Mesure de l'indicateur de qualité à contrôler de la pièce i prélevée
\bar{x}_e	Moyenne des mesures de l'indicateur de qualité pour l'échantillon e
n	Taille de l'échantillon
μ	Moyenne du processus
μ_1	La moyenne qui correspond à la variabilité du processus suite au passage à l'état 'hors-contrôle'
σ	L'écart type du processus
\mathcal{D}	Amplitude moyenne de la dérive de l'indicateur de qualité en cas de dépassement des limites de contrôle
I_s	Intervalle d'échantillonnage
$q(I_s)$	Niveau du lot à l'instant I_s
m	Nombre d'échantillons contrôlés lors de la production d'un lot
UCL	Limite de contrôle supérieure
LCL	Limite de contrôle inférieure
k	Nombre d'écart-types de la distribution de l'échantillon entre la ligne centrale de la carte de contrôle et les limites de contrôle
J	Nombre moyen d'échantillons pour détecter l'état 'hors-contrôle'
$F(.)$	Fonction de répartition de la loi Normale centrée réduite
μ_c	Durée moyenne d'une action de maintenance corrective
p_{sc}	Probabilité que le système soit dans l'état 'sous-contrôle'
p_{hc}	Probabilité que le système soit dans l'état 'hors-contrôle'
β	Proportion d'articles défectueux produits durant le dernier intervalle d'échantillonnage avant le passage à l'état de surveillance
C_s	Coût de détention d'un item par unité de temps

C_p	Coût de pénurie d'une pièce perdue
C_i	Coût de l'inspection d'un produit
C_r	Coût d'un produit défectueux
C_{MP}	Coût d'une action de MP
C_{MC}	Coût d'une action de MC
Γ_{reg}	Coût moyen de réglage
Γ_{S1}	Coût moyen d'inventaire pour le scénario où le stock tampon est atteint avant le passage à l'état hors contrôle de la machine (scénario 1)
Γ_{S2}	Coût moyen d'inventaire pour le scénario où le stock tampon n'a pas été complètement bâti avant le passage à l'état hors contrôle (scénario 2)
Γ_s	Coût total moyen d'inventaire
Γ_M	Coût total moyen de maintenance
Γ_Q	Coût total moyen de qualité
Γ_{ech}	Coût total moyen d'échantillonnage
Γ_{NC}	Coût total moyen de non-conformité des produits
Γ_{FA}	Coût total moyen de fausses alarmes
Γ_{tot}	Coût total moyen par unité de temps

3 MODELE MATHEMATIQUE

Le modèle de calcul de coût devrait prendre en compte tous les scénarios possibles. Deux scénarios sont possibles selon que le niveau du stock tampon est atteint ou non avant le passage à l'état de surveillance ou à l'état critique. Le premier scénario se produit lorsque le stock tampon h est complété pendant le cycle. La condition pour être dans le scénario 1 est :

- Dans le cas de l'état de surveillance (action de MP)

$$j_p \times I_s \geq \frac{h}{P - D}$$

- Dans le cas de l'état critique (action de MC)

$$j_c \times I_s \geq \frac{h}{P - D}$$

Le deuxième scénario se produit lorsque le niveau de stock de tampon h n'a pas été complété par la production suite au passage à un état hors contrôle du processus (dépassement des limites de surveillance ou de contrôle). Dans ce cas, un certain niveau h' ($h' < h$) est construit. La condition pour être dans le scénario 2 est :

- Dans le cas de l'état de surveillance (action de MP)

$$j_p \times I_s < \frac{h}{P - D}$$

- Dans le cas de l'état critique (action de MC)

$$j_c \times I_s < \frac{h}{P - D}$$

Deux possibilités seront envisagées plus tard pour chacun de ces deux scénarios selon que la pénurie se produit ou non.

Notre objectif est de déterminer les paramètres optimaux de la carte de contrôle qui sont : la taille de l'échantillon n , l'intervalle d'échantillonnage I_s , ainsi que le coefficient des limites de contrôle, k . D'autre part, nous proposons d'optimiser la taille du stock tampon h et l'indice de l'échantillon α auquel l'action de MP doit être entreprise. L'objectif est de minimiser le coût total moyen par unité de temps pendant un cycle de restauration. Le coût total moyen correspond à la somme du coût de réglage, du coût total moyen de l'inventaire, du coût total moyen de maintenance et du coût total moyen de la qualité, y compris les coûts d'échantillonnage, de fausses

alarmes et de rejet. La politique de production consiste à constituer au début de chaque cycle de production un stock tampon de taille h à un taux $(P-D)$. Le cycle de production se termine lorsque les limites de contrôle de la carte de contrôle sont dépassées. Par conséquent, le cycle de production se termine après $(J \times I_s)$ unités de temps, où J est le nombre moyen d'échantillons pour détecter l'état 'hors contrôle'.

3.1 Durée moyenne du cycle de restauration

Soit j_p le nombre moyen d'échantillons pour détecter le passage à l'état de surveillance. Il est exprimé comme suit [Montgomery, 2004].

$$j_p = \frac{1}{1 - \alpha_1} \quad (1)$$

α_1 est la probabilité de non-détection de l'état de surveillance. Dans ce cas, la moyenne de l'indicateur de qualité se situe dans une fourchette acceptable entre les limites de surveillance ($LSL \leq \bar{x}_e \leq USL$), alors que le processus est dans un état de surveillance et que la machine doit être soumise à une action de MP.

$$\alpha_1 = \left\{ \text{Prob}(LSL \leq \bar{x}_e \leq USL) \mid \mu_1 = \mu + \delta_p \times \sigma \right\} \quad (2)$$

$$\alpha_1 = F\left(\frac{USL - \mu_1}{\sigma/\sqrt{n}}\right) - F\left(\frac{LSL - \mu_1}{\sigma/\sqrt{n}}\right) \quad (3)$$

Ainsi :

$$\alpha_1 = F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) \quad (4)$$

En utilisant les équations 1 et 4, nous obtenons :

$$j_p = \frac{1}{1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n})} \quad (5)$$

Soit j_c le nombre moyen d'échantillons pour détecter le passage à l'état critique. Il est exprimé comme suit [Montgomery, 2004].

$$j_c = \frac{1}{1 - \alpha_2} \quad (6)$$

α_2 étant la probabilité de non-détection de l'état critique. Dans ce cas, la moyenne de l'indicateur de qualité ne dépasse pas les limites de contrôle ($LCL \leq \bar{x}_e \leq UCL$) alors que le processus est dans l'état critique et qu'une action de MC est nécessaire.

$$\alpha_2 = \left\{ \text{Prob}(LCL \leq \bar{x}_e \leq UCL) \mid \mu_2 = \mu + \delta_c \times \sigma \right\} \quad (7)$$

$$\alpha_2 = F\left(\frac{UCL - \mu_2}{\sigma/\sqrt{n}}\right) - F\left(\frac{LCL - \mu_2}{\sigma/\sqrt{n}}\right) \quad (8)$$

$$\alpha_2 = F(k_c - \delta_c \times \sqrt{n}) - F(-k_c - \delta_c \times \sqrt{n}) \quad (9)$$

Et à partir des équations 6 et 9, nous avons :

$$j_c = \frac{1}{1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n})} \quad (10)$$

La durée moyenne du cycle de restauration, RCD , est donnée par:

$$RCD = \left\{ \begin{array}{l} j_p \times \Psi_1 + j_c \times (1 - \Psi_1) \times I_s \\ + \mu_p \times \Psi_1 + \mu_c \times (1 - \Psi_1) \end{array} \right\} \quad (11)$$

ψ_1 est un paramètre qui indique si les limites de surveillance ont été dépassées avant les limites de contrôle. Il est égal à 1 si on va entreprendre une action de MP, c'est-à-dire si les limites de surveillance sont dépassées. Dans le cas de dépassement des limites de contrôle, cet indicateur est égal à zéro et une action de MC est envisagée. Cet indicateur est exprimé par :

$$\psi_1 = \text{indicateur}(j_p < j_c) = \begin{cases} 1 & \text{si } j_p < j_c \\ 0 & \text{sinon} \end{cases} \quad (12)$$

Dans la section suivante, nous développerons l'expression du coût moyen de l'inventaire en considérant tous les scénarios possibles et en utilisant les valeurs de j_p et j_c exprimées par les équations (5) et (10).

3.2 Probabilités des trois scénarios

- $P_{\text{sous-contr}}$ est la probabilité que le système soit dans l'état 'sous contrôle'. Elle est formulée comme suit :

$$P_{\text{sous-contr}} = \text{Prob}(LSL \leq \bar{x}_e \leq USL) \quad (13)$$

$$P_{\text{sous-contr}} = F\left(\mu + k_p \times \frac{\sigma}{\sqrt{n}}\right) - F\left(\mu - k_p \times \frac{\sigma}{\sqrt{n}}\right) \quad (14)$$

- p_{mp} est la probabilité qu'une action de MP soit effectuée.

Ce paramètre est décrit par la fonction suivante :

$$P_{mp} = \left\{ \begin{array}{l} \text{Prob}(USL \leq \bar{x}_e \leq UCL) \\ + \text{Prob}(LCL \leq \bar{x}_e \leq LSL) \end{array} \right\} \quad (15)$$

$$P_{mp} = \left\{ \begin{array}{l} F\left(\mu + k_c \times \frac{\sigma}{\sqrt{n}}\right) - F\left(\mu + k_p \times \frac{\sigma}{\sqrt{n}}\right) \\ + F\left(\mu - k_p \times \frac{\sigma}{\sqrt{n}}\right) - F\left(\mu - k_c \times \frac{\sigma}{\sqrt{n}}\right) \end{array} \right\} \quad (16)$$

- p_{mc} est la probabilité qu'une action de MC soit effectuée.

Elle est donnée par :

$$p_{cm} = \text{Prob}(\bar{x}_e \leq LCL) + \text{Prob}(\bar{x}_e \geq UCL) \quad (17)$$

$$p_{mc} = F\left(\mu - k_c \times \frac{\sigma}{\sqrt{n}}\right) + 1 - F\left(\mu + k_c \times \frac{\sigma}{\sqrt{n}}\right) \quad (18)$$

3.3 Coût moyen d'inventaire

La politique de production consiste à produire un lot de taille Q au taux de production maximal P , et à constituer, au début de chaque cycle de production, un stock tampon de taille h à un taux $(P-D)$ afin d'assurer la satisfaction continue de la demande pendant les actions de maintenance préventive ou corrective. Le cycle de production se termine lorsque les limites de surveillance ou de contrôle de la carte de contrôle sont dépassées. C'est-à-dire que le cycle de production se termine après $j_p \times I_s$ unités de temps au cas où le processus passe à l'état de surveillance, et il se termine après $j_c \times I_s$ si le processus passe à l'état critique.

Le coût de l'inventaire est divisé en deux termes : le coût lié au premier scénario sans pénurie et celui relatif au deuxième scénario avec rupture de stock. Pour chaque scénario, il est possible d'effectuer une action de MP ($\psi_1 = 1$) ou une action de MC ($\psi_1 = 0$).

Nous développons ci-dessous les expressions des coûts totaux moyens d'inventaire Γ_{S1} et Γ_{S2} pour le premier et le deuxième scénario respectivement.

Dans ce scénario, deux situations sont possibles : la pénurie se produit ou ne se produit pas.

▪ **Scénario 1 : Cas sans pénurie**

La Figure IV.4 montre l'évolution du stock pour le cas sans pénurie à l'état de surveillance (une action de MP est effectuée) et à l'état critique (une action de MC est entreprise). C'est le cas lorsque les périodes d'interruption de service μ_p ou μ_c ne dépassent pas la période de consommation du stock tampon.

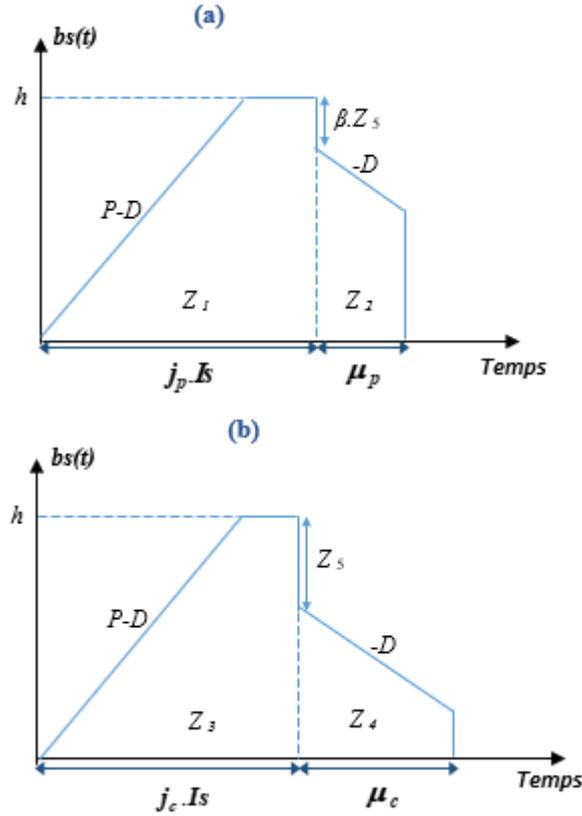


Figure 4. Evolution du niveau du stock : cas sans pénurie (action de MP (a) et action de MC (b))

Soit Z_i ($i = 1, 2, 3, 4$) l'inventaire moyen retenu dans chaque période, et la quantité produite durant le dernier intervalle d'échantillonnage (Is) avant de dépasser les limites de surveillance ou de contrôle.

Le coût moyen de possession du stock correspondant au cas sans pénurie dans ce scénario est donné par :

$$\Gamma_{S1} = C_S \times \left\{ \begin{array}{l} p_{sous-contr} \times (Z_1 \times \psi_1 + Z_3 \times (1 - \psi_1)) \\ + p_{mp} \times (Z_2 \times \psi_1) + p_{mc} \times (Z_4 \times (1 - \psi_1)) \end{array} \right\} \quad (19)$$

Avec :

$$Z_1 = \frac{h^2}{2 \times (P - D)} + h \times \left(j_p \times Is - \frac{h}{(P - D)} \right) \quad (20)$$

$$Z_2 = \mu_p \times (h - \beta \times Z_5) - \frac{\mu_p^2 \times D}{2} \quad (21)$$

$$Z_3 = \frac{h^2}{2 \times (P - D)} + h \times \left(j_c \times Is - \frac{h}{(P - D)} \right) \quad (22)$$

$$Z_4 = \mu_c \times (h - Z_5) - \frac{\mu_c^2 \times D}{2} \quad (23)$$

$$Z_5 = D \times Is \quad (24)$$

Ainsi :

$$\Gamma_{S1} = C_S \times \left\{ \begin{array}{l} p_{sous-contr} \times \left[\begin{array}{l} \left(\frac{h^2}{2 \times (P - D)} + h \times \left(j_p \times Is - \frac{h}{(P - D)} \right) \right) \times \psi_1 \\ + \left(\frac{h^2}{2 \times (P - D)} + h \times \left(j_c \times Is - \frac{h}{(P - D)} \right) \right) \times (1 - \psi_1) \end{array} \right] \\ + p_{mp} \times \left(\frac{\mu_p \times (h - \beta \times D \times Is)}{-\frac{\mu_p^2 \times D}{2}} \right) \times \psi_1 \\ + p_{mc} \times \left(\frac{\mu_c \times (h - D \times Is)}{-\frac{\mu_c^2 \times D}{2}} \right) \times (1 - \psi_1) \end{array} \right\} \quad (25)$$

▪ **Scénario 2 : Cas avec pénurie**

Pour un cycle de restauration avec pénurie, les périodes d'interruption de service dépassent la période de consommation du stock de tampon. Dans ce cas, le coût d'inventaire moyen inclut le coût de possession de l'inventaire et le coût de la pénurie. La Figure 5 présente l'évolution du stock pour l'état de surveillance (action de MP) et l'état critique (action de MC).

Nous utilisons la même approche que pour le cas sans pénurie dans la sous-section précédente. Pour un cycle de restauration avec pénurie, l'expression du coût moyen d'inventaire dans le cas avec pénurie est obtenue en additionnant le coût de stockage et le coût de rupture.

$$Z_6 = Z_1 \quad (26)$$

$$Z_7 = \frac{(h - \beta \times D \times Is)^2}{2 \times D} \quad (27)$$

$$Z_8 = \left(\mu_p - \frac{(h - \beta \times D \times Is)}{D} \right) \times D \quad (28)$$

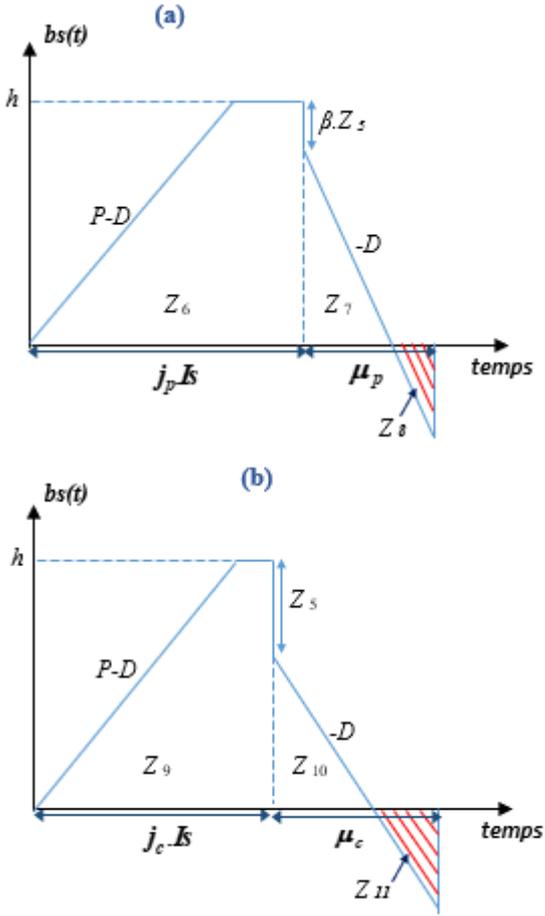


Figure 5. Evolution du niveau du stock : cas avec pénurie (action de MP (a) et action de MC (b))

$$Z_9 = Z_3 \quad (29)$$

$$Z_{10} = \frac{(h - D \times Is)^2}{2 \times D} \quad (30)$$

$$Z_{11} = \left(\mu_c - \frac{(h - Z_5)}{D} \right) \times D \quad (31)$$

Et :

$$\Gamma_{S2} = \left\{ \begin{array}{l} C_S \times \left(\begin{array}{l} p_{\text{sous-contr}} \times \left(\begin{array}{l} Z_6 \times \psi_1 \\ + Z_9 \times (1 - \psi_1) \end{array} \right) \\ + p_{mp} \times (Z_7 \times \psi_1) \\ + p_{mc} \times (Z_{10} \times (1 - \psi_1)) \end{array} \right) \\ + C_p \times \left(\begin{array}{l} p_{mp} \times (Z_8 \times \psi_1) \\ + p_{mc} \times (Z_{11} \times (1 - \psi_1)) \end{array} \right) \end{array} \right\} \quad (32)$$

Par la suite, nous obtenons l'expression du coût d'inventaire relatif au deuxième scénario formulée dans l'équation 33.

$$\Gamma_{S2} = \left\{ \begin{array}{l} C_S \times \left(\begin{array}{l} p_{\text{sous-contr}} \times \left(\begin{array}{l} \left(\frac{h^2}{2 \times (P-D)} \right) \times \psi_1 \\ + h \times \left(j_p \cdot Is - \frac{h}{(P-D)} \right) \end{array} \right) \\ + \left(\begin{array}{l} \left(\frac{h^2}{2 \times (P-D)} \right) \times (1 - \psi_1) \\ + h \times \left(j_c \cdot Is - \frac{h}{(P-D)} \right) \end{array} \right) \end{array} \right) \\ + p_{mp} \times \frac{(h - \beta \times D \times Is)^2}{2 \times D} \times \psi_1 + p_{mc} \times \frac{(h - D \times Is)^2}{2 \times D} \times (1 - \psi_1) \end{array} \right\} \\ + C_p \times \left(\begin{array}{l} p_{mp} \times (\mu_p \times D - (h - \beta \times D \times Is)) \times \psi_1 \\ + p_{mc} \times (\mu_c \times D - (h - D \times Is)) \times (1 - \psi_1) \end{array} \right) \quad (33)$$

$$\varphi_1 = \text{indicateur}(\mu_c > \frac{h - D \times Is}{D}) = \begin{cases} 1 & \text{si } \mu_c > \frac{h - D \times Is}{D} \\ 0 & \text{sinon} \end{cases}$$

pour $\psi_1 = 0$ (34)

Le coût total moyen de l'inventaire Γ_S tenant compte des deux cas (avec et sans pénurie) est exprimé par :

$$\Gamma_S = \Gamma_{S1} \times (1 - \varphi_1) + \Gamma_{S2} \times \varphi_1 \quad (35)$$

La taille du lot de production Q présente la quantité produite pendant le temps de production réel pendant un cycle de restauration. Il est exprimé en fonction du niveau de stock maximum h et donné par :

$$Q = P \times \frac{h}{P - D} + D \times \left(\begin{array}{l} \left(\begin{array}{l} j_p \times \psi_1 \\ + j_c \times (1 - \psi_1) \end{array} \right) \times Is \\ - \frac{h}{P - D} \end{array} \right) \quad (36)$$

3.3 Coût total moyen de maintenance

Compte tenu du fait que dans chaque cycle, une action de maintenance préventive ou corrective peut être effectuée, le coût d'entretien total moyen est exprimé comme suit :

$$\Gamma_M = C_{MP} \times \psi_1 + C_{MC} \times (1 - \psi_1) \quad (37)$$

3.4 Coût total moyen de qualité

Le coût moyen total de la qualité correspond à la somme du coût des fausses alarmes, du coût d'échantillonnage et du coût de non-conformité des produits pendant la période de production d'un lot.

$$\Gamma_Q = \Gamma_{FA} + \Gamma_{éch} + \Gamma_{NC} \quad (38)$$

3.4.1 Coût moyen de fausses alarmes

La notion de fausse alarme est introduite dans le cas de détection d'un dépassement des limites de la carte bien que le processus soit sous contrôle. Dans ce cas, un coût supplémentaire appelé coût d'une fausse alarme est intégré dans le modèle. Il est donné par [Bahria et al., 2018] :

$$\Gamma_{FA} = 2 \times C_F \times \left(\sum_{i=0}^m i \times (p_{sc})^i \times (1 - p_{sc}) \right) \times F(-k) \quad (39)$$

3.4.2 Coût moyen d'échantillonnage

Le coût total moyen de l'échantillonnage et de l'inspection est exprimé par :

$$\Gamma_{éch} = Ci \times n \times m \quad (40)$$

3.4.3 Coût moyen de non-conformité des produits

En référence à la Figure 4, la proportion d'unités non conformes rejetées du fait du passage à l'état critique est égale à la demande au cours du dernier intervalle d'échantillonnage I_5 ($Z_5 = D \times I_5$). Cependant, le nombre d'articles défectueux rejetés lors du passage à l'état de surveillance est proportionnel à la quantité produite durant le dernier intervalle d'échantillonnage avant le passage à cette condition ($\beta \times Z_5 = \beta \times D \times I_5$). Ainsi, l'expression du coût des items non conformes pour ce scénario est :

$$\Gamma_{NCI} = C_r \times ((\beta \times D \times I_5) \times \psi_1 + (D \times I_5) \times (1 - \psi_1)) \quad (41)$$

3.4.4 Coût total moyen par unité du temps

L'expression du coût total moyen par unité de temps est donnée comme suit :

$$\Gamma_{tot} = \frac{\Gamma_{rég} + \Gamma_{stoc} + \Gamma_M + \Gamma_Q}{RCD} \quad (42)$$

4 EXEMPLE NUMERIQUE

Notre objectif est de minimiser le coût total attendu par unité de temps pour les deux scénarios 1 et 2 dans les deux cas d'absence et de pénurie (Equation 42). Afin de résoudre le modèle proposé, et en raison de la complexité du problème, nous avons développé une procédure numérique pour déterminer les valeurs optimales du stock tampon, la taille de l'échantillon, l'intervalle d'échantillonnage, les limites de surveillance et de contrôle qui correspondent à la politique optimale. En utilisant un algorithme d'optimisation développé, les meilleures stratégies d'inventaire, de qualité et de maintenance sont obtenues.

Dans cette partie, nous discutons d'un exemple décrivant un système de fabrication par lots. Les valeurs des paramètres de l'exemple sont choisies arbitrairement en considérant des paramètres réalistes : $\mu = 5$, $\sigma = 1.5$, $P = 30$ unités/heure, $D = 20$ unités/heure, $\mu_c = 2$ heures, $\Gamma_{rég} = 500\$$, $C_s = 0.1$ \$/unité/heure, $C_p = 5$ \$/unité/heure, $C_{MP} = 200\$$, $C_{MC} = 1000\$$, $C_i = 15$ \$/produit, $C_r = 70$ \$/unité défectueuse, $\delta = 0.8$. Une procédure d'optimisation numérique itérative a été développée et codée avec le logiciel MATLAB®.

La meilleure stratégie consiste donc à prélever un échantillon de taille 27 toutes les 3,4 heures. De plus, en ce qui concerne la conception de la carte de contrôle, le nombre optimal d'écart-types entre la ligne centrale et les limites de surveillance est de 1,9, et le nombre optimal d'écart-types entre la ligne centrale et les limites de contrôle est de 2,6.

Cette stratégie consiste à effectuer une action de maintenance préventive vu que le dépassement des limites de surveillance survient avant celui des limites de contrôle ($j_p < j_c$). Le passage à l'état de surveillance se produirait en moyenne après 27,2 heures.

Il sera également nécessaire de constituer un stock tampon de 124 unités. La taille du lot correspondante est égale à 668 unités. Le coût total moyen par unité de temps correspondant à la politique optimale est de 483,63 \$/heure.

5 CONCLUSION

Ce papier propose un modèle pour une stratégie qui intègre l'inventaire, la production, la maintenance et le contrôle de la qualité, pour un système de production unitaire qui doit satisfaire une demande constante et continue. La stratégie intégrée est destinée aux processus de fabrication qui peuvent se déplacer de manière aléatoire vers un état hors de contrôle en raison de la dégradation de l'équipement. Deux états 'hors

contrôle' peuvent être détectés par la carte de contrôle. Le premier appelé 'État de surveillance' déclenche une action de maintenance préventive et le second 'État critique' déclenche une maintenance corrective de plus longue durée. Le modèle mathématique développé permet la détermination simultanée des valeurs optimales des paramètres de la carte de contrôle et la taille du stock tampon. Ce stock est construit pour assurer la satisfaction de la demande pendant les actions de maintenance. L'objectif est de minimiser le coût total attendu par unité de temps, qui comprend le réglage, l'inventaire, la maintenance, et les coûts attendus de qualité.

D'autres extensions consistent à considérer un modèle de maintenance imparfaite. Une deuxième perspective est de changer le type de la carte de contrôle et d'utiliser des cartes de types CuSum et EWMA en vue d'étudier la robustesse du modèle.

6 REFERENCES

- Bahria, N., Chelbi, A., Dridi, I.H. and Bouchriha, H. (2018). Maintenance and quality control integrated strategy for manufacturing systems, *European Journal of Industrial Engineering*, vol. 12(3), pp. 307-331.
- Ben-Daya, M. (2002). The economic production lot-sizing problem with imperfect production processes and imperfect maintenance. *International Journal for Production Economics*, vol. 76, pp. 257-264.
- Bouslah, B., Gharbi, A., & Pellerin, R. (2016). Integrated production, sampling quality control and maintenance of deteriorating production systems with AOQL constraint. *Omega*, vol. 61, pp. 110-126.
- Chakraborty, T., Giri, B.C., (2012). Joint determination of optimal safety stocks and production policy for an imperfect production system. *Applied Mathematical Modeling*, vol. 36, pp. 712-722.
- Chelbi, A., Rezg, N., & Radhoui M. (2008). Simultaneous determination of production lot size and preventive maintenance schedule for unreliable production system. *Journal of Quality in Maintenance Engineering*, vol. 14(2), pp. 161-176.
- Chen, Y. C. (2013). An optimal production and inspection strategy with preventive maintenance error and rework. *Journal of Manufacturing Systems*, vol. 32(1), pp. 99-106.
- Dhouib, K., Gharbi, A., Ben Aziza, M.N., (2012). Joint optimal production control/preventive maintenance policy for imperfect process manufacturing cell. *International Journal of Production Economics*, vol. 137, pp. 126-136.
- Montgomery, D. C. (2004) Introduction to statistical quality control, John Wiley & Sons, New York, NY.
- Pandey, D., Kulkarni, M.S., Vrat, P. (2011). A methodology for joint optimization for maintenance planning, process quality and production scheduling. *Computers and Industrial Engineering*, vol. 61, pp. 1098-1106.
- Radhoui, M., Rezg, N., Chelbi, A., (2010a). Joint quality control and preventive maintenance strategy for imperfect production processes. *Journal of Intelligent Manufacturing*, vol. 21, pp. 205-212.
- Radhoui, M., Rezg, N., Chelbi, A. (2010b). Integrated maintenance policy based on quality. *Computers and Industrial Engineering*, vol. 58, pp. 443-451.
- Salmasnia, A., B. Abdzadeh, and M. Namdar. 2017. "A Joint Design of Production run Length, Maintenance Policy and Control Chart with Multiple Assignable Causes." *Journal of Manufacturing Systems* 42: 44-56.
- Suliman, S. M., & Jawad, S. H. (2012). Optimization of preventive maintenance schedule and production lot size. *International Journal of Production Economics*, vol. 137(1), pp. 19-28.
- Rivera-Gómez, H., Gharbi, A., Kenné, J. P., Montañó-Arango, O., & Corona-Armenta, J. R. (2020). "Joint optimization of production and maintenance strategies considering a dynamic sampling strategy for a deteriorating system." *Computers & Industrial Engineering*, 140, 106273.

Génération de modèles de fiabilité opérationnelle

CHANGYI XU¹, EMIL DUMITRESCU¹, ERIC NIEL¹...

¹ AMPERE, CNRS, UNIV LYON, INSA LYON
VILLEURBANNE, FRANCE

prenom.nom@insa-lyon.fr

Résumé – Les faits marquants de ce travail se situent dans une approche collaborative de l'ingénierie et de la sûreté de fonctionnement au travers du concept MBSA (model-based safety assessment). Les aspects formels sur lesquels reposent la proposition (composition d'automates réguliers) sont cohérents à la rigueur de conception d'une part et aux exploitations qui y sont conséquentes d'autre part, notamment en termes de propriétés de validation. A ce titre, ce travail relève autant de la méthodologie que de la modélisation.

Abstract – The highlights of this work lie in a collaborative approach to engineering and operational safety through the MBSA concept (model based safety assessment). The formal aspects on which the proposal is based (composition of regular discrete-event automata) are consistent with the rigour of the design on the one hand and the operations that are consequent on it on the other, particularly in terms of validation properties. As such, this work is as much about methodology as it is about modelling.

Mots clés – MBSA, systèmes à événements discrets (SED), simulation

Keywords – MBSA, discrete-event systems (DES), simulation

1 INTRODUCTION

Les systèmes industriels sont conçus et pilotés pour fournir un service global délivré à partir de plusieurs composants en interaction et éventuellement distribués. Les interactions entre ces composants reposent sur des échanges d'information au travers d'interdépendances matérielles. Chaque composant est sujet à des défaillances conduisant à des états de panne et certains d'entre eux, moyennant des moyens appropriés peuvent être réparés. L'impact d'une défaillance d'un composant sur la prestation du service global attendu par le système peut varier en fonction des interdépendances relatives à ce composant. Ainsi, l'aboutissement d'une panne de composant n'est pas toujours synonyme d'une panne totale du système et ainsi de perte définitive du service global.

La garantie de continuité de service d'un système est généralement critique dans tout contexte industriel. Afin de fournir une telle garantie, les défaillances des composants individuels et leurs impacts possibles sur la prestation de services doivent être évalués et des opérations appropriées de d'isolation (de propagation de défaillance) doivent être anticipées. Ainsi, l'ingénierie de conception et de pilotage d'un système en vue d'assurer une délivrance de service global malgré l'occurrence de défaillance doit être complétée par un corpus d'action lié à l'analyse de la sûreté de fonctionnement. Les travaux abordés dans ce contexte concernent notamment l'assistance à la modélisation et l'évaluation de la fiabilité, la disponibilité et la maintenabilité d'un système structuré par des composants soumis à des défaillances.

Dans un tel contexte, le flux de conception de systèmes regroupe à la fois les points de vue fonctionnels et dysfonctionnels et l'évaluation des impacts de défaillance nécessite plusieurs étapes de modélisation : (1) une modélisation dysfonctionnelle locale, où l'étape de modélisation des composants sujets aux défaillances fournit des connaissances qualitatives et quantitatives sur les comportements défectueux et leur probabilité d'occurrence [Serfozo, 2009][Spitzer, 1971][Gagniuc, 2017]; (2) une modélisation structurelle du système où l'étape de conception du système définit la structuration des composants, support de propagation de défaillance, au travers de leurs interdépendances ; (3) une approche d'évaluation de la fiabilité basée sur un modèle, reposant sur les 2 modèles précédents. L'évaluation peut être qualitative ou quantitative et s'appuie sur des outils formels dédiés pour définir l'assurance de la continuité du service global. Cette dernière étape fournit à la fois des résultats quantitatifs (temps moyen de défaillance, de réparation, probabilités d'atteignabilité des états de panne, etc.) [Mondares et al. 2016] [Laprie, 2004] et qualitatifs (scénarios de propagation de défaillance et impact du choix des spécifications de reprise sur le service attendu).

La manipulation de ces différents types de modèles s'inscrit dans une démarche systémique classique de la sûreté de fonctionnement, encore appelée MBSA (Model Based Safety Analysis) [Prosvirnova et al., 2013].

L'originalité de cette proposition est de fournir une approche unifiée pour l'évaluation de la sûreté de fonctionnement (notamment fiabilité dans un premier temps) des systèmes structurés en tenant compte des exigences de service. L'objectif consiste alors à générer un modèle intégrant

dysfonctionnement, structure et exigences. Plusieurs étapes formelles de modélisation et de transformation sont préconisées, traitant successivement du comportement dysfonctionnel et de la structure du système au regard des propagations de défaillance et des exigences de service. Une première étape vise à relier de manière exhaustive les configurations de panne globale du système aux comportements dysfonctionnels des composants locaux. Chaque comportement dysfonctionnel d'un composant est formellement modélisé comme un système à événements discrets (« Event-driven automaton », EDA). Cette étape est fondamentale dans la gestion de l'aspect qualitatif et comportemental de chaque composant : composition, ajout de spécifications, identification d'états. Une deuxième étape consiste à identifier formellement, notamment à distinguer les états de panne locale et globale au niveau du système. À ce stade, des services spécifiques tels que les exigences de réparation peuvent être intégrés. Ceux-ci sont définis comme des modèles comportementaux, décrivant des séquences ou des priorités de réparation souhaitées, ou d'autres comportements opérationnels. Cette troisième étape produit un modèle global qui présente les comportements dysfonctionnels liés aux exigences de service. La quatrième et dernière étape cible les configurations de panne globale dans le modèle comportemental en vue des objectifs de l'analyse, fiabilité, disponibilité ou maintenabilité. Sans perte de généralité, seul est présenté dans ce travail l'identification de la panne totale, comme état absorbant du modèle à des fins de fiabilité dite opérationnelle *i.e.* tenant compte des spécifications de service.

Le système qui en résulte présente des trajectoires déclenchées par des événements menant à une panne globale définitive. Ces événements sont en général décrits de manière stochastique par les taux d'occurrence. Par conséquence, une chaîne de Markov en temps continu est dérivée du modèle comportemental étiqueté et offre les calculs habituels pour l'évaluation de la fiabilité.

Dans la suite de ce document, la démarche de conception est d'abord présentée dans la section 2. Une illustration sur un exemple simple est fournie section 3. L'extension vers la prise en compte de la fiabilité opérationnelle est présentée dans la section 4 puis illustrée section 5.

2 DEMARCHE PROPOSEE

2.1 Position du problème

Le domaine de la fiabilité exploite par excellence des données quantitatives afin de produire des résultats quantitatifs : des indicateurs de performance exprimant des probabilités, ou encore des temps moyens pour tout système structuré en composants. Cependant, la dimension qualitative du comportement intrinsèque à chacun des composants d'un système s'est avérée incontournable pour un calcul pertinent d'indicateurs locaux à chacun d'entre eux, préalable à une estimation globale de ces mêmes indicateurs. Pour cela, de nombreux modèles intègrent ces aspects comportementaux, tout en conservant l'expressivité d'aspects quantitatifs. Sans perte de généralité, les Chaînes de Markov à Temps Continu (CTMC) sont ici considérées comme modèle de référence.

La démarche classique d'ingénierie consiste à exprimer les aspects comportementaux et qualitatifs manuellement, pour un composant OU un sous-système composé de plusieurs composants, conjointement et au sein d'un même modèle CTMC. Cette démarche n'est pas extensible : (1) l'ajout d'un composant nécessite la réécriture

totale du modèle ; (2) l'ajout d'une exigence comportementale régissant deux ou plusieurs composants nécessite elle aussi une réécriture manuelle et intégrale. Cette démarche est coûteuse en temps de modélisation, et sujette aux erreurs (humaines).

Ce travail propose d'exploiter le cadre fonctionnel et formel de la modélisation comportementale à événements discrets afin de non pas réécrire des modèles, mais les générer, avec des garanties de correction par construction, comme le montre la Figure 1. L'hypothèse centrale est celle de l'existence d'une bijection entre les états/transitions d'un modèle CTMC d'un composant et son modèle comportemental EDA équivalent. Ceci permet d'obtenir un modèle EDA à partir de chaque CTMC et de basculer ainsi le champ de l'étude d'un domaine mixte quantitatif/qualitatif vers un domaine qualitatif offrant de outils à la fois rigoureux et puissants. Il devient ainsi possible d'exprimer des exigences de gestion opérationnelle (ETS) et de les intégrer de manière automatique dans le modèle comportemental global (GRA).

En exploitant la même hypothèse, le GRA peut être transformé en CTMC, prêt pour un calcul d'indicateurs de performance intégrant des aspects opérationnels, et obtenu automatiquement. Cet aspect n'est pas développé dans ce document.

2.2 Le flot de conception global

La figure 1 illustre le cadre de modélisation modulaire de cette proposition. Le modèle cible est appelé 'GRA opérationnel'. Il y a trois entrées de cette proposition : (1) les comportements des composants défectueux, tels que les modèle CTMC des composants, (2) la structure du système et (3) les exigences de gestion opérationnelle, telles que les exigences de planification de la séquence de réparation des composants. Sur la base de la conception du système et des comportements des composants défectueux, le « Global Reliability Automaton » (GRA) opérationnel est généré. Cette génération de modèle GRA se fait en six étapes :

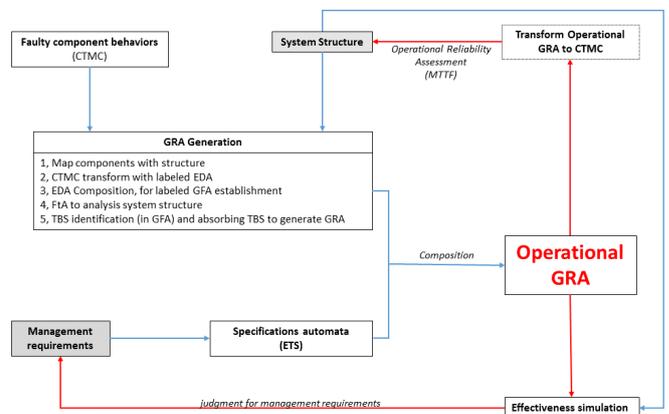


Figure 1. Flot de conception global

1. associer le modèle CTMC avec la structure correspondante ;
2. transformer le modèle CTMC en EDA (Event Driven Automaton) étiqueté. L'étiquetage correspond à l'association de propositions booléennes pour décrire les caractéristiques des états ;
3. composer des EDA constituant le système, afin de générer un automate global GFA (Global Faulty Automaton) ;
4. utiliser l'analyse d'arbres de faute (FtA) [Clemens, 2002] pour analyser la structure du système afin d'obtenir une expression logique de défaillance ;

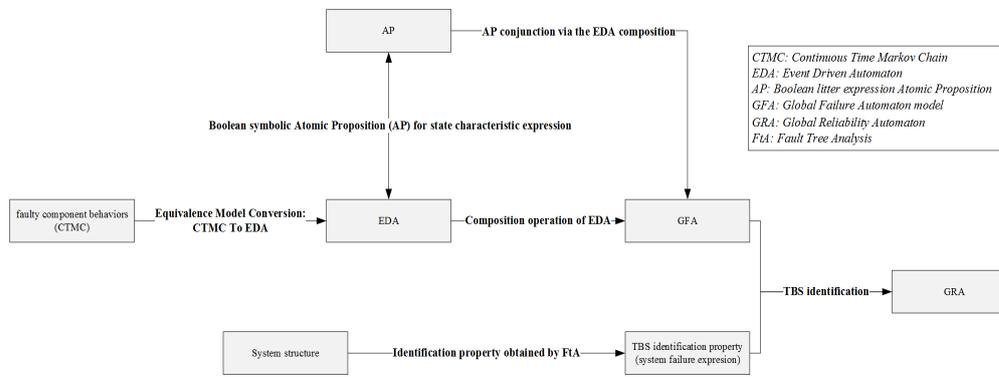


Figure 2. Génération du modèle GRA

5. identifier l'état de panne totale, TBS (Total Broken Down State), en s'appuyant sur l'expression logique de défaillance. Le TBS est rendu « absorbant » pour générer le GRA.
6. concevoir en fonction des exigences de gestion, des automates de spécifications (ETS). La composition du GRA et des automates de spécification établit le GRA Opérationnel.

Une étape ultérieure de simulation permet d'estimer les performances du système et d'améliorer éventuellement ses caractéristiques.

2.3 La génération du modèle de fiabilité (GRA)

La Figure 2 détaille l'approche de génération en six étapes proposée plus haut. Il est considéré que les taux de passage des transitions dans chaque modèle CTMC correspondent aux événements de transition dans chaque EDA (Event Driven Automaton). Les travaux de [Garg et al., 1999],[Garg, 1992], et de [Ionescu, 2016] évoquent la transformation du modèle CTMC en EDA.

A travers une opération de composition, un modèle global de faute (GFA) est obtenu. Celui-ci possède un état initial ainsi qu'un ou plusieurs états de panne totale du système (TBS). En anticipation des calculs quantitatifs ultérieurs, les états TBS dans GFA doivent être rendus *absorbants*. Le modèle obtenu ainsi est le GRA.

L'identification des états de panne (TBS) totale s'appuie sur l'emploi d'étiquettes sous forme de propositions Booléennes atomiques (AP), associées aux états des composants, comme détaillé à l'étape 4 ci-dessous.

La FtA (Fault Tree Analysis) [Clemens, 2002] est utilisée pour analyser la structure du système. Elle permet d'établir une expression logique exprimant les états de défaillance du système et s'appuyant sur les AP associés aux états. L'ensemble de ces six étapes sont détaillées dans la suite.

Étape 1 : Conversion des modèles d'équivalence : CTMC et EDA

Il s'agit de définir les équivalences des modèles stochastiques CTMC et automates réguliers EDA afin de permettre la composition formelle de modèles et d'accéder ainsi du comportement dysfonctionnel du local (composant) au global (système). Soit une CTMC définie par :

$$CTMC = (Q_{mc}; E_{mc}; \delta_{mc}; q0_{mc})$$

où Q_{mc} est l'ensemble d'états, E_{mc} est l'ensemble de transition, δ_{mc} est la fonction de transition, $q0$ est l'état initial. La dynamique du modèle peut être présentée par l'union des séquences d'événements associées aux états, l'ensemble des états associés x_i est équivalent à l'état défini dans CTMC :

$\{x_i\} = Q_{mc}$. La fonction de transition dans une CTMC décrit l'évolution des états conduite par les événements.

Soit un EDA (Event Driven Automaton) défini par :

$$EDA = (Q_{EDA}; E_{EDA}; \delta_{EDA}; q0_{EDA}; qm_{EDA})$$

où Q_{EDA} est l'ensemble d'états, E_{EDA} est l'ensemble de transition, δ_{EDA} est la fonction de transition, $q0_{EDA}$ est l'état initial et qm_{EDA} est l'état marqué. Par le même principe que précédemment, on peut avoir : $\{x_i\} = Q_{EDA}$

Étant donné que les modèles CTMC et EDA sont construits pour le même système, les transitions et les états sont identiques :

$$Q_{EDA} = \{x_i\} = Q_{mc} \text{ et } \delta_{mc} = \delta_{EDA}$$

pour le modèle EDA: $\delta_{EDA}(x, e_{xy}) = y$ et pour le modèle CTMC : $\delta_{mc}(x, e_{xy}) = y$, de taux : $R \cdot e_{xy}$.

Étape 2 : Opération de composition d'EDA

Pour la raison que le CTMC peut être transformé en EDA, la composition des comportements defectueux des composants qui est décrite par CTMC peut être émise par son opération de composition EDA équivalente. Cette partie étudie la méthode formelle de l'opération de composition d'automates.

Conformément à la définition du EDA, la composition de EDA est basée sur l'opération de composition des automates dans la théorie des systèmes d'événements discrets [Cassandras, 2009]. Soient deux automates $EDA1 = (Q1; E1; \delta1; q01; qm1)$ et $EDA2 = (Q2; E2; \delta2; q02; qm2)$. L'opération de composition (notée '//') est définie :

$$EDA1 // EDA2 = (Q1 \times Q2; E1 \cup E2; \delta; q01 \times q02; qm1 \times qm2;)$$

tel que :

$$\delta((x,y), t) = \begin{cases} (x', y) & \text{Si } t \in E1 \setminus (E1 \cap E2) \text{ et } \delta1(x,t) = x' \\ (x, y') & \text{Si } t \in E2 \setminus (E1 \cap E2) \text{ et } \delta2(y,t) = y' \\ (x', y') & \text{Si } t \in (E1 \cap E2) \text{ et } \delta1(x,t) = x' \text{ et } \delta2(y,t) = y' \end{cases}$$

Sur la base de la définition ci-dessus, si la transition n'est pas commune aux deux modèles, elle appartient aux deux premières conditions. La simultanéité est traduite par la troisième.

Étape 3 : Identification des états à l'aide de Propositions atomiques booléennes (AP)

Seul le nom d'un état n'est pas suffisamment explicatif pour l'identifier. La proposition atomique booléenne (AP) appliquée aux modèles EDA permet d'étiqueter les états et ainsi assurer une identification *symbolique*. Les modèles EDA sont alors exprimés de la sorte :

$$EDA = (Q; E; \delta; q0; qm; AP; L)$$

où AP est l'ensemble des propositions atomiques exploités dans l'identification des états, fautifs ou non. Le symbole booléen F (signifie 'fails') associé à un état signifie que celui-ci est un état de panne, et sa négation $\neg F$ indique un état de fonctionnement correct. Il est alors possible d'obtenir par conjonction d'étiquettes (AP), la caractérisation des états combinés résultant d'opération de composition de modèles.

Supposons l'ensemble AP :

$$AP = \{ap_0, ap_1, ap_2, \dots, ap_n\}$$

L'ensemble des formules de AP est $2^{AP} = \{\emptyset, ap_0, ap_1, ap_2, ap_0 \wedge ap_1, ap_0 \wedge ap_2, ap_1 \wedge ap_2, ap_0 \wedge ap_1 \wedge ap_2, \dots\}$

L'expression des symboles AP relève de situations réelles du système, et la fonction d'association est motivée par la signification de l'état. Si un état a plus d'une caractéristique, il s'agira d'un élément combiné (par exemple, $ap_0 \wedge ap_1$) dans 2^{AP} associé à cet état. En outre, si chaque nom d'état est unique, mais il est possible que plusieurs états partagent le même AP.

Étape 4 : Composition des étiquettes locales

État donné que les états système sont le résultat produit cartésien des états des composants, ces états système résultants possèdent toutes les caractéristiques des composants correspondants. L'opération de composition des AP appliquée à la composition de modèles EDA définit le Global Faulty Automaton GFA comme suit :

$$GFA = EDA1 // EDA2 = AC (Q1 \times Q2; E1 \cup E2; \delta; q01 \times q02; qm1 \times qm2; AP1 \cup AP2; L1 \wedge L2)$$

$AP1 \cup AP2$ est l'union de l'ensemble de propositions atomiques. $L1 \wedge L2$ est la fonction d'association d'AP pour $EDA1 // EDA2$ est :

$$L1 \wedge L2: Q1 \times Q2 \rightarrow 2^{AP1 \cup AP2}$$

Étape 5 : Expression de la propriété de défaillance

Les états sont ainsi identifiés par leur symbole. Les diverses combinaisons logiques d'AP sont l'identification des états composés. Tout état peut être ainsi identifié, il en va de même pour l'identification du TBS. Le TBS est un état redouté dans le système, il résulte de la propagation d'un événement (ou séquence) redouté. Les arbres de fautes (FtA) expriment cette propagation par la combinaison logique d'événement initiaux. Cette combinaison résulte de la structure collaborative des composants constituant le système. L'entrée des feuilles du FtA portera un symbole AP typé défaillance, et l'expression logique de l'événement redouté est alors la combinaison des AP de composants. Ainsi, le TBS (Total Breakdown State) peut être identifié dans le GFA.

Étape 6 : Identification du TBS (Total Breakdown State)

L'ensemble des états satisfaisant une formule Φ_i , est noté $IS(\Phi_i)$ étant défini comme :

$$IS(\Phi_i) = \{x \mid x \in Q, L(x) \models \Phi_i\}$$

Ici, le symbole ' \models ' désigne le fait que l'étiquette $L(x)$ de l'état x est une condition suffisante de Φ_i .

$IS(\Phi_i)$ est un sous-ensemble de l'ensemble d'états Q : $IS(\Phi_i) \subseteq Q$. Si aucun état AP ne satisfait la propriété d'identification, $IS(\Phi_i)$ est un ensemble vide.

La génération du GRA résulte de l'identification du TBS sur le GFA et de la transformation des états TBS en états absorbants *i.e.* supprimer sur le modèle GFA, les transitions de sortie de l'ensemble des états étiquetés TBS. Comme certaines transitions de sortie ont été supprimées, il est possible qu'il existe plusieurs états inaccessibles à partir de l'état initial. Ces états n'ont aucun sens dans le GRA et sont donc supprimés.

3 EXEMPLE : GENERATION D'UN GRA

L'exemple considéré ici est constitué de trois composants G1, G2 et G3. Les états W1, W2 et W3 représentent les états de fonctionnement sain du système ; les états B1, B2 et B3 représentent les états de panne du système ; chaque composant

a un comportement de panne unique de f_1, f_2 et f_3 , et une opération de réparation unique de r_1, r_2 et r_3 . Les ensembles de PA booléens $\{FG1 ; FG2 ; FG3\}$ sont utilisés pour indiquer l'état de panne des composants, tandis que la symbolique négative de ces PA présente les composants en bon état de fonctionnement. Ce système est illustré dans la Figure 3. Le composant G1 est critique, alors que G2 et G3 offrent un schéma de redondance, tel que représenté dans la Figure 4.

Pour déterminer de quelle manière la panne totale du système peut se produire, l'analyse de l'arbre de défaillance (FtA) est appliquée à cette structure afin d'obtenir l'expression logique de la panne". L'expression logique de défaillance" est la propriété d'identification de l'état du TBS, car chaque état est déjà associé à l'AP. Comme le montre la figure 5 la FtA utilise une porte "et" reliant la panne (désignée par AP) de G2 et G3 en raison de la structure parallèle ; elle utilise une porte "ou" reliant la panne de G1 et la porte "et" précédente en raison de la structure en série, l'expression logique de défaillance est donc $\Phi_3 = FG1 \vee (FG2 \wedge FG3)$.

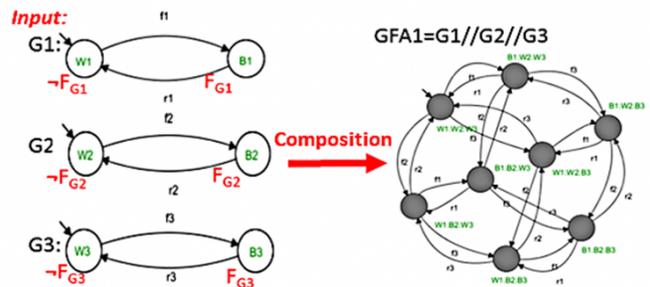


Figure 3. Système à trois composants

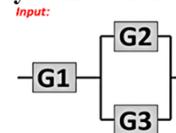


Figure 4. Structure « série-parallèle »

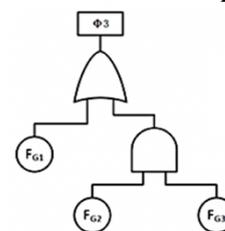


Figure 5. Arbre de défaillance correspondant

Les états de panne totale (TBS), identifiés conformément à la procédure décrite plus haut sont $IS(\Phi_3) = \{ B1.B2.B3 ; B1.B2.W3 ; B1.W2.B3 ; B1.W2.W3 ; W1.B2.B3 \}$, et ils sont marqués dans le cercle rouge de la Figure 7. Ces TBS sont ensuite « absorbés » pour générer le GRA1, puis les états inaccessibles sont supprimés. Le GRA1 est illustré à la Figure 8.

4 LA FIABILITE OPERATIONNELLE

La Figure 6 illustre l'enrichissement du modèle GFA (Global Faulty Automaton) avec la notion d'opérationnalité. On s'appuie pour cela sur la spécification d'une trajectoire d'événements (issue des exigences de gestion). Il s'agit de contraindre le comportement du système par l'expression d'une spécification relative à la « Sûreté de Fonctionnement » : ici une trajectoire de reprise de fonctionnement après panne.

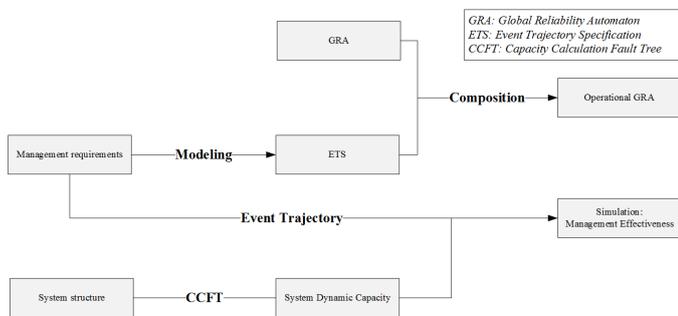


Figure 6. Ajout de spécifications

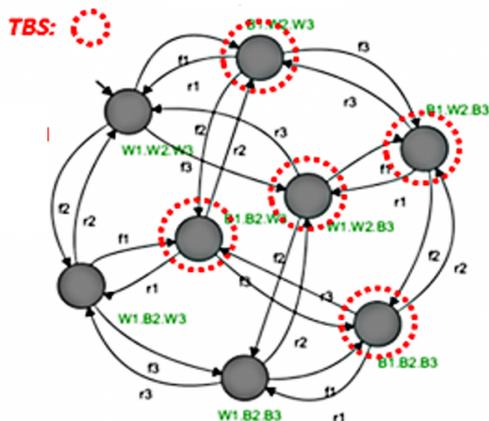


Figure 7. Identification du TBS

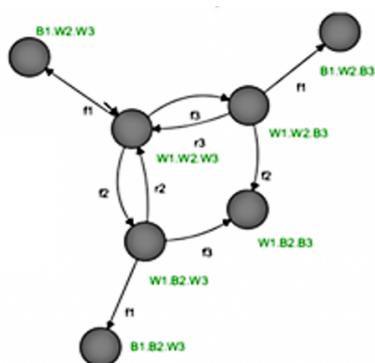


Figure 8. GRA résultant

On obtient ainsi un automate de spécification de trajectoires d'événements (ETS, Event Trajectory Specification Automaton) [Cassandras, 2009][Lo, 2006]. Les exigences de gestion sont définies en fonction des différentes situations du système, qui forme la trajectoire de réparation dans le modèle. Par composition du GRA et de l'ETS, le GRA opérationnel est établi.

La fiabilité opérationnelle a été examinée en ciblant le TBS et en le considérant comme un état final. Cela exclut toute opération de réparation pour un système totalement défaillant. Cependant les trajectoires de réparation font partie des modèles de l'EDA et peuvent donc être facilement incluses lors du ciblage du TBS. Ainsi, la prise en compte de la fiabilité opérationnelle est réalisée en planifiant les trajectoires de réparation possibles dans le modèle global de l'EDA. D'autres exigences de gestion peuvent être spécifiées afin d'exprimer les priorités de réparation, la maintenance pour le service, etc. Le modèle qui en résulte est exprimé sous la forme d'un EDA supplémentaire, qui doit être composée avec le modèle global

GFA. La fiabilité opérationnelle peut être évaluée, qualitativement et quantitativement, en s'appuyant sur le modèle résultant.

Il est important de souligner qu'il s'agit là d'un des points forts de cette proposition, par rapport au corpus de méthodes d'évaluation de la sûreté de fonctionnement. En outre, son aspect formel offre la possibilité d'introduire la vérification formelle de propriétés telles que la vivacité, l'accessibilité, non développées à l'état actuel de ce travail mais totalement accessibles en tant que perspectives. En outre, cette notion de traitement de la structure d'une part et des contraintes opérationnelles d'autre part, reste proche des méthodes et outils d'ingénierie existants.

5 EXEMPLE : PRISE EN COMPTE DE LA FIABILITE OPERATIONNELLE

Considérons le système de la figure 4, réduit à deux composants, tel qu'illustré dans la figure 9.

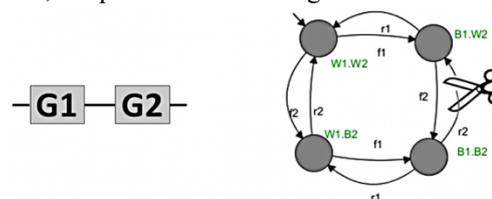


Figure 9. Système à deux composants : structure et EDA

L'exigence de fiabilité opérationnelle envisagée dans cet exemple stipule que « l'élément G1, le plus important, est prioritaire pour la réparation ». Ainsi, lorsque les deux composants sont en panne, la réparation de G1 est prioritaire. Cette exigence serait satisfaite si la transition B1.W2 vers B1.W2 pouvait être inhibée, comme le montre la figure 9.

Afin de renforcer la satisfaction de cette exigence, elle est d'abord modélisée par un ETS, tel que représenté dans la figure 10.

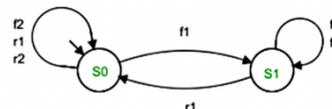


Figure 10. ETS : spécification de la priorité à la réparation

En considérant qu'après l'occurrence de f1, il faut réparer G1 puis aller à l'état initial (la réparation de G2 est interdite ici), la transition interdite (réparation de G2, lorsque les deux composants tombent en panne) est rejetée. En composant l'ETS avec l'EDA modélisant les deux composants, on obtient un GRA qui satisfait l'exigence modélisée par l'exigence de priorité à la réparation (Figure 11).

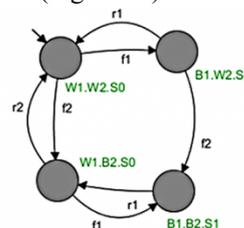


Figure 11. GRA résultant : spécification de la priorité à la réparation

Des modèles similaires peuvent être exprimés pour prendre en compte plusieurs composants, ou pour exprimer des spécifications différentes, telles que les FIFOs.

6 CONCLUSIONS ET PERSPECTIVES

La théorie des systèmes à événements discrets (SED) est appliquée dans un cadre de type « MBSA », sur la base de ses principales capacités pour la génération de modèles. L'existence d'un lien entre EDA (Event Driven Automaton) et CTMC (Continuous Time Markov Chain) est illustrée. Ici, un EDA valide la séquence et la structure à travers des déclarations de propriétés telles que l'observabilité et l'accessibilité à des états spécifiques tels que TBS, tandis qu'une CTMC valide la fonction d'évaluation du modèle. Grâce à l'opération de composition un modèle intégrant une spécification est construit automatiquement. Ainsi, dans le modèle généré, les comportements défectueux, les configurations complexes et les exigences opérationnelles sont pris en compte.

Les faits marquants de ce travail se situent dans une approche collaborative de l'ingénierie et de la sûreté de fonctionnement au travers du concept MBSA (model based safety assessment). Les aspects formels sur lesquels reposent la proposition (composition d'automates réguliers) sont cohérents à la rigueur de conception d'une part et aux exploitations qui y sont conséquentes d'autre part, notamment en termes de propriétés de validation. A ce titre, ce travail relève autant de la méthodologie que de la modélisation.

De plus, malgré son caractère exploratoire, cette approche est limitée dans les aspects suivants : pas d'interface de données directe avec l'outil de calcul des indicateurs de fiabilité, la complexité du modèle résultant, la capacité d'analyse de la structure, la description des comportements des composants et la capacité de calcul de la capacité. Bien qu'un modèle soit réalisé, ce modèle ne peut pas être directement et automatiquement entré dans un outil de calcul d'indicateurs de temps de fiabilité, par exemple GRIF ou AltaRica.

Les travaux en cours expriment l'intérêt pour les avancées ultérieures des points suivants.

L'interface de données avec l'outil de calcul des indicateurs de fiabilité : le développement des plans de travail futurs d'un programme de traduction de modèle formel, afin de réaliser une interface de données du modèle obtenu aux outils de calcul des indicateurs de fiabilité, tels que GRIF ou AltaRica.

Ensuite, la complexité du modèle de système : plus de limitations d'état devraient être définies dans les hypothèses, ou des technologies d'optimisation de modèle possibles devraient être recherchées, afin d'éviter le problème potentiel d'explosion d'état.

La structure du système : cette approche utilise les outils d'analyse de structure normale (FtA et RBD) pour obtenir l'expression logique de défaillance. La limitation de ces outils restreint la capacité d'obtenir l'expression logique de défaillance, ainsi les TBS ne peuvent pas être entièrement identifiables dans le cas où la structure du système ne serait pas analysable par ces outils. Par exemple, les structures mixtes ne peuvent pas être entièrement analysées par la FtA ou la RBD, de sorte que ce type de systèmes structurels ne peut pas être évalué formellement par notre approche. Ainsi, le scénario de l'analyse de la structure, de l'obtention de l'expression logique de défaillance et de l'identification des états sera amélioré par les travaux en cours.

Le modèle de disponibilité et modèle de maintenabilité : La génération du modèle de fiabilité est introduite dans cette thèse. De plus, le modèle de disponibilité et le modèle de maintenabilité peuvent également être générés par un domaine similaire. Les travaux en cours étudieront l'approche de

génération du modèle de maintenabilité, principalement en identifiant les « états de fonctionnement total' » et en supprimant leurs transitions de sortie par l'application de l'analyse de l'arbre de réussite. En effet, l'évaluation de la maintenabilité démarre avec l'état de panne totale et se termine par les états de fonctionnement. Les travaux sur le modèle de disponibilité se concentreront sur le fonctionnement de la composition des modèles locaux, notamment dans lequel la partie compatible et incompatible sera étudiée lors de la mise en place des modèles de composants. Intuitivement, les événements et les états de l'EDA seront formellement distingués comme étant compatibles ou incompatibles entre les différentes chaînes de Markov.

7 REFERENCES

- M. Modarres, M. P. Kaminskiy, and V. Krivtsov, *Reliability engineering and risk analysis: a practical guide*, CRC press, 2016.
- J.-C. Laprie, "Sûreté de fonctionnement informatique: concepts, défis, directions," *Computing*, vol. 1, no. 1, pp. 11-33, 2004.
- P. Clemens, "Fault tree analysis," *JE Jacobs Severdurup*, pp. 13, 2002.
- R. Serfozo, *Basics of applied stochastic processes*, Springer Science & Business Media, 2009.
- F. L. Spitzer, *Random fields and interacting particle systems*, Mathematical Association of America Oberlin-Ohio, 1971.
- P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*, John Wiley & Sons, 2017.
- T. Prosvirnova, M. Batteux, P.-A. Brameret, A. Cherfi, T. Friedlhuber, J.-M. Roussel, and A. Rauzy, "The altarica 3.0 project for model-based safety assessment," *IFAC Proceedings Volumes*, vol. 46, no. 22, pp. 127-132, 2013.
- V. K. Garg, R. Kumar, and S. I. Marcus, "A probabilistic language formalism for stochastic discrete-event systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 280-293, 1999.
- V. K. Garg, "An algebraic approach to modeling probabilistic discrete event systems." *Decision and Control, 1992.*, Proceedings of the 31st IEEE Conference on, IEEE, pp. 2348-2353, 1992.
- D.-R. Ionescu, "Evaluation quantitative de séquences d'événements en sûreté de fonctionnement à l'aide de la théorie des langages probabilistes," *Automatique / Robotique*, Université de Lorraine, 2016.
- C. G. Cassandras, and S. Lafortune, *Introduction to discrete event systems*, Springer Science & Business Media, 2009.
- D. Lo, and S.-C. Khoo, "QUARK: Empirical assessment of automaton-based specification miners." *13th Working Conference on Reverse Engineering, IEEE*, pp. 51-60, 2006.

Session 3.1 : Conception, Innovation, Fabrication

Chairman : Bernard Yannou - Guillaume Thomann

CIGI 2021

Comment hybrider les avantages des approches agiles avec le phasage traditionnel pour les projets industriels ?

THIERRY GIDEL¹, WILLIAM ZONGHERO²

¹Université de Technologie de Compiègne
Adresse postale, Pays
thierry.gidel@utc.fr

² Consultant
19 chemin de Belvoir 05100 PUY SAI NT PIERRE
wzonghero@gmail.com

Résumé – Cet article présente les principes permettant une hybridation du management classique et agile des projets industriels. Après avoir présenté les raisons qui ont conduit à élaborer les principes du management de projet classique à partir des années 70, nous présentons les principes sous-jacents au management de projet agile. Cela nous permet de mettre en évidence les points communs, mais aussi les divergences entre ces principes et les organisations, méthodes et outils associés. Partant de ces éléments théoriques et en s'appuyant sur des études de cas d'hybridations, nous proposons un cadre favorisant cette hybridation. En conclusion, nous identifions les perspectives de recherches qui permettraient d'affiner cette première proposition.

Abstract – This paper presents the principles allowing the hybridisation of classic and agile management of industrial projects. After presenting the reasons which led to the development of the principles of classic project management from the 70s, we present the underlying principles of agile project management. This allows us to highlight the common points, but also the divergences between these principles and the associated organisations, methods and tools. Starting from these theoretical elements and drawing on hybridisation case studies, we propose a framework that promotes this hybridisation. In conclusion, we identify the research perspectives that would make it possible to refine this first proposition.

Mots clés – Management, Projet, Agile, Hybride, Classique.

Keywords – Management, Project, Agile, Hybrid, Classic.

1 À L'ORIGINE DU MANAGEMENT DE PROJET CLASSIQUE

Pour analyser les avantages et inconvénients des approches classique et agile du management de projet, il faut d'abord revenir aux raisons qui ont conduit à l'introduction des projets dans les entreprises et à l'élaboration des principes, méthodes et outils du management de projet que nous qualifions de classique.

Au départ, le management de projets a été élaboré pour mobiliser et organiser les différentes ressources nécessaires à la réalisation d'infrastructure importantes (Garel 2011) puis, pour des activités innovantes et non répétitives dans les entreprises telles que la conception et lancement d'un nouveau produit, la mise en place d'un nouveau process de production ou d'action spécifique par exemple sur la qualité (Gidel et Zonghero 2020). Cela avec des résultats supérieurs à ceux obtenus avec une organisation d'entreprise par métiers/fonctions « en silos » (Giard et Midler 1993).

En effet, les entreprises sont organisées pour produire en continu et de manière stabilisée afin de générer un chiffre d'affaires et une rentabilité assurant la pérennité de l'entreprise. Pour arriver à ce but, une organisation, des méthodes et outils ont été développés pour gérer les processus de production répétitifs, gérables statistiquement, entraînant des décisions réversibles pour l'entreprise.

Dans ce contexte, toutes les activités nouvelles, uniques, non répétitives, viennent perturber l'organisation qui a été mise en place et viennent aussi puiser dans les ressources habituellement dédiées à la production. Par ailleurs, réaliser ces activités nouvelles pour l'entreprise nécessitent des investissements et des prises de décisions souvent irréversibles ainsi qu'un suivi spécifique qui ont conduit à l'élaboration de plusieurs principes fondamentaux du management de projet. Un des premiers principes est relatif à l'identification des acteurs des projets permettant l'instauration de négociations entre les demandeurs (client ou maîtrise d'ouvrage, MOA) et les réalisateurs (chef de projet ou maîtrise d'œuvre, MOE), redonnant une place et une visibilité à l'approche client (Giard 1991). Ensuite le principe

de phasage et jalonnement des projets a été proposé. Celui-ci va de pair avec le concept d'engagement progressif (Chvidchenko et Chevalier 1997) qui permet l'élaboration et le figeage progressif des livrables du projet, à commencer par l'expression précise des besoins (matérialisés par la validation du CdC) lors de la phase d'avant-projet. Ce principe sous-entend aussi que les premières phases sont consacrées aux activités les moins coûteuses de planification et de définition du projet et permettent de prendre les décisions critiques d'investissement engageant financièrement l'entreprise (Gidel et Zonghero 2006). Ce principe est influencé par les approches qualité en particulier les notions d'actions préventives qui préconisent d'investir en amont plutôt que de devoir corriger les problèmes à postériori.

Pour mettre en œuvre ces principes du management de projet, des organisations, méthodes et outils ont été développés à partir des années 70 (Garel 2011). Pour « casser » les silos conduisant au cloisonnement des métiers de l'entreprise, des organisations matricielles ont été proposées (Clark, Hates, et Wheelwright 1988). Les rôles de coordinateur, chef ou directeur de projet ainsi que celui de la maîtrise d'ouvrage ou comité de pilotage ont été instaurés. Des nouveaux types de contrats ont été formalisés (par ex. forfaitaires ou en régie). Des réunions entre les membres de l'équipe projet ainsi que des revues de projet réalisées lors des jalons en présence des clients (MOA) sont organisées. Des méthodes et outils de structuration du projet ont été formalisés voir normalisés (*Product Breakdown Structure* PBS, *Organisation Breakdown Structure* OBS, *Work Breakdown Structure* WBS, ou Organigramme des Tâches OT, etc.) (PMI 2017). Des outils de planification ont été mis en œuvre (PERT, Gantt...), ainsi que des outils de maîtrise : mesure de l'avancement physique et délais, méthode de la valeur acquise, coûts, qualité, santé, environnement, risques, etc. (Kerzner 2017).

Pour réduire encore la durée des projets et compenser une certaine linéarité de ces premières approches, de nouveaux principes ont été introduits à partir des années 90. Les concepts d'ingénierie concourante ou simultanée ont été développés permettant plus de flexibilité dans la mise en œuvre du phasage et jalonnement (Garel, Giard, et Midler 2004). Cela a nécessité de faire évoluer les organisations matricielles en plateau projet avec des équipes dédiées (Clark et al. 1988) et des outils de travail partagés (cf. maquette numérique).

Aussi, partant de ces principes, beaucoup de méthodes et outils spécifiques à certaines industries et contextes ont été développés, standardisés, voire normalisés (NF EN 9200 : 2005 ; ISO 21500 : 2012). Des organisations ont été imaginées, par exemple avec architecte industriel, des contrats ont été formalisés, par exemple coût plus honoraires incitatifs, prix maximum garanti (PMG) ou marché en main, des méthodes développées pour maîtriser les délais, les coûts, les performances ou les risques (Courtot 1998). La diffusion des méthodes de management de projet dans les entreprises (les *plannings* glissants pour les constructions de routes ou voies ferrées, etc.) et la multiplication des projets ont aussi conduit au développement des méthodes de management multi-projet avec des portefeuilles, plateformes ou programmes (Cooper, Edgett, et Kleinschmidt 1999; Cusumano et Nobeoka 1999) et pour le management de l'innovation (Fernex-Walch et Romon 2013).

Parallèlement à ces évolutions, il faut noter que beaucoup de praticiens et de chercheurs ont dénoncé certaines dérives bureaucratiques liées à la mise en œuvre de ces principes,

méthodes et outils, conduisant à perdre de vue les objectifs initiaux au profit d'une standardisation excessive. Le *planning* n'est plus un outil de pilotage, mais une procédure à respecter au détriment de l'objectif initial du client. Par exemple, les tâches simples sont réalisées au début afin d'améliorer les indicateurs de performance du projet (taux d'avancement) au lieu de réaliser les tâches les plus complexes associées aux fonctions essentielles du produit du projet (Jolivet 2003; Lenfle et Loch 2010).

2 ÉMERGENCE DES PRINCIPES, ORGANISATIONS, METHODES ET OUTILS DU MANAGEMENT AGILES DE PROJET

Partant de ces principes, développés pour le management classique des projets industriels à partir des années 70, comment expliquer l'engouement pour les méthodologies agiles développées pour le secteur informatique à partir des années 2000 (Bianchi, Marzi, et Guerini 2020; Campanelli et Parreiras 2015). Autrement dit, quels sont les principes nouveaux que les approches agiles apportent et quels sont les principes existants qui sont remis en cause ou détournés ? Qu'est-ce que les approches agiles conservent ou modifient à ces principes ou approches développées pour les projets industriels ?

Pour répondre à ces questions, il faut d'abord revenir sur le contexte et son évolution. Puis nous proposons d'identifier les principes, organisations, méthodes et outils développés pour gérer les projets de manière plus agile.

2.1 Évolution du contexte

Les méthodes agiles ont fait leur apparition dans le secteur informatique, en particulier pour le développement d'applications (Hoda et al. 2017). Dans ce secteur dont les technologies évoluent rapidement, il est à la fois difficile de définir les besoins des utilisateurs en amont du projet et dans le même temps assez facile d'intégrer des modifications en aval des projets. Ce contexte, confronté au constat d'une dérive procédurière du management de projet classique, a conduit un groupe d'informaticiens à formaliser un manifeste agile au début des années 2000 à partir des pratiques existantes (Beck et al. 2001). L'objectif de cette proposition étant de remettre les utilisateurs au cœur du processus en focalisant d'une part sur les fonctions du logiciel plutôt que sur le contrat ou la documentation, et d'autre part, en responsabilisant les membres de l'équipe du projet plutôt que de s'appuyer sur le respect de procédures pas toujours comprises ou acceptées par l'équipe. Plusieurs principes fondamentaux émergent de ces propositions.

2.2 Principes agiles

La planification des projets agiles est basée sur le phasage à durée fixe, des itérations ou *sprint* de 2 à 4 semaines. Cela implique une identification des livrables intermédiaires utiles aux utilisateurs à court terme, c'est-à-dire exploitable. Ces livrables devant être présentés aux clients / utilisateurs à la fin de chaque itération afin de recueillir leur avis.

La planification par la demande est aussi une caractéristique des projets agiles : les besoins des parties intéressées fixent les priorités, l'équipe de travail peut également modifier l'ordre des tâches en commençant, par exemple, par les plus risquées.

Les fonctions du projet se définissent progressivement à mesure que les parties intéressées expriment leurs besoins selon les résultats des livrables intermédiaires.

Un autre principe des méthodes agiles, est l'autonomie de l'équipe de travail : il n'y a pas de méthode imposée, les membres de l'équipe définissent eux-mêmes leur méthode de travail et les font évoluer en cours de projet en se basant sur le retour d'expérience. C'est une des raisons qui rendent ces

méthodes insaisissables : il existe plusieurs dizaines de méthodes agiles (Denning 2018). Mais c'est aussi ce principe de laisser le choix des méthodes aux équipes qui permet de renforcer leur appropriation et par conséquent l'engagement et la motivation des équipes.

Pour mettre en œuvre ces principes agiles, des organisations, méthodes et outils ont été formalisés.

2.3 Organisations agiles

Les organisations agiles préconisent d'intégrer les clients, utilisateurs ou parties intéressées à l'équipe projet élargie : ils participent à minima aux revues d'itération, mais peuvent aussi intervenir en cours d'itération. Cette organisation a plusieurs conséquences. D'une part le recueil des besoins est facilité et les priorités décidées en commun. D'autre part, il n'y a plus de contrôle qualité en tant que tel, car les livrables intermédiaires sont présentés et validés au fur et à mesure de l'avancement par les parties intéressées.

Les équipes sont de petite taille (inférieures à 10 personnes), stabilisées et s'auto-organisent : auto-assignation des rôles et des tâches. Il est aussi préconisé d'avoir des équipes co-localisées pour faciliter la communication, en particulier en s'appuyant sur le management visuel.

Les rôles clés sont au nombre de trois : les équipiers d'une part. Ceux-ci identifient l'un d'entre eux pour animer les réunions et revues, coacher les équipiers et résoudre les problèmes ou obstacles, notamment logistique de l'équipe en s'assurant qu'ils ont les moyens nécessaires pour travailler. Cet équipier est souvent appelé *Scrum Master* (SM), il s'assure de la cohésion de l'équipe et encourage l'équipe à optimiser ses processus de travail.

Enfin, une personne est chargée de faire l'interface avec les clients et parties intéressées (PI). Il facilite la traduction des besoins et assure un rôle d'interface avec le reste de l'organisation. Il est souvent appelé *Product Owner* (PO).

D'autres experts peuvent être ponctuellement appelés par l'équipe de travail, par exemple : *Architect owner*, mentor agile (*agile coach*), expert des métiers du client, spécialiste technique, testeur indépendant, intégrateur.

L'équipe de travail réalise également l'ensemble des tâches de réalisation du projet : conception, développement, contrôle qualité, déploiement et de maîtrise. Les tâches administratives (qui ne contribuent pas à la création de livrables pour les PI) sont réduites au minimum.

La première itération, parfois appelée initiation, permet aussi d'identifier les parties intéressées : utilisateurs principaux et clients, utilisateurs secondaires, participants aux réunions, actionnaires et toutes les personnes ayant un lien indirect avec le projet. Elle permet aussi d'organiser l'espace physique de travail (Rola, Kuchta, et Kopczyk 2016) pour faciliter le travail individuel et collaboratif ainsi que la communication et l'entraide : zone de travail, zone de rencontre de l'équipe, salle de réunion et zone de rencontre des utilisateurs.

2.4 Méthodes agiles

Même si les méthodes permettant de mettre en œuvre les principes agiles ne sont pas normalisées et que par principe elles doivent être choisies et adaptées aux projets, aux compétences et aux souhaits de l'équipe, il est néanmoins possible d'identifier des constantes parmi les méthodes disponibles les plus répandues.

La planification est réalisée lors des revues d'itération qui réunissent l'ensemble des membres de l'équipe avec les clients et utilisateurs. Cela permet de définir et prioriser les fonctions (*features*) qui vont répondre aux besoins des utilisateurs, puis

d'identifier les tâches nécessaires à la réalisation de ces fonctions. Le poids de ces tâches est ensuite évalué et assigné aux membres de l'équipe.

Le contrôle de l'avancement des tâches se fait lors de réunions journalières de coordination (*scrum* ou *stand-up meeting*) qui permettent de confirmer que le travail prévisionnel a pu être réalisé et de discuter d'éventuels problèmes ou blocages qui auraient pu subvenir lors de la réalisation des tâches.

La validation du travail effectué est réalisée directement par les clients, utilisateurs ou *Product Owner* (PO) lors des revues d'itération.

Les méthodes agiles ne nécessitent pas de contrôle des délais et des coûts car le rythme de travail ainsi que l'effectif de l'équipe sont connus (itérations, saisons), le projet s'arrête lorsque tous les besoins des PI sont satisfaits ou bien lorsque le budget alloué est atteint. Une mesure de l'efficacité de l'équipe peut cependant être réalisée (mesure de la quantité de travail par itération, parfois appelée *vélocité*).

Enfin, l'amélioration continue ou adaptation de la méthode se fait lors de revues de retour d'expérience réalisées à la fin de chaque itération.

L'ensemble de ces revues ou réunions sont organisées en suivant des rituels décidés par les équipiers. Ces rituels rythment le projet et permettent à chacun de s'exprimer soit à l'oral, soit par écrit souvent avec des post-its papier ou virtuel.

2.5 Outils agiles

Il existe des dizaines d'outils utilisés dans le cadre des méthodes agiles. Les outils principaux sont basés sur le management visuel pour permettre l'implication de chacun et le travail collaboratif (entre-aide, écoute, coordination). En effet, l'utilisation de canevas facilite l'expression de tous, la communication interpersonnelle et le travail collaboratif. Par ailleurs, l'affichage du travail à réaliser, de l'avancement et de la performance de l'équipe permettent un meilleur engagement. Le canevas emblématique des méthodes agiles est le bac de fourniture (*product backlog*). Ce canevas permet d'identifier et faire mûrir les fonctions attendues du produit (*features*) qui traduisent les besoins des utilisateurs (figure 1). Pour cela, les fonctions sont d'abord proposées dans un bac à sable. La quantité de travail que représente chaque fonction est estimée. La priorité des fonctions est discutée et les fonctions essentielles sont identifiées (*Minimum Viable Product*, MVP) et certaines peuvent être mises en attente dans un bac à glace. Après validation par les PI, les fonctions sont transférées dans un bac d'affinage et affichées sous forme d'un Kanban ordonnant les fonctions. Ensuite, elles peuvent être décomposées en macro-tâches (*epics*) puis en tâches (*user story*).

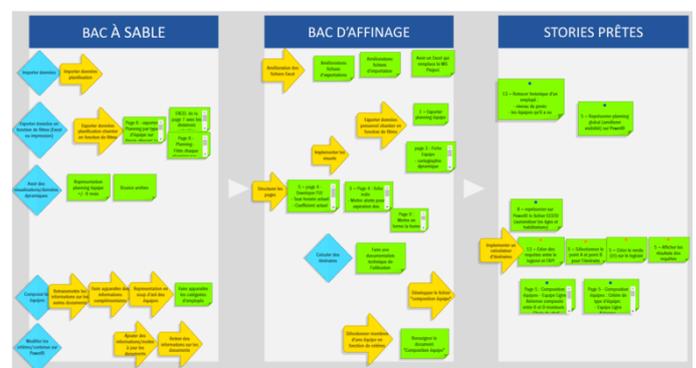


Figure 1. Bac de fourniture (*product backlog*)

Un deuxième canevas est utilisé pour gérer les tâches (*user stories*) dont le poids sera estimé. Cela peut se faire en utilisant

un jeu de cartes avec des poids prédéfinis (*poker planning*). Chaque membre de l'équipe fait sa propre estimation en choisissant une carte, puis les cartes sont mises en commun, discutées avant que l'équipe ne s'accorde sur un poids de la tâche, souvent un nombre de points (*story point*). Ce bac de travail ou de tâches (*sprint backlog*) permet aussi de suivre l'avancement : les tâches sont affichées sur le canevas d'abord dans la réserve de tâches, puis elles passent en statut à faire, en cours, fait, à tester/vérifier et enfin terminée / fait (figure 2). Ces canevas sont principalement utilisés lors des réunions, mais ils restent aussi affichés sur le plateau projet et les membres de l'équipe viennent les consulter et les mettre à jour au fur et à mesure de l'avancement. Cela contribue à la stimulation sociale : lorsqu'un des équipiers se lève de son poste de travail pour faire passer une tâche de « en cours » à « à tester » tous les autres membres le remarque. Les recherches en sociologie ont souligné l'importance de ce phénomène de stimulation sociale dans la participation et l'engagement de chacun. Dans ce contexte, la mise en place d'autocontrôle et de vérification par des membres de l'équipe est aussi généralement préférée à des tests externalisés.



Figure 2. Bac de tâches (*sprint backlog*)

D'autres canevas ou tableaux sont utilisés pour afficher l'avancement physique des tâches sous forme de graphiques : chaque fois qu'une tâche est réalisée, la courbe est incrémentée du nombre de points (*burn-up chart*) (figure 3). C'est parfois l'inverse qui est fait : partant de la quantité de points à réaliser au cours de l'itération, la courbe décroît jusqu'à arriver à zéro (*burn-down chart*).

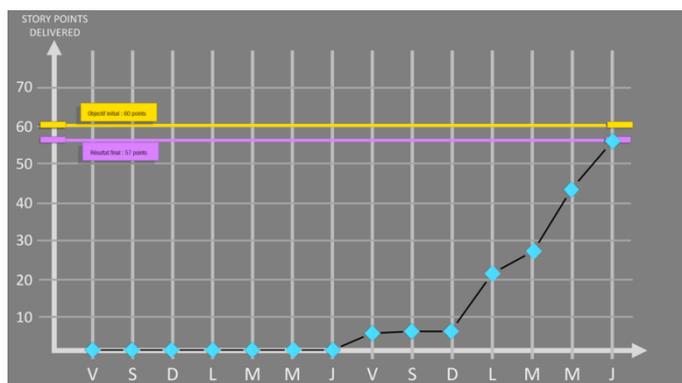


Figure 3. Tableau d'avancement (*burn-up chart*)

Un tableau des obstacles est souvent utilisé pour repérer les difficultés et suivre la progression de leur résolution (figure 4).

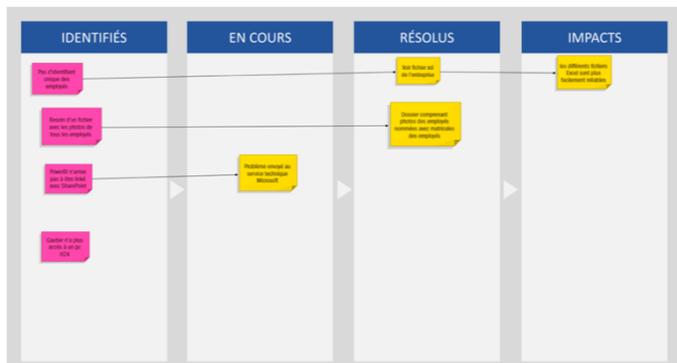


Figure 4. Tableau des obstacles

Enfin, un canevas est utilisé pour le retour d'expérience (REX ou *retrospective*) à la fin de chaque itération. Là aussi, le but est que chacun des équipiers puisse s'exprimer sur les pratiques qu'il veut garder, développer, ajouter ou au contraire réduire voire supprimer (figure 5).

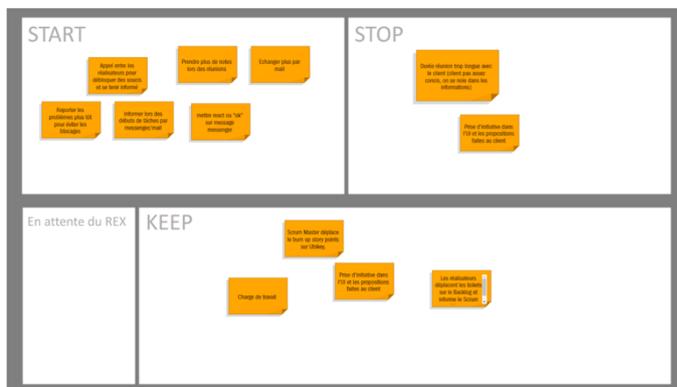


Figure 5. Canevas retour d'expérience (*retrospective*)

3 APPROCHES CLASSIQUE ET AGILE : POINTS COMMUNS ET DIVERGENCES

Avant de discuter les points communs et divergences entre les principes, organisation, méthodes et outils des deux approches, il convient de revenir sur les contextes favorables à leur mise en œuvre et les présupposés ou hypothèses sous-jacents à la mise en place de ces méthodes.

En effet, un des premiers arguments freinant le déploiement des méthodes agiles vient à la fois de la limite de taille des équipes de projet agiles : 10 personnes maximum et des contraintes associées au produit qui doit être conçu : les fonctions doivent être suffisamment indépendantes pour d'une part, être livrées / présentées séparément, et d'autre part avoir de la valeur pour les utilisateurs en dehors des autres fonctions. Ces contraintes importantes devront être dépassées pour permettre la mise œuvre de démarches hybrides.

Cette analyse permet aussi de cibler les types de projets qui bénéficieraient le plus de chacune des approches.

Les grands projets, critiques, ayant des exigences stables et pouvant être formalisées dans un CdC, réalisés dans des entreprises ayant une expertise pour formaliser les procédures de projet et des membres d'équipe projet ayant une culture de l'organisation sachant les mettre en application vont privilégier des approches classiques en étant vigilant à ne pas tomber dans une standardisation excessive nuisant à la flexibilité nécessaire. À contrario, les petits projets mobilisant moins d'une dizaine de personnes, non critiques, avec pas ou peu de sous-traitance, des besoins flous, difficiles à formaliser ou susceptibles d'évoluer

en cours de projet, un produit dont les fonctions sont relativement indépendantes les unes des autres et une équipe stable plutôt composée de personnes d'expérience capables de s'adapter facilement aux changements privilégiera les méthodes agiles.

Au-delà de ces deux extrêmes, pour beaucoup d'industrie, et particulièrement pour les entreprises technologiques qui souhaitent innover, les projets impliquent souvent plus de dix personnes et ils nécessitent de la sous-traitance et des expertises qu'il n'est pas toujours possible de regrouper sur le même site. Aussi, pour ces projets il est souvent difficile de figer les besoins des utilisateurs à l'avance, lors de la phase d'avant-projet. Les recherches sur la technologie et les nouveaux produits ont démontré la co-évolution des besoins avec les solutions (Dorst et Cross 2001; Maher et Tang 2003). Le concept de l'énaction et de la technologie comme anthropologiquement constitutive explicitent les raisons de cette co-évolution (Steiner 2010; Stiegler 1998). Retarder le figage des exigences, réduire le temps de planification amont afin de mieux prendre en compte les résultats de l'exploration des pistes de solutions matérialisées par des maquettes ou prototypes réalisés très en amont du processus de conception devient nécessaire, c'est ce qui est proposé par les méthodes de *design thinking* (Brown 2009; Cross 2008; Plattner, Meinel, et Leifer 2013a, 2013b, 2014). Par ailleurs, les fonctions du produit du projet peuvent être relativement dépendantes les unes des autres.

Tableau 1. Comparaison des principes, organisations, méthodes et outils classiques et agiles.

	Classique	Agile
Contexte et hypothèses	Projets de toutes tailles. Le produit n'a de valeur pour le client que s'il est livré complet du fait de l'interdépendance entre les fonctions. Les besoins sont connus au début du projet et stables pendant la durée du projet.	Petit projet (10 p. max.) Chaque fonction et donc chaque livrable intermédiaire est exploitable indépendamment des autres. Les besoins vont évoluer au cours du temps et en fonction de la présentation des livrables intermédiaires.
Principes	Phase/jalon en fonction des livrables, qui sont « figés » lors des revues avec décisions GO/NO GO. Planification en amont avant de passer à l'action, anticipation sur plusieurs mois, <i>planned-based</i> . Engagement progressif, possibilité théorique d'arrêt du projet à chaque jalon.	Itérations (<i>sprint</i>) à durée fixe et revue d'itération pour présenter les livrables intermédiaires qui peuvent évoluer. Planification à court terme (2/4 semaines), à la demande et tout au long du projet, <i>planning on the fly</i> Décision d'investissement préalable, pivot si nécessaire.
Organisations	MOA/MOE, discussions contractuelles, relation client/fournisseur. Chef de projet nommé par la hiérarchie. Il est impliqué dans le	Client/utilisateurs/PO/SM intégration des parties intéressées à l'équipe. SM coopté par et parmi les développeurs. Possibilité de changer de

	recrutement et l'évaluation des équipiers. Communication formalisée et plutôt centralisée dans l'équipe et avec le client (parties intéressées). Petite équipe dédiée au management du projet (3 à 10 p.) avec des ressources pour les tâches de réalisation dans et à l'extérieur de l'entreprise (environ 100p. pour des projets moyens). Achat, travaux et sous-traitance souvent importante.	SM en cours de projet si l'équipe le décide. Communication horizontale encouragée (même si les PO/SM doivent « protéger » les développeurs). Petite équipe (6 à 10 p. max) dédiée et co-localisée qui réalise à la fois les tâches de réalisation et de maîtrise sans recourir à la sous-traitance. Tâches de management réparties entre les membres de l'équipe.
Méthodes	Les méthodes et procédures de projet sont formalisées et rédigées (parfois imposées par le MOE ou le contrat) et validées par le client. Elles peuvent être assez complexes. Réunions régulières (souvent hebdomadaires) et revues de projet aux passages des jalons. Réunions ad hoc de résolution de problème, analyse des risques, etc.	Les méthodes sont décidées par l'équipe elle-même. Elles sont simples, rarement écrites et évoluent en cours de projet afin de s'adapter aux contextes et REX. Rituels sous forme de réunion journalière « debout », planning d'itération. Contacts horizontaux entre membres de l'équipe et parties intéressées pour le REX et résoudre les problèmes.
Outils	Note de clarification, CdCF, CdC, outils de structuration (PBS, OBS, WBS, PDP). Référentiel du projet. Planning Gantt, PERT, dépenses et temps passés, suivi des fournisseurs et sous-traitants, mesure de l'avancement physique... Audits produit/processus, gestion des non-conformités, amélioration continue. Le REX est généralement réalisé en fin de projet).	Sprint de lancement permettant d'alimenter le product backlog (rédigé sous forme de besoin des utilisateurs) Kanban des backlogs de fourniture et de travail, diagramme des bacs, burn-up ou burn-down charts d'itération, capacité des itérations ... Autocontrôle. Tableau des obstacles ou problèmes géré par le SM. Canevas keep more, keep less, start, stop, mis à jour à chaque itération pour un REX en continu.

Partant de ces premiers constats sur le contexte et les présupposés, nous pouvons maintenant confronter les principes, organisations, méthodes et outils qui en découlent (Tableau 1). Les approches classiques reposent sur le principe de phases et jalons permettant un engagement progressif, mais nécessitant d'une part de figer les besoins assez tôt et d'autre part beaucoup de planification amont. Cela conduit à de nombreuses

vérifications et une gestion des changements en aval pour avoir un produit du projet conforme aux exigences finales. Les conséquences de cette approche sont une forte focalisation sur le contrat et les procédures parfois au détriment des objectifs long terme.

Les organisations classiques ont permis de clarifier les relations maîtrise d'ouvrage / maîtrise d'œuvre et les approches agiles ont permis de renforcer l'implication des PI dans le processus. Cela se fait soit directement, en contrepartie d'un investissement important de la part de parties intéressées, soit en passant par un *Product Owner* (PO). Cette implication des parties intéressées dans le processus permet une conception au juste nécessaire : seules les fonctions utiles et demandées sont réalisées. Au sein de l'équipe projet, le mode de management agile est participatif avec souvent une cooptation de l'animateur ou *Scrum Master* (SM) alors qu'en management de projet classique, le mode de management est choisi par le chef de projet qui est nommé par la direction. Les organisations agiles ne prévoient pas les cas de sous-traitance ou les tailles d'équipe supérieures à 10 personnes. Les méthodes de management classique permettent une capitalisation des procédures de projet alors que les processus agiles sont identifiés sans être nécessairement formalisés par les équipes.

Concernant le temps passé en réunions notamment pour la coordination et la planification, les différentes études réalisées à ce sujet divergent. Plusieurs auteurs évaluent un temps passé pour la planification plus important pour les méthodes agiles du fait l'implication de tous les membres de l'équipe lors de la période d'initiation (*prelude* qui dure souvent une itération complète), lors des réunions de planification d'itération (structuration des *backlog*) et lors des réunions journalières (Serrador et Pinto 2015). Proportionnellement, le temps passé à la planification serait donc moins important dans les approches classiques, car réalisé par quelques experts en impliquant les membres de l'équipe dans des réunions plus ciblées. Il faudrait néanmoins que d'autres études confirment ces résultats en prenant aussi en compte les temps induits, par exemple la gestion des erreurs de planification liées au manque d'implication des acteurs dans le processus ou le temps passé à la gestion des modifications ou changements.

Les outils développés pour les méthodes agiles reprennent sous une forme visuelle et ludique le contenu d'outils déjà présents dans les méthodes classiques. Les canevas, tableaux de bord et indicateurs définis et mis à jour collectivement puis affichés à la vue de tous permettent à chacun d'identifier les fonctions du produit, les priorités, les tâches à réaliser et la progression du travail. Ces canevas favorisent un travail collaboratif et une vision partagée par les membres de l'équipe.

Les rituels tels que les réunions journalières « debout » (*stand-up daily scrum meeting*) donnent un rythme au projet et permettent la coordination et le suivi de l'avancement des tâches. Les réunions de retour d'expérience à la fin de chaque itération permettant une amélioration en continu aussi préconisées dans les méthodes classiques sans que cela soit toujours réalisé.

4 PROPOSITION D'UNE APPROCHE HYBRIDE

Ce qui a été exposé permet de mettre en évidence les contextes favorables à la mise en œuvre des deux approches ainsi que les avantages de chacune des méthodes.

Plusieurs recherches (Cooper et Sommer 2016) (Karlström et Runeson 2006) (Sommer et al. 2015) (Könnölä et al. 2016) (Conforto et Amaral 2016) (Heimicke et al. 2019) (Olszewska (née Płaska) et al. 2016) ont permis de mettre en évidence les avantages des approches hybrides : meilleure communication

interne, management visuel qui permet de mieux prendre conscience de l'avancement du projet, l'implication du client dans la planification permet plus d'efficacité, l'amélioration de l'implication et du moral de l'équipe.

Mais aussi les limitations de ces approches hybrides : difficulté d'acceptation par les clients de ces nouvelles modalités (« prise de risque »), limitations liées à la taille des projets, difficultés de communication / capitalisation avec le reste de l'entreprise (équipe en autarcie), difficulté à conserver des équipes dédiées sur la durée, prolifération du nombre de réunions.

L'analyse de ces éléments permet de proposer des orientations pour une intégration des approches classiques et agiles pour le management des projets industriels de moyenne ou grande dimension (Tableau 2).

4.1 Contexte des approches hybrides

Pour justifier le déploiement d'approches hybrides, certains éléments de contexte sont nécessaires. Nous avons déjà évoqué la taille des projets. La possibilité de décomposer le produit du projet en fonctions ou disciplines relativement indépendantes doit aussi être validée pour permettre de répartir le travail entre les sous-équipes et d'identifier des livrables intermédiaires à présenter au client. Le client (MOA/PI) doit enfin pouvoir s'impliquer fortement dans le processus de pilotage du projet : priorisation des livrables, (ré-)orientation du projet, évaluation et validations intermédiaires.

Par ailleurs, le déploiement de ce type de démarche présente un intérêt si les besoins sont difficiles à formaliser : le cahier des charges n'est pas figé, il doit pouvoir être précisé progressivement au fur et à mesure de l'avancement du projet.

4.2 Principes des approches hybrides

Ce qui ressort des études, c'est que le principe cardinal et différenciant des méthodes agiles, qui semble être la base des approches hybrides, est l'itération (*sprint*) de 2 à 4 semaines. Cela permet un micro-planning à la demande en trouvant un compromis entre les demandes des clients, des équipes et les contraintes de coordination entre sous-équipes.

Ces itérations sont combinées avec un phasage traditionnel pour assurer le macro-planning qui permet de donner une visibilité sur le long terme à l'équipe, aux clients et au management du réalisateur. La micro-planification est réalisée au sein d'une phase en identifiant les livrables intermédiaires qui seront présentés à chaque itération.

Le référentiel préliminaire permet de définir le cadre général du projet (macro-planning). Ce référentiel de coordination est décidé avec les sous-équipes (ou leurs représentants). La structuration de détail est décidée progressivement par les sous-équipes. La planification des lots de travaux est faite par les sous-équipes à la demande.

4.3 Organisation pour les approches hybrides

Les deux approches préconisent, lorsque c'est possible, d'avoir des équipes de personnes dédiées au projet, stabilisées pour la durée du projet et regroupées sur un plateau pour faciliter la cohésion et la communication. Pour gérer les projets de taille moyenne ou grande (équipe entre 10 et 100 personnes), le projet est décomposé en fonctions ou disciplines suffisamment autonomes et d'une taille permettant leur gestion par une sous-équipe en mode agile. Chaque équipe réalise toutes les tâches depuis la conception jusqu'au suivi de la réalisation y compris les contrôles (qualité, risques, coûts, délais). Il est nécessaire que les équipes ainsi constituées soient chacune d'une taille inférieure ou égale à 10 personnes. Il est essentiel de veiller à préserver l'autonomie des sous-équipes (principe de subsidiarité) : laisser le niveau de décision, en particulier le

choix des méthodes, au plus près du terrain. C'est un gage de motivation et d'efficacité.

Tableau 2. Proposition des principes, organisation, méthodes et outils d'une approche hybride.

<p>Contexte et hypothèses : Petit ou moyen projet industriel avec un produit pouvant être décomposé en fonctions relativement indépendantes. Grands projets si décomposables en sous-projets. Disponibilité MOA et PI pour s'impliquer dans le projet. Besoin difficile à figer : CdCF précisé au cours du projet.</p>
<p>Principes : Combinaison d'un phasage classique pour le projet (macro-planning) et d'itérations agiles (micro-planning à la demande). Revues des livrables intermédiaires à chaque itération et rituels pour le contrôle de l'avancement et la gestion des problèmes. Structuration en fonctions principales avec découpage et distribution du travail à plusieurs sous-équipes.</p>
<p>Organisation : Équipiers à temps plein pour la durée du projet, équipes localisées pour faciliter la communication et avec une autonomie maximum (principe de subsidiarité) et réalisant toutes les tâches depuis la conception jusqu'au contrôle qualité en passant par la réalisation. Pilotage (coordination) des sous-équipes par un chef (directeur) de projet. Relations MOA-MOE classiques, mais implication forte des clients/utilisateurs dans le processus de décision des sous-équipes, si nécessaire en nommant un PO. Les fournisseurs et travaux sont gérés par chaque sous-équipe.</p>
<p>Méthodes : Référentiel du projet et maîtrise (risques, coûts et délai) assurés au niveau des sous-équipes. Consolidation des indicateurs de chaque équipe au niveau global du projet par le CP. Processus choisis par les équipiers et adaptés aux particularités du projet. REX à la fin de chaque itération. Réduction au minimum les tâches contractuelles et administratives. Tâches de réalisation des livrables privilégiées.</p>
<p>Outils : Outils de management visuel permettant de mieux comprendre et partager les objectifs, le travail à faire, et l'avancement ; ils contribuent à une meilleure implication, engagement et motivation de l'équipe.</p>

Les projets composés de plusieurs sous-équipes devront disposer d'une structure de coordination. Le chef de projet aura alors pour mission de coordonner ces sous-équipes. L'équipe de coordination pouvant être composée de représentants de chaque sous-équipe. Les décisions prises au niveau de la coordination, même si elles sont partagées avec les sous-équipes, sont limitées à celles qui sont indispensables à la coordination générale.

L'organisation mise en place pour piloter le projet reste une question ouverte. Les études disponibles montrent que plusieurs types d'organisation sont possibles. Certains conservent un chef de projet nommé même s'il applique un management participatif. D'autres préfèrent utiliser les rôles préconisés dans les méthodes agiles. Dans ce cas, il n'y a plus de chef de projet, mais un animateur ou facilitateur (*Scrum Master*, SM) identifié parmi les membres des sous-équipes pour animer, coordonner et coacher les sous-équipes et aussi en charge de résoudre les problèmes d'interface. C'est généralement une personne compétente dans la discipline principale du projet et formée aux méthodes agiles qui assume ce rôle.

Les relations MOA-MOE restent classiques, mais une implication forte des clients/utilisateurs dans le processus de décision est nécessaire pour orienter le projet, évaluer et valider les livrables intermédiaires. Une personne est aussi souvent identifiée pour servir d'interface avec les clients ou utilisateurs et l'équipe (*Product Owner*, PO). Son rôle est de faire accoucher les besoins, de les traduire et/ou de les formaliser. Le cas échéant, ces deux rôles peuvent être assumés par le chef de projet.

Les fournisseurs et travaux sont gérés par chaque sous-équipe, mais, s'il y a plusieurs sous-équipes réalisant des achats, une coordination semble nécessaire.

Le problème qui se pose est la coordination de cet ensemble en préservant l'autonomie des sous-équipes, en minimisant le nombre de niveaux hiérarchiques de management ainsi que les coûts d'administration. Il est aussi parfois compliqué d'estimer les prévisions à terminaison du projet avec ce mode de fonctionnement.

4.4 Méthodes pour les approches hybrides

Le référentiel du projet et la maîtrise (risques, coûts et délai) sont assurés au niveau de l'équipe. S'il y a plusieurs équipes, il sera nécessaire de consolider les indicateurs de chaque équipe au niveau global du projet. Un poste de coordinateur maîtrise auprès du CP peut être créé pour cela. Le leader (chef ou directeur, SM ou PO) du projet se soucie de réduire au minimum les tâches contractuelles et administratives. Il privilégie les tâches réalisant des livrables.

Les processus et procédures de travail sont choisis par les équipiers au début du projet : ils peuvent être adaptés aux particularités du projet. Un retour d'expérience est réalisé à la fin de chaque itération pour, entre autres, faire évoluer les méthodes de travail. Ces processus peuvent être capitalisés pour les projets suivants. Dans les projets industriels, les équipes sont pluridisciplinaires, c'est-à-dire que chaque équipier traite d'une discipline différente : le retour d'expérience de l'équipe ne peut donc pas être technique. La question est différente si les équipes sont divisées par discipline, dans ce cas un REX technique peut être réalisé en plus du REX méthodologique.

La maîtrise des contrats forfaitaires ou financièrement contraints exige que le CdCF soit figé dès le départ : ce type de maîtrise ne peut donc pas être mis en œuvre pour les approches hybrides. Cependant il est possible d'utiliser les méthodes et outils de maîtrise des coûts et délais en s'appuyant sur les planifications macro et micro présentées précédemment. Par exemple, la maîtrise du budget interne et du délai peut être assurée en précisant à l'avance le nombre d'itérations par phase pour éviter les dérives. S'il doit y avoir de la sous-traitance ou des achats, un système classique de maîtrise des coûts et délais spécifiques peut être mis en œuvre. Comme il ne sait pas définir précisément ses besoins, le client accepte le risque de ne pas avoir un engagement ferme sur les coûts et les délais. Ce risque est maîtrisé par son implication forte à chaque revue d'itération. Dans ce contexte, la décision de réaliser le projet est généralement prise à l'avance et le principe d'engagement progressif est abandonné. Le principe adopté est celui de la réorientation du projet (pivot) plutôt que l'arrêt du projet en cas de problème.

4.5 Outils pour les approches hybrides

Le management visuel basé sur l'utilisation de canevas support à l'auto-contrôle du travail au quotidien et à la résolution des problèmes en équipe couplés aux rituels agiles tels que les *daily scrum meeting* ont un rôle clé dans la motivation et l'engagement des équipiers. Plusieurs études montrent que

l'introduction de cet état d'esprit agile dans un grand projet classique permet de mieux identifier les priorités ce qui conduit à améliorer la communication et le moral, la réactivité et la flexibilité et permet un gain productivité et de temps (Gren, Torkar, et Feldt 2017).

Il est essentiel d'impliquer tous les membres de l'équipe ainsi que les parties intéressées dont les clients et utilisateurs. Cela se fait en particulier lors des revues d'itération qui permettent de bien comprendre, identifier, caractériser et prioriser les besoins en s'appuyant sur des livrables intermédiaires les plus concrets possibles : maquette, prototype, démonstrateurs, simulations, etc. L'implication active des clients permet d'avoir un retour des utilisateurs, l'ajustement des besoins en fonction des contraintes de conception, une adaptation de la demande... Les livrables à présenter lors de ces revues d'itération doivent être précisément discutés, identifiés et caractérisés : faut-il d'abord privilégier une maquette d'aspect ou une maquette fonctionnelle ... Dans les approches hybrides, ces revues permettent généralement d'élargir la vision principalement centrée sur le « produit » des méthodes agiles (avec des livrables orientés fonctions, concepts, technologie...) à celle du « projet » en identifiant des états intermédiaires de représentation du projet (connaissances, risques, coûts, marché...).

Le principe d'itérations ou phase à durée fixe avec présentation de livrables intermédiaires est aussi la clé de l'implication, l'engagement et la motivation des membres de l'équipe. En effet, les livrables intermédiaires n'étant pas finalisés, cela facilite leur évolution en créant une dynamique de projet prenant en compte les dernières avancées et les retours des clients/utilisateurs.

5 CONCLUSION ET PERSPECTIVES

Les études réalisées sur la mise en place de méthodes hybrides semblent confirmer l'intérêt de cette combinaison dans le contexte de petits ou moyens projets industriels dont les fonctionnalités sont relativement indépendantes.

Cependant, il faudrait poursuivre les recherches afin de préciser d'une part les contextes favorables à ce type de démarches hybrides : taille et criticité des projets, type de produits, type de contrat et de relations client/réalisateur.

D'autre part, il faudrait préciser les effets et impacts associés aux principes, organisations, méthodes et outils choisis dans un contexte donné. Si l'importance du principe d'itération ou phasage à durée fixe semble aujourd'hui reconnu, de même que l'organisation en plateau projet avec des équipes stables et à temps plein sur le projet utilisant des outils de management visuel et des rituels favorisant l'implication des parties prenantes et la motivation des équipes, des questions subsistent sur certains aspects :

- Quel dosage entre planification classique et planification itérative ?
- Quels livrables intermédiaires présenter en fin d'itération (produit, projet...)? Quels sont les livrables intermédiaires qui facilitent les validations des besoins par les clients et utilisateurs ?
- Quel équilibre entre implication de tous les acteurs dans tous les processus de décision et multiplication des réunions ?
- Quels organisations et rôles privilégier ? SM et PO ou Cdp ?
- Comment assurer la coordination de plusieurs équipes agiles (dans le cadre des projets industriels de taille moyenne) en préservant les principes essentiels agiles et sans créer une administration pléthorique ?
- Faut-il développer de nouveaux canevas ? En particulier, en plus du bac produit, un bac projet pour identifier les risques, le marché...

-Comment gérer les approvisionnements externes et les parties sous-traitées des projets ?

- Comment décider de lancer les projets et pourquoi un projet agile ?

Enfin, au-delà des limitations déjà identifiées liées aux contextes d'application, il faut noter que le succès de ces approches nécessite des changements d'état d'esprit importants, aussi bien au niveau des clients et utilisateurs que du management et des équipiers. Ces changements peuvent être plus ou moins bien acceptés en fonction des cultures d'entreprise et des habitudes des secteurs d'activité (Javdani Gandomani et Ziaei Nafchi 2016) (Dikert, Paasivaara, et Lassenius 2016). C'est aussi un élément qui nécessite des recherches complémentaires afin de l'expliquer et de le quantifier pour mieux cibler d'une part les entreprises et secteurs propices à la mise en place de ces méthodes et d'autre part les actions de conduite du changement qui peuvent être entreprises pour faciliter l'adoption de cet état d'esprit.

6 REFERENCES

- Beck, Kent, Mike Beedle, Arie van Bennekum, Alistair Cockburn, Ward Cunningham, Martin Fowler, James Grenning, Jim Highsmith, Andrew Hunt, Ron Jeffries, Jon Kern, Brian Marick, Robert Martin, Steve Mellor, Ken Schwaber, Jeff Sutherland, et Dave Thomas. 2001. « Manifesto for Agile Software Development ». Consulté 1 novembre 2020 (<http://agilemanifesto.org/>).
- Bianchi, Mattia, Giacomo Marzi, et Massimiliano Guerini. 2020. « Agile, Stage-Gate and Their Combination: Exploring How They Relate to Performance in Software Development ». *Journal of Business Research* 110:538-53.
- Brown, Tim. 2009. *Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation*. New York: HarperBusiness.
- Campanelli, Amadeu Silveira, et Fernando Silva Parreiras. 2015. « Agile Methods Tailoring – A Systematic Literature Review ». *Journal of Systems and Software* 110:85-100.
- Chvidchenko, I., et J. Chevalier. 1997. *Conduite et Gestion de Projets*. Cépaduès Editions. Toulouse, France.
- Clark, Kim B., R. H. Hates, et Steven C. Wheelwright. 1988. *Dynamic Manufacturing, Creating the Learning Organization*. The Free Press. New York.
- Conforto, Edivandro C., et Daniel C. Amaral. 2016. « Agile Project Management and Stage-Gate Model—A Hybrid Framework for Technology-Based Companies ». *Journal of Engineering and Technology Management* 40:1-14.
- Cooper, Robert G., Scott J. Edgett, et Elko J. Kleinschmidt. 1999. « New product portfolio management: practices and performance ». *Journal of Product Innovation Management* 16(4):333-51.
- Cooper, Robert G., et Anita F. Sommer. 2016. « Agile-Stage-Gate: New Idea-to-Launch Method for Manufactured New Products Is Faster, More Responsive ». *Industrial Marketing Management* 59:167-80.
- Courtot, Hervé. 1998. *La gestion des risques dans les projets*. Economica. Paris, France.
- Cross, Nigel. 2008. *Engineering design methods: strategies for product design*. Chichester, England: John Wiley & Sons.
- Cusumano, Michael A., et Kentaro Nobeoka. 1999. *Le management multiprojets. Optimiser le développement de produits*. Dunod. Paris, France.
- Denning, Stephen. 2018. « The Challenge of Leadership in the Age of Agile ». *Leader to Leader* 2018(89):20-25.
- Dikert, Kim, Maria Paasivaara, et Casper Lassenius. 2016. « Challenges and Success Factors for Large-Scale Agile

- Transformations: A Systematic Literature Review ». *Journal of Systems and Software* 119:87-108.
- Dorst, Kees, et Nigel Cross. 2001. « Creativity in the design process: co-evolution of problem–solution ». *Design Studies* 22(5):425-37.
- Fernez-Walch, Sandrine, et François Romon. 2013. *Management de l'innovation, de la stratégie aux projets*. Vuibert. Paris, France.
- Garel, Gilles. 2011. *Le management de projet*. Éditions la Découverte. Paris, France.
- Garel, Gilles, Vincent Giard, et Christophe Midler. 2004. *Faire de la recherche en management de projet*. Vuibert. Paris.
- Giard, Vincent. 1991. *Gestion de projets*. Economica. Paris, France.
- Giard, Vincent, et Christophe Midler. 1993. *Pilotages de projet et entreprises*. Economica. Paris.
- Gidel, Thierry, et William Zonghero. 2006. *Management de projet, vol.2, Approfondissements*. Vol. 2. Paris: Hermes Science/Lavoisier.
- Gidel, Thierry, et William Zonghero. 2020. *Management de projet, vol.1, Introduction et fondamentaux*. Vol. 1. 2ème. Paris: Hermes/Lavoisier.
- Gren, Lucas, Richard Torkar, et Robert Feldt. 2017. « Group Development and Group Maturity When Building Agile Teams: A Qualitative and Quantitative Investigation at Eight Large Companies ». *Journal of Systems and Software* 124:104-19.
- Heimicke, J., M. Niever, V. Zimmermann, M. Klippert, F. Marthaler, et A. Albers. 2019. « Comparison of existing agile approaches in the context of mechatronic system development: Potentials and limits in implementation ». P. 2199-2208 in Vol. 2019-August.
- Hoda, Rashina, Norsaremah Salleh, John Grundy, et Hui Mien Tee. 2017. « Systematic Literature Reviews in Agile Software Development: A Tertiary Study ». *Information and Software Technology* 85:60-70.
- Javdani Gandomani, Taghi, et Mina Ziaei Nafchi. 2016. « Agile Transition and Adoption Human-Related Challenges and Issues: A Grounded Theory Approach ». *Computers in Human Behavior* 62:257-66.
- Jolivet, F. 2003. *Manager l'entreprise par projets. Les métarègles du management par projet*. éditions EMS. Paris, France.
- Karlström, Daniel, et Per Runeson. 2006. « Integrating Agile Software Development into Stage-Gate Managed Product Development ». *Empirical Software Engineering* 11(2):203-25.
- Kerzner, Harold. 2017. *Project Management, a Systems Approach to Planning, Scheduling and Controlling*. John Wiley and Sons. New York.
- Könnölä, Kaisa, Samuli Suomi, Tuomas Mäkilä, Tero Jokela, Ville Rantala, et Teijo Lehtonen. 2016. « Agile Methods in Embedded System Development: Multiple-Case Study of Three Industrial Cases ». *Journal of Systems and Software* 118:134-50.
- Lenfle, Sylvain, et Christophe Loch. 2010. « Lost Roots: How Project Management Came to Emphasize Control over Flexibility and Novelty ». *California Management Review*.
- Maher, Mary, et Hsien-Hui Tang. 2003. « Co-Evolution as a Computational and Cognitive Model of Design ». *Research in Engineering Design* 14(1):47-64.
- Olszewska (née Płaska), Marta, Jeanette Heidenberg, Max Weijola, Kirsi Mikkonen, et Ivan Porres. 2016. « Quantitatively Measuring a Large-Scale Agile Transformation ». *Journal of Systems and Software* 117:258-73.
- Plattner, Hasso, Christoph Meinel, et Larry Leifer. 2013a. *Design Thinking Research: Building Innovation Ecosystems*. New York: Springer International Publishing AG.
- Plattner, Hasso, Christoph Meinel, et Larry Leifer. 2013b. *Design Thinking: Understand - Improve - Apply*. Berlin; Heidelberg u.a.: Springer-Verlag Berlin and Heidelberg GmbH & Co. K.
- Plattner, Hasso, Christoph Meinel, et Larry Leifer. 2014. *Design Thinking Research: Building Innovators*. Springer International Publishing AG.
- PMI. 2017. *PMBOK® Guide*. Project Management Institute; Sixth Edition.
- Rola, Paweł, Dorota Kuchta, et Dominika Kopczyk. 2016. « Conceptual Model of Working Space for Agile (Scrum) Project Team ». *Journal of Systems and Software* 118:49-63.
- Serrador, Pedro, et Jeffrey K. Pinto. 2015. « Does Agile Work? — A Quantitative Analysis of Agile Project Success ». *International Journal of Project Management* 33(5):1040-51.
- Sommer, A. F., C. Hedegaard, I. Dukovska-Popovska, et K. Steger-Jensen. 2015. « Improved product development performance through Agile/Stage-Gate hybrids—The next-generation Stage-Gate process? » *Research-Technology Management* 158(1):1-10.
- Steiner, Pierre. 2010. « Philosophie, technologie et cognition. Etats des lieux et perspectives ». *Intellectica* 53(1):7-40.
- Stiegler, Bernard. 1998. *Technics and Time: The Fault of Epimetheus*. Stanford University Press.
- =====
- [ISO 21500 : 2012] ISO 21500 : 2012, Lignes directrices sur le management de projet, 2012.
- [NF EN 9200 : 2005] NF EN 9200, Série aérospatiale - Management de programme - Recommandation pour une spécification de management de projet, AFNOR, 2005.

What are the origins of eco-innovations? Results from a case based reasoning approach on 300 eco-innovations

OLIVIER PIALOT¹, DOMINIQUE MILLET²

¹ QUARTZ

Supmeca Toulon, 83000 Toulon, France
olivier.pialot@supmeca.fr

² COSMER

SeaTech-UTLN, 83000 Toulon, France
dominique.millet@univ-tln.fr

Résumé – Les conséquences environnementales de la consommation de masse obligent à éco-innover, ce qui signifie repenser complètement notre façon de concevoir, fabriquer et consommer en proposant sur le marché des produits et services à forte ambition environnementale. Afin d'aider les entreprises dans les premières phases de conception de l'éco-innovation, cet article explore les origines de l'éco-innovation avec une approche de raisonnement par cas basée sur 300 cas. Grâce à une expérimentation mobilisant 8 experts en éco-conception, 5 problèmes de départ et 12 principes d'éco-innovation pour résoudre ces problèmes de départ sont identifiés. Les résultats montrent la contribution des méthodes de conception pour éco-innover dans une vue d'ensemble des origines de l'éco-innovation, en particulier celle d'un ensemble de 7 méso-mécanismes de stimulation de l'éco-idéation, dénommés ESM.

Abstract – The environmental consequences of mass consumption require to eco-innovate, which means completely rethinking our way of designing, manufacturing and consuming by proposing on the market products and services with a high environmental ambition. In order to help companies in early design phases of eco-innovation, this article investigates the origins of eco-innovation with a case based reasoning approach based on 300 cases. Thanks to an experiment mobilizing 8 ecodesign experts, 5 Starting Problems and 12 Eco-innovation principles to solve these starting problems are identified. Results show the contribution of design methods to eco-innovate compared to the eco-innovation origins identified, in particular the potential input of a set of 7 Eco-ideation Stimulation Meso-mechanisms, named ESM.

Mots clés - éco-innovation, raisonnement par cas, phases amont de conception.

Keywords – eco-innovation, case based reasoning, early design phases.

1 INTRODUCTION- CONTEXT

Our society is increasingly concerned with the environmental consequences of mass consumption. Resource use has more than tripled since 1970, and continues to grow. Models based on a propensity to consume and throw away are having devastating effects on our planet. It is established that 90 percent of biodiversity loss and water stress is due to resource extraction and processing. These same activities contribute to nearly half of the world's greenhouse gas emissions. [United Nations, 2019] Further than eco-designing approach, an answer to remedy to this problem is to eco-innovate, which means completely rethinking our way of designing, manufacturing and consuming by proposing on the market products and services with a high environmental ambition. Despite the increasing interest in eco-innovation noticed in institutions and academia in the past few years [Díaz-García et al., 2015; Hojnik & Ruzzier, 2016], its dissemination in companies remains limited. A recent survey in a small panel of French companies underlined that eco-innovation is still ambiguous for industrial practitioners, and therefore they cannot identify examples of eco-innovative products [Cluzel et al., 2014]. In fact, eco-innovation means to design a new

system integrating the sustainable development dimensions (environment, social, technology, stakeholders). This multi-dimensional aspect makes it more complex to characterize and to design.

To remedy to this, the aim of the ALIENNOR research project is to develop eco-ideation mechanisms [Tyl et al., 2016], which are intended to be a synthesis of recent research works in eco-innovation. This project also includes the creation of a database of 300 eco-innovation cases. This base of concrete cases offers the opportunity to question the real origins of eco-innovations with an analysis operating in a bottom-up way, following the approach of case based reasoning. Then it is possible to compare the tools proposed and the reality, in order to validate or put into perspective the axes pushed by eco-innovation methodologies, in particular the ESMs. This is the purpose of this paper.

This article starts with a state of the arts of eco-innovation definitions, strategies and tools, in particular the Eco-innovation Stimulation Meso-mechanisms (ESM), and with the results of a workshop about eco-innovation operable definition (section 2). In section 3, an experiment based on a case based reasoning approach is presented to investigate Starting

problem to eco-innovate, and eco-innovation principles. Results showing origins of eco-innovation are presented in section 4, and are discussed in section 5 particularly in front of the ESMs set proposed.

2 STATE OF THE ARTS

2.1 State of the art on eco-innovation definition

To select 300 cases of eco-innovations, the question of the attributes defining an eco-innovation arises.

Diaz-Garcia et al. summarized 8 different definitions of eco-innovation appearing in key studies between 1996 and 2013 [Díaz-García et al., 2015]. It concerns a new product or service which significantly reduces the environmental impacts all along its life cycle. Since the first definition was given by Fussler and James [Fussler & James, 1996], the concept has drifted from a product/service to a potentially more organizational focus; from a purely environmental to a mixed environmental, social and even institutional contribution [Mathieu et al., 2015]. Mathieu et al. add that eco-innovation creates positive externalities on one or several dimensions of sustainable development [Mathieu et al., 2015].

With O'Hare and Mc Aloone [O'Hare & McAlloone, 2014], the concept of eco-innovation is discussed with regard to three anchoring domains: engineering design; strategy and management; environmental science. The recent contribution of the UNEP guide emphasizes the importance of the business model issue associated with eco-innovation [O'Hare et al., 2014]. In a practical way, Bocken et al. [Bocken et al., 2014] unify bodies of knowledge into eight sustainable business model archetypes. Rennings shows that eco-innovation can be technological, organizational, social or institutional, and developed by a wide range of stakeholders, from companies to NGOs [Rennings, 2000].

Diaz-Garcia et al. [Díaz-García et al., 2015] stress that the focus of eco-innovation is either on the effect (i.e. the contribution to environmental improvement), on the motivation (the goal of sustainable development) or both. For Mathieu et al. [Mathieu et al., 2015], eco-innovation may be intentional or not, as it is the contribution (and not the objective) that has to be relevant to sustainable development. This means that eco-innovation may be appreciated ex post, after launch to market.

These elements describe the evolutions of eco-innovation definition taking into account different dimensions, but don't give really actionable criteria to characterize what an eco-innovative system is.

Concerning the criteria of eco-innovation, in addition to the originality or feasibility of ideas, an environmental criterion is necessarily required. This characterization is often associated with a quantitative evaluation, such as life cycle analysis (LCA) or simplified LCA [Hunt et al., 1998]. Bocken et al. [Bocken et al., 2012] emphasize that most environmental assessment tools are intended for the downstream design phases. So the characterization of eco-innovative ideas becomes more critical, subjective and uncertain [Vallet et al., 2013]. In other words, at the end of the eco-ideation process in the early design phases, there is a difficulty to assess the environmental potential. The system is not well-defined yet and it is not possible to do an LCA analysis due to the lack of information. When it is necessary to characterize an eco-innovative project, there is the same problem: the available information about these eco-innovation cases are often just described with outlines, performance claims, interests and functionalities in use, a new technology, etc.

Specific research on the evaluation of ideas in eco-innovation is rare. From a qualitative experiment based on focus groups and on a case based reasoning approach emerged the double intent of "environmental gain" and "mass effect" to characterize an eco-innovation [Pialot & Millet, 2018]. The authors propose to transform these characterization elements of eco-innovation used a posteriori into two criteria that can be used from the early design phases of design:

Criterion Environmental Potential: To achieve large-scale environmental gain, two pathways seem possible (Table 1). What can be considered:

- "direct" environmental gains generated by a change of conceptual models in the design of the existing system (disruptive elements bringing a gain on one or more phases of the lifecycle without changing the functional set of system)
- "indirect" environmental gains by a transition towards a new system/mode of consumption with softer environmental impact considering a new functional set as reference (democratization of a new system that challenge a more impacting consumption pattern, pooling of objects etc.) or with a change of consumers behaviour (more parsimonious usages, local consumption, increasing of the product lifetime due to a better maintenance etc.). For example, an electric bicycle with protection against rain and accidents provides an alternative to private car use, especially in the city. There is no direct environmental gain due to the added modules integrated but it is obtained a potential strong indirect environmental gain.

Table 1. Environmental Potential Criterion illustrated on the bicycle system

Case	Type of gain	"direct" environmental gains	"indirect" environmental gains
Bicycle in bio-materials		Use of « green » Materials	NO
Electric bicycle with protection against rain and accidents		NO	Democratization of biking (soft mobility strategy in the city compared to a private car use)
Upgradable bicycle with system to rent it if not used		Rationalization of the materials use over time (due to upgradability)	Change of consumption behaviour with pooling of objects

If there is no environmental gain (neither direct nor indirect) easy to identify, the idea of eco-innovation is stopped. This aspect can be characterized very early.

Then, the "mass effect" underlies in fine a commercial success of the future eco-innovative product and no environmental drift. To achieve this status, it is proposed to check during the early design phases if the new eco-innovative concept satisfies certain requirements grouped under the name "viability of diffusion".

Criterion Viability of diffusion: it is proposed to check if the new eco-innovative concept with all the changes it generates can reach a satisfactory level of viability, in terms of technical feasibility (technical viability), attractiveness of the value proposition -i.e. value for the client / price- (attractiveness viability), stakeholder satisfaction (stakeholders viability), and potential negative rebound effects (environmental viability). The authors consider that if these four viabilities are satisfied, it is then possible to speak of eco-innovation. If not, it can be a project / concept of eco-innovation.

2.2 State of the art on eco-innovation tools

In the literature, creativity in eco-innovation is widely considered as critical. Consequently, some researchers have analysed how to support eco-ideation courses, through the development of the specific eco-ideation tools.

Eco-ideation sessions were first supported by diagrams or radars, such as the LiDS Wheel [Brezet & Van Hemel, 1997] or the Eco-Compass [Fussler & James, 1996]. Cluzel et al. [Cluzel et al., 2016] presents the eco-design strategy wheel from [Brezet and Van Hemel, 1997] to generate eco-innovative ideas. The creative operation roughly consists in performing a brainstorming session on each axis of the diagram or the wheel (Table 2). These tools use macro level mechanisms of eco-ideation, that is to say innovative axes to explore are proposed but without developed methodological guidelines.

Table 2. Presentation of the strategies of the LiDS wheel [Brezet & Van Hemel, 1997]

Lifecycle Design Strategies (LiDS)
LiDS 0: New concept development
LiDS 1: Selection of low-impact materials
LiDS 2: Reduction of materials usages
LiDS 3: Optimisation of production techniques
LiDS 4: Optimisation of distribution system
LiDS 5: Reduction of impact during use
LiDS 6: Optimisation of initial lifetime
LiDS 7: Optimisation of end-of-life system

A wide part of the literature on eco-ideation methods and tools is based on TRIZ methodology. TRIZ is a systematic creative method to solve design contradictions [Altshuller, 1998; Kobayashi, 2006], but also mixed with biological patterns [Bogatyrev & Bogatyreva, 2014]. These tools use micro level mechanisms of eco-ideation because they need precise data to be used.

Tyl proposes to use "meso" mechanisms to support eco-ideation sessions, to efficiently stimulate the design team during the whole eco-innovative process with tools at the same time generic, that require no precise data, but didactic with a methodological guide [Tyl et al., 2016]. Tyl [Tyl et al., 2014] developed the EcoASIT tool, adaptation of the ASIT tool for eco-innovation. In line with recent developments in eco-innovation to consider business model innovation as a way to generate sustainable ideas, the Value Mapping Tool proposes to cover the different values for key stakeholders and to transform them from missing or destroyed values into opportunities [Bocken et al., 2013].

In order to eco-innovate, the exploration field of the designer goes beyond the scope of the product, by integrating in system design the sustainable development dimensions (environment, social, technology, stakeholders). Building on the research work of UNEP approach [O'Hare et al., 2014], the 8 archetypes of sustainable business models [Bocken et al., 2014], the adaptation of the Business Model Canvas [Osterwalder & Pigneur, 2010] for sustainability by Joyce [Joyce et al., 2015], etc., Tyl proposes an original set of Eco-ideation Stimulation Meso-mechanisms, named ESM. This set of ESMs leads to an exploration of systemic dimensions related to the sustainable development, still under-exploited, but very promising (locality, stakeholders, etc.) [Tyl et al., 2016]. After trials and in order to have a homogeneous content, the original 8 ESMs set has been reviewed and restricted to 7 (the concept of rebound effects will be ultimately used in evaluation of ideas). Table 3 describes each of these ESMs.

These ESMs seem to deliver ideas that are sufficiently well defined and cover the sustainable development dimensions as they appear in the literature.

In order to validate the completeness of the ESMs set or to position it on a wider field, it is necessary to explore the

different possible origins of eco-innovations in another way. The approach proposed by this article is to analyse 300 cases of eco-innovation.

Table 3. Presentation of the set of 7 ESMs [Tyl et al., 2016]

ESM	Justification
ESM1: Innovate through value creation considering all stakeholders	This ESM explores the value creation for all stakeholders (customers, business, environment and society).
ESM2: Innovate through biomimicry	This ESM explores natural strategies of development at several system levels (organ, organism, ecosystem).
ESM3: Innovate through end-user and sustainable uses	This ESM explores the issues of frugality, eco-usage and energy efficiency.
ESM4: Innovate through services and functional economy	This ESM explores the possibility to add services in the offer to avoid the ownership transfer.
ESM5: Innovate through local and collaborative networks	This ESM explores the possibilities of territorial and local resources, and collaborative networks (crowdfunding, fab-lab concept, etc.).
ESM6: Innovate through lifetime lengthening and closed loop thinking	This ESM explores the different possibilities to optimize use of materials remanufacturing, upgradability, or recycling.
ESM7: Innovate through new trends "materials-technologies-models-process"	This ESM explores the new technical advances (new material, new process, new technological model, etc.).

3 RESEARCH METHODOLOGY

Faced with this issue, an experiment was conceived, based on a panel of concrete cases of eco-innovation, operating in a bottom-up way, following the approach of case based reasoning. Starting from eco-innovative examples and "best practices" is relevant to better understand eco-innovation. Some academic works have already been proposed accordingly [Bocken et al., 2014; Hellström, 2013 ; Carrillo-Hermosilla et al., 2010].

This experiment was built on 300 eco-innovative cases. Any eco-innovation project / concept are eligible. In other words, only the Criterion Environmental Potential is required to select a case because it is not question in this paper to assess the viability or the dissemination of a eco-innovative project for the future but to identify in a broad way the possible origins (stimuli, problems, motivations, technological discoveries, etc.) to eco-innovate. A second criteria mobilized is the originality. In third criteria, the cases retained must cover different human needs (food, transport, energy, etc.) and different industrial / economic sectors (ICT, transport, household appliances, web-platform, etc.). The cases originated in particular from the French database Efficycle scanning online social and environmental oriented- projects (see Table 4).

Table 4. Presentation of 5 eco-innovation cases

Case study	Short description
Glowee Lighting	Biointeracting system without electricity consumption, thanks to natural properties of bioluminescent cells.
Obiflam log	Heat logs manufactured from sawdust (80%) and coffee grounds (20%).
Fairphone	Smartphone integrating ethical, social and environmental criteria (no conflict minerals, fair supply chain, modular and repairable)
Bike sharing	Large-scale public bicycle sharing system in Paris
Eco-cup	Sharing system of reusable and customizable cups for festivals and others events

The experiment take the form of the analysis of the 300 eco-innovations retained, dealing with the potential origins of case in imagining the starting problem to resolve to eco-innovate and the potential eco-innovation principles mobilized to solve the starting problem. The selection of the 300 cases is not intended to accurately reflect reality. Rather, the goal is to have a broad spectrum of potential type of eco-innovation. The experimental approach was conducted with 8 eco-design experts split in three groups of 2-3 participants corresponding to the 3 partners of ALIENOR project.

The experiment was structured in three parts:

- During Part 1, the analysis concerns only 50 cases and aims at identifying the list of “starting problem” and the list of “eco-innovation principles mobilized to solve the starting problem”. To do this, participants analysed individually and with their expertise the different cases and modelled the categories of starting problem and eco-innovation principles. Then, a collective discussion began with the pooling of individual results and was conducted in order to develop a common view within each group and between each group. The lack of information does not always make it easy to say what the original problem is or the real principles that have been mobilized. In some cases, there are several potential eco-innovation origins. The lists of “starting problem” and “eco-innovation principles mobilized to solve the starting problem” are considered valid if each item can be associated at least to one case and if each eco-innovation project / concept can be associated at least one origin.
- During Part 2, participants associated “starting problem” and “eco-innovation principles mobilized to solve the starting problem” with the 250 remaining cases of eco-innovation project / concept
- During Part 3, participants compared the lists of “starting problem” and “eco-innovation principles” with the set of ESMs to determine if these tools are sufficient to eco-innovate on each identified dimension or if they need to be completed.

4 EXPERIMENT AND RESULTS

4.1 List of “Starting problem” (SP)

From the analysis of Part 1 focused on 50 cases of eco-innovation emerge 5 different “starting problem” (SP). These problems invite to imagine a new system or to improve the existing system. They are the followings:

- The first SP is the observation of waste out of the scope of the system that is designed.
- The second SP is a lack in a group of systems at macro level
- The third SP is a lack of green intention of the existing system
- The fourth SP is a functional mismatch of the system in considering a frugal need or an augmented need (Figure 1).
- The fifth SP is a new principle of invention, that means a new intention concerning the existing system, a new invention at micro level or a new need or new problem that requires a new system (Figure 1).

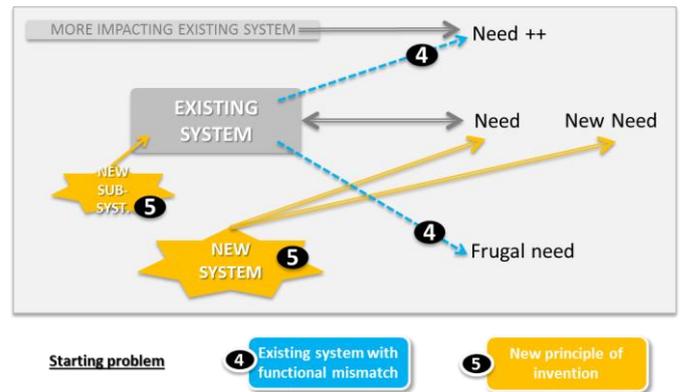


Figure 1. Presentation of Starting problem to eco-innovate

Starting problem	Eco-innovative principles	Links with ESM (Tyl et al. 2016)	Links with LiDS (Brezet & Van Hemel, 1997)
1 Observed waste 36	1a Transforming Waste flows Transform waste into raw material or as resources 23	ESM 1	LiDS 0
	1b Avoid waste (design of new system, better collection ...) 13	ESM 1	LiDS 0
2 Lack at the Macro-system level 107	2a New System that gives substance to Macro-system New system to generate a Macro-system that is of “sharing and pooling economy” type 38	ESM 1	LiDS 0
	2b New system to improve environmental performance at the Macro-system level 36	ESM 3	LiDS 0
	2c New system to generate a Macro-system that is of “new consumption behaviour” type 33	ESM 4	LiDS 0
3 Existing system with a lack of green intention 56	3a System with green intention Environmental quality of production and distribution modes 19	ESM 5	LiDS 3, 4
	3b Efficient Design and Use 32	ESM 2, 3, 7	LiDS 1, 2, 5
	3c Extended LifeTime & closed loop at the end-of-life 5	ESM 6	LiDS 6, 7
4 Existing system with functional mismatch 23	4a System considering a change of target Added features to compete with a more impacting system 4	ESM 3	LiDS 0
	4b Frugality for democratization of Good systems & products 19	ESM 3	LiDS 0
5 New principle of invention 78	5a New Techno Brick / New concept of System New discovery / techno development for new raw material or new technical module 34	ESM 2	LiDS 0
	5b New concept of system meeting a new intention (willingness to have much better performance, to solve a new problem, to change environmental consciousness of consumer ...) 44	ESM 2, 7	LiDS 0

Figure 2. Full results on the origins of eco-innovations.

4.1 List of “eco-innovation principles mobilized to solve the starting problem” (EP)

In front of these “starting problem”, participants identify “Eco-innovative principles” (EP). For each SP correspond(s) one or several EP. The results are presented in Figure 2 (Figure 2).

All eco-innovation cases are related to at least one starting problem and one eco-innovative principle. The quantitative results in terms of number of cases concerned by each item (SP, EP) are presented in (Figure 2) next to the item title. The eco-innovation cases related to a lack at the macro-system level are the most numerous (107). At the other side, those related to the starting problem “existing system with functional mismatch” represent a part of 8%.

4.2 Comparison between the lists of “starting problem” and “eco-innovation principles” and the set of ESMs

At the right of Figure 2 are presented the links with the ESMs proposed by Tyl et al. 2016] and with the Lifecycle Design Strategies (LiDS) of Brezet [Brezet & Van Hemel, 1997]. The link is strong when the meso-mechanism or the Lifecycle Design strategy covers the considered eco-innovation principle. This strong link is evidenced by a dark grey background. The link is partial, that is to say not complete, or “possible / rare” when the ESM or LiDS can contribute to the emergence of the type eco-innovation considered but when there is also more or less others paths to do this. These weak links are evidenced by a light grey background.

From the set of ESMs emerge for the most part eco-innovations related to the SP 2 and 3. Each of 7 ESMs except one is concerned in the form of strong link. The eco-innovations corresponding to the SP 2 and 3 represent more than 50% of cases. But for the others eco-innovations, the contribution of the set of ESMs is possible or rare because these tools are one potential path but not the most natural.

With regard to the Lifecycle Design Strategies (LiDS), the 7 LiDS are focused on the SP 2. Only the LiDS 0 related to “new concept development” deals with the others types of eco-innovation (SP 1, 2, 4, 5), with a limited contribution, except for “shared use of the product” that brings naturally Eco-innovative principle 2a. More generally, the LiDS wheel is a “macro” type tool that is to say eco-innovation axes to explore are given but without real methodological guidelines. So the contribution of ESMs to eco-innovate is stronger than that of LiDS approach.

More broadly, starting point to use a meso-mechanism or a LiDS is an existing system that it is subsequently “deformed” from different eco-innovative angles proposed. That's why certain eco-innovations are hard to generate for these methods. The watching of a plastic discharge into the sea or the discovery of a new material in the laboratory or the idea to build a vegetable garden floating above a river are examples of new opportunities to imagine new system whose starting point is not “an existing system” to improve. More generally, when the starting problem is out of the scope of an existing system, it is very hard to find tools in the literature to support these opportunities.

In the light of those, the real weakness of the set of ESMs and of LiDS is to not propose eco-innovation axes faced to the SP “existing system with functional mismatch” because the starting point concerns an existing system.

5 CONCLUSION AND DISCUSSION

In conclusion, this article focuses on a practical case-based approach to eco-innovation. 300 cases have been selected to be analysed by eco-design experts. Results reveal 5 Starting Problem (SP) to consider and 12 Eco-innovative Principles (EP) to use in order to eco-innovate. It is also showed that the scope of the Lifecycle Design Strategies are mainly focused on the EP related to the SP 3 and that the set of ESMs deals with the EP related to the SP 2 and 3.

For a design method starting from an existing system and using different eco-innovative angles to “deform” it, this paper identifies a no exploited possibility to eco-innovate: the EP 4a and 4b related to the SP “Existing system with functional mismatch”. This type of eco-innovations seem interesting to us, but there is very little literature on the subject. The term « downsizing » is common in industry, but no paper explains its methodological path. Likewise, Jugaad Innovation [Radjou et al., 2013] theorizes about frugal innovation but it is not very operable from a methodological point of view for a company. In design field, Tchertchian [Tchertchian et al., 2014] proposes the concept of functional negotiation for complex systems. In reality, the heart of the eco-innovations coming from EP 4a and 4b is the definition of the basket of functions, and that means to put into question the marketing practices.

On the basis of current knowledge, this paper therefore consolidates the status of the 7 ESMs to eco-innovate. Multiple tests are in progress to improve the set of ESMs towards an efficient use.

As perspective to complete the set of ESMs to eco-innovate, maybe it could be considered to scan the “problems related to a usage dimension” in addition to deforming the solutions space from an existing system. For example, in addition to using ESMs to eco-innovate from a transport system, others ESMs would lead a design team to eco-innovate in working on the problems of mobility (observation of waste out of the scope of the transport system designed, new need or new problem that requires a new system which does not yet exist, a lack in the management of transport systems that affect the global performance of mobility or the consumers behaviour ...).

6 ACKNOWLEDGEMENTS

The ALIENNOR project is an industrial research project financially supported by the French research agency (reference: ANR-15-CE10-0001).

7 RÉFÉRENCES

- Altshuller, G.S., "Creativity as an Exact Science, Gordon and Breach", ISSN 0275- 5807, New York, 1998.
- Bocken N.M.P, Allwood J.M., Willey A.R., King J.M.H. 2012. Development of a tool for rapidly assessing the implementation difficulty and emissions benefits of innovations, Technovation 32: 19–31.
- Bocken, N., Allwood, J., Willey, A., King, J., "Development of an eco-ideation tool to identify stepwise greenhouse gas emissions reduction options

- for consumer goods", JCP, Vol.19, No.12, 2011, pp. 1279-1287.
- Bocken, N., Short, S., Rana, P., Evans, S., "A literature and practice review to develop Sustainable Business Model Archetypes", Journal of Cleaner Production, Vol.65, 2014, pp. 42–56.
- Bocken, N., Short, S., Rana, P., Evans, S., "A value mapping tool for sustainable business modelling", Corporate Governance, Vol.13, No.5, 2013, pp.482 – 497
- Bogatyrev, N., Bogatyreva, "BioTRIZ: a win-win methodology for eco-innovation", In Eco-Innovation and the Development of Business Models, Springer International Publishing, pp. 297-314, 2014
- Brezet, H., Van Hemel, C., "Ecodesign: A Promising Approach to Sustainable Production and Consumption", UNEP, Paris, 1997.
- Carrillo-Hermosilla J., del Rio P., Könnöla T. (2010) Diversity of eco-innovations: Reflections from selected case-studies. Journal of Cleaner Production, 18 (10-11), pp. 1073-1083.
- Cluzel F., Vallet F., Tyl B., Bertoluci G., Leroy Y. (2014) Eco-design vs eco-innovation: an industrial survey, International design conference – Design 2014. Dubrovnik, Croatia.
- Cluzel, F., Yannou, B., Millet, D., Leroy, Y. (2016). "Eco-ideation and eco-selection of R&D projects portfolio in complex systems industries", Journal of Cleaner Production, Vol.112, pp. 4329-4343.
- Díaz-García, C., González-Moreno, Á., Sáez-Martínez, F. J. (2015). Eco-innovation: insights from a literature review. Innovation, Organization and management 17(1), pp. 6-23.
- Fussler, C., James, P., "Driving Eco-innovation. A Breakthrough Discipline for Innovation and Sustainability", Pitman Ed., London, 1996.
- Hellström, T. (2013) Dimensions of environmentally sustainable innovation: the structure of eco-innovation concepts. Sustainable Development 2013, 15(3), pp. 148-159.
- Hojnik, J., Ruzzier, M. (2016). What drives eco-innovation? A review of an emerging literature. Environmental Innovation and Societal Transitions 19, pp. 31-41.
- Hunt R.G., Boguski T.K., Weitz K. and Sharma A. (1998): Case Studies Examining LCA Streamlining Techniques. International Journal of LCA, Vol. 3, N°1, pp.36–42
- Joyce, A., Paquin, R., Pigneur, Y., " The triple layered business model canvas: a tool to design more sustainable business models" ARTEM Organizational Creativity International Conference, Nancy, 2015.
- Kobayashi, H., "A systematic approach to eco-innovative product design based on life cycle planning", Advanced Engineering Informatics, Vol.20, 2006, pp113–125.
- Mathieu, A., Reynaud, E., Chandon, J.L.: Les déterminants internes de l'éco innovation : Analyse de 118 éco innovations selon le référentiel gestionnaire et la stratégie RSE de l'entreprise, (2015).
- O'Hare, J.A., McAlloone, T.C., Pigosso, D.C.A., Howard, T.J., "Eco-Innovation Manual – Tools instruction", United Nations Environment Programme / DTU, 2014.
- O'Hare, J.A., McAlloone, T.C.: Eco-innovation: the opportunities for engineering design research, in: DS 77: In: Proceedings of the DESIGN 2014, Dubrovnik, Croatia, pp. 1631–1640 (2014)
- Osterwalder, A., Pigneur, Y., "Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers", John Wiley & Sons, Hoboken, New Jersey, 2010.
- Pialot O., Millet D., Towards Operable Criteria of Eco-innovation and Eco-ideation Tools for the Early Design Phases; 25th CIRP Life Cycle Engineering (LCE) Conference, 30 April – 2 May 2018, Copenhagen, Denmark, Procedia CIRP, Volume 69, Pages 692-697, 2018
- Radjou N., Prabhu J., Ahuja S., Boillot J-J., Innovation Jugaad. Redevons ingénieurs ! Ed. Broché – 2013
- Rennings, K.: Redefining innovation-eco-innovation research and the contribution from ecological economics. 32(2), 319–332 (2000)
- Tchertchian N., Millet D., Yvars P-A., How to improve environmental performance by negotiating functional specifications of complex system?; 21st CIRP Conference on Life Cycle Engineering, Trondheim, Procedia CIRP 15, June 2014
- Tyl B., Legardeur J., Millet D., Vallet F., 2014. "A comparative study of ideation mechanisms used in eco innovation tools", Journal of Engineering Design, Vol.25, No.10-12, 2014, pp.325-345.
- Tyl, B., Vallet, F., Pialot, O., Millet, D., Le Duigou, J., Graves, G. (2016) The ESM approach: 8 mechanisms to efficiently support eco-ideation. In: Proceedings of Design 2016. Dubrovnik, Croatia.
- United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects: The 2019 Revision, Key Findings and Advance Tables.
- Vallet F., Tyl B., Millet D. and Eynard B. (2013) A method to select best nuggets from eco-innovation sessions. In Green Design, Materials and Manufacturing Processes, H. Bartolo et al. (Eds), pp 647-654. CRC Press/Balkema, Leiden Netherlands.

CIGI 2021

L'optimisation hybride multidisciplinaire pour concevoir les produits de demain

Résumé – Utilisée en aéronautique pour améliorer une bonne solution aérodynamique et structurelle (forme d'aile) par l'algorithme du gradient (méthode adjointe), l'optimisation multidisciplinaire (MDO : *Multidisciplinary Design Optimisation*) cherche à trouver les meilleurs compromis entre les divers métiers. Mais elle ne peut trouver des solutions radicalement nouvelles que par l'emploi de méthodes d'optimisation globale (algorithmes génétiques, recuit simulé...) dont les temps de calcul sont le plus souvent rédhibitoires. Cependant l'hybridation de méthodes d'optimisation globale et locale pallie cette difficulté en accélérant considérablement la vitesse de convergence. Cet article montre l'apport de méthodes d'optimisation hybride à la conception d'un drone à voilure fixe (Cabiota). Il en souligne également les limites ainsi que les caractéristiques attendues des méthodes d'optimisation multidisciplinaire pour les rendre applicables à tous les domaines d'ingénierie. Elles concernent notamment l'aéronautique civile qui doit améliorer le bilan carbone des avions de ligne, dont la configuration n'a guère évolué depuis les premiers avions à réaction (Boeing 707).

Abstract – Used in Aeronautics to improve a good aerodynamic and structural solution (wing shape) by the gradient algorithm (adjunct method), Multidisciplinary Design Optimisation (MDO) seeks to find the best compromises between various disciplines. However, it can only obtain radically better solutions when involving global optimisation methods (e.g. genetic algorithms, simulated annealing, etc.) for which the computation times are usually prohibitive. Some hybridisation between global and local optimisation methods can nevertheless overcome such difficulty by considerably accelerating the speed of convergence. This article shows the application of hybrid optimisation methods to the design of a fixed-wing drone (Cabiota). It also tries to highlight the advantages of such methods as well as their current limitations so that MDO can be applied to many more engineering fields. These methods can, in particular, be of interest in civil aeronautics, where carbon footprint of airliners must be reduced significantly in the near future and where aeroplane designs have not changed much since the first jets (i.e. Boeing 707).

Mots clés – Conception, MDO, Optimisation multidisciplinaire, Optimisation hybride, drone.

Keywords – Design, MDO, Multidisciplinary Design Optimisation, Hybrid optimisation, drone, UAV.

1 INTRODUCTION

L'Optimisation de Conception Multidisciplinaire (*Multidisciplinary Design Optimisation*, soit MDO en anglais) est un domaine d'ingénierie qui cherche à résoudre des problèmes de conception optimale mettant en œuvre plusieurs disciplines [Vanderplaats, 2007]. L'optimum trouvé est alors meilleur que celui obtenu en optimisant chacune des disciplines indépendamment des autres, car il prend en compte leurs interactions. Recouvrant différents métiers, cette recherche d'un optimum global entraîne cependant de la complexité et un surcoût en temps de calcul souvent rédhibitoire.

L'optimisation multidisciplinaire est déjà utilisée en aéronautique pour améliorer une bonne solution aérodynamique et structurelle (forme d'aile) par l'algorithme du gradient (calculé par la méthode adjointe). Mais une méthode d'optimisation locale ne permet pas de trouver des solutions radicalement nouvelles placées aux optima globaux, car elle reste bloquée sur l'optimum local situé au voisinage des solutions connues. Par ailleurs, les variables de nature diverse (réel, entier, alphanumérique) interdisent tout calcul de dérivée et seuls les résultats de la fonction critère et des

fonctions contraintes sont généralement accessibles, par calcul ou simulation, pour chacune des configurations de variables.

Des méthodes d'optimisation globale, telles que les algorithmes génétiques ou le recuit simulé, sont susceptibles de trouver de bonnes solutions, sans cependant garantir de leur optimalité. Mais leur vitesse de convergence est très lente. Outre leur caractère stochastique, fondé sur une recherche aléatoire de l'optimum, elles doivent effectuer un nombre d'évaluations des objectifs et des contraintes qui croît exponentiellement en fonction du nombre de variables de conception. Cependant l'hybridation de méthodes d'optimisation globale et locale pallie cette difficulté en accélérant considérablement la vitesse de convergence.

Une méthode d'optimisation hybride a ainsi été utilisée par la société Cab Innovation pour concevoir un drone à voilure fixe (Cabiota) qui présente des performances sensiblement supérieures à celles des produits existants et des caractéristiques non entièrement prévues des concepteurs eux-mêmes.

En dépit des marges de conception offertes par l'absence de personne à bord, ce type de drone n'a fait l'objet, jusqu'à présent, que de peu de recherches sur le plan scientifique et technique et sa configuration aérodynamique reste le plus

souvent semblable à celle des avions pilotés. De même, les avions de ligne n'ont guère évolué depuis les premiers avions à réaction (notamment le Boeing 707 mis en service en 1958) mais doivent désormais changer radicalement de design pour améliorer leur bilan carbone.

Cet article montre l'application de méthodes d'optimisation hybride à la conception d'un drone à voilure fixe. Il en souligne les apports mais aussi les limites afin de déterminer les caractéristiques attendues des méthodes d'optimisation hybride multidisciplinaire pour les rendre applicables à tous les domaines d'ingénierie.

Son aspect scientifique ne porte pas sur l'hybridation de méthodes d'optimisation globales et locales, connue de la littérature mais quasiment absente de l'ingénierie des systèmes, mais sur la mise en lumière d'un potentiel, largement inconnu des concepteurs, à travers une application concrète d'une certaine complexité. Cette dernière permet, par ailleurs, de souligner les insuffisances des outils de MDO existants, ce qui constitue un préalable aux projets de Recherche et Développement collaboratifs que les auteurs de cet article appellent de leurs vœux.

2 L'OPTIMISATION MULTIDISCIPLINAIRE

L'activité de conception des systèmes recouvre des métiers très différents (marketing, structure, aérodynamique, contrôle, fiabilité, etc.). Cependant, une question primordiale reste à résoudre : comment réaliser les meilleurs compromis entre ces diverses disciplines sachant qu'un optimum global résulte rarement de la juxtaposition des optima locaux ? L'optimisation multidisciplinaire cherche à résoudre ce problème [Cramer et al, 1994]. Illustré par la figure 1, un système multidisciplinaire comprend des variables globales (z), des variables locales propres à chacune des disciplines (x) et des variables de partage (y) qui sont des sorties des différentes disciplines.

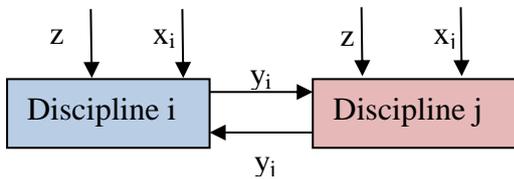


Figure 1. Système multidisciplinaire

L'optimum recherché est meilleur que la configuration trouvée en optimisant chacune des disciplines indépendamment des autres, car il prend en compte leurs interactions. En effet, quand l'optimisation s'arrête aux frontières des disciplines, la gestion des interfaces conduit à considérer des pires cas et à cumuler des marges cachées qui peuvent avoir des effets de seuil très pénalisants, tels que le changement de catégorie de satellite (nano, micro, mini, gros...) ou de lanceur (Véga, Soyouz, Ariane...) dans le domaine spatial, pour la mise en orbite d'une charge utile.

2.1 Méthodes mono et multiniveaux

L'optimisation multidisciplinaire fait appel à des méthodes mono ou multiniveaux telles que la méthode MDF (*Multidisciplinary Feasible Design*) ou la méthode collaborative CO (*Collaborative optimisation*) représentée à la figure 2 [Braun et al, 1996].

La Méthode MDF met en œuvre un optimiseur unique qui agit sur un système intégré comprenant un MDA (*Multidisciplinary Design Analysis*). Ce dernier détermine les sorties des disciplines à chaque itération du processus d'optimisation en

résolvant l'équation d'état que doit vérifier chacune des variables de couplage.

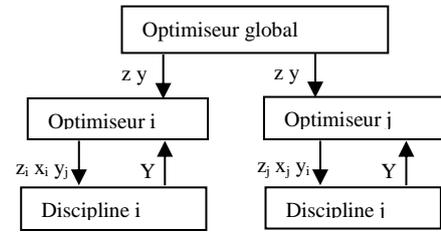


Figure 2. Méthode collaborative CO

La méthode CO met en œuvre plusieurs optimiseurs qui reproduisent le fonctionnement d'une entreprise où la direction donne des consignes que les différents services tentent de respecter. Au niveau système, un optimiseur joue sur les variables globales et de partage pour minimiser la fonction de coût tout en respectant les contraintes système ainsi que des contraintes d'égalité assurant le respect des consignes au niveau inférieur.

2.2 Optimisation locale par la méthode du gradient

La méthode du gradient et ses dérivées sont utilisées en optimisation multidisciplinaire pour améliorer localement une solution [Tedford et Martin, 2010]. Les dérivées de la fonction de performance n'étant généralement pas fournies par les codes de calcul, le gradient peut être estimé par le calcul des différences finies (méthode du pseudo gradient) qui s'avère très pénalisant en temps de calcul quand il porte sur un grand nombre de paramètres de conception. Le gradient peut être également estimé plus efficacement par la méthode GSE (*Global Sensitivity Equation*) qui est une méthode adjointe utilisant l'information apportée par l'équation d'état et sa dérivée dans la méthode MDF [Martins et Hwang, 2013].

La fonction à optimiser est de la forme : $f = F(p, Y(p))$ où p désigne le vecteur des paramètres de design (z et x) et $Y(p)$ le vecteur des variables d'échange (y) qui respecte les équations d'état :

$$R(p, Y(p)) = 0$$

La dérivée totale de la fonction f a pour expression :

$$\frac{df}{dp} = \frac{\partial f}{\partial p} + \frac{\partial f}{\partial Y} \frac{dY}{dp}$$

Le calcul de $\frac{dY}{dp}$ est difficile car il n'y a en général pas de solution analytique de l'équation d'état $R(p, Y(p)) = 0$, qui est résolue grâce à des méthodes numériques.

Ce problème est contourné en considérant que la dérivée totale du résidu R est nulle :

$$\frac{dR}{dp} = \frac{\partial R}{\partial p} + \frac{\partial R}{\partial Y} \frac{dY}{dp} = 0 \text{ d'où l'expression } \frac{dY}{dp} = - \left[\frac{\partial R}{\partial Y} \right]^{-1} \frac{\partial R}{\partial p}$$

que l'on peut remplacer dans celle de la dérivée :

$$\frac{df}{dp} = \frac{\partial f}{\partial p} - \frac{\partial f}{\partial Y} \left[\frac{\partial R}{\partial Y} \right]^{-1} \frac{\partial R}{\partial p}$$

On introduit alors le vecteur des variables adjointes Λ , tel que :

$$\left[\frac{\partial R}{\partial Y} \right]^T \Lambda = - \left[\frac{\partial f}{\partial Y} \right]^T$$

Que l'on remplace dans l'expression de la dérivée :

$$\frac{df}{dp} = \frac{\partial f}{\partial p} - \Lambda^T \frac{\partial R}{\partial p}$$

L'algorithme résout d'abord numériquement les équations d'état $R(p, Y(p)) = 0$ pour un vecteur p donné, puis calcule numériquement l'adjoint Λ avant d'obtenir la dérivée. L'algorithme d'optimisation peut ainsi sélectionner le vecteur p de l'étape suivante.

La méthode GSE est moins coûteuse en temps de calcul car elle ne nécessite pas un calcul couplé par variable de conception. Il y a un seul calcul couplé et les dérivées sont fournies par chacune des disciplines. L'équation d'état peut être résolue de manière précise afin d'obtenir un gradient de bonne précision. Cette méthode d'optimisation est notamment utilisée pour la conception aérodynamique et structurelle des ailes d'avion. [Zhoujie et al, 1994].

2.3 L'optimisation multidisciplinaire aujourd'hui

L'optimisation multidisciplinaire est actuellement en vogue dans la recherche en aéronautique, informatique et mathématiques appliquées. Elle permet d'améliorer localement la forme des avions, mais rencontre des freins à son développement.

Dans les différents domaines de l'ingénierie, les outils, les méthodes, les techniques voire les technologies sont souvent cloisonnées par métier ou domaine de travail, ce qui limite leur adoption par d'autres corps de métier. Ainsi, les dernières technologies de calculs massivement parallèles et distribués, introduites par le *Big Data*, sont peu utilisées en optimisation, malgré leur intérêt évident pour réduire les temps de calcul. De même, les méthodes de virtualisation, d'abstraction ou d'interfaces, communes aux domaines du *Cloud computing* et du *Web* en général, semblent inconnues des concepteurs de logiciel de MDO pourtant confrontés aux mêmes défis, à savoir l'intégration d'outils hétérogènes devant travailler de concert.

Dans les sciences des données (*Data science*), les méthodes d'optimisation globales restent confidentielles et leur hybridation à des méthodes locales semble inconnue. La plupart des bibliothèques de *machine learning* se fondent sur des méthodes simples, voire simplistes, de recherche locale sans exploiter l'état de l'art de la spécialité.

Par ailleurs, l'optimisation multidisciplinaire n'aura un véritable impact sur les produits de demain que si les outils qui la mettent en œuvre se démocratisent et deviennent accessibles, voire incontournables, pour l'ensemble des ingénieurs en conception à la manière des réseaux de neurones pour l'intelligence artificielle [Cabarbaye et al, 2016].

3 APPORT DE L'OPTIMISATION HYBRIDE A LA CONCEPTION D'UN DRONE

Bénéficiant de l'expérience acquise durant la réalisation du drone météorologique CABFLY de 2,5 kg dans le cadre du projet ANR NEPHELAE (CNRM, LAAS, ENAC, ENM) [Cabarbaye et al, 2019], le drone Cabiota a été développé par la société Cab Innovation au moyen d'un outil d'optimisation hybride, Gencab, également développé par la même société. Cet outil générique associe une méthode globale (algorithmes génétiques) à une méthode locale (simplexe non linéaire ou algorithme de Nelder-Mead) pour accélérer la convergence vers l'optimum global.

3.1 Présentation du contexte

Né depuis une dizaine d'années, le secteur des drones civils a principalement été porté par le marché des drones dits de loisir jusqu'en 2018.

A partir de cette date, en raison de la faiblesse du marché de renouvellement, l'industrie s'est repositionnée sur le segment professionnel, principalement pour la photographie et la cartographie aérienne.

Cependant, le domaine spatial offre des solutions abordables qui entrent directement en concurrence avec les drones mais avec des coûts d'utilisation plus faibles et à beaucoup plus grande échelle.

Les drones présentent toutefois les avantages suivants par rapport aux satellites :

- type de charges utiles non limitées aux seuls capteurs passifs ou à énergie dirigée,
- vol sous le couvert nuageux pour des applications en temps réel,
- altitude de vol variable,
- sites visités en dehors de la période de défilement des satellites en orbite basse.

De nouvelles applications peuvent être, par ailleurs, envisagées comme par exemple :

- le transport de marchandise à haute valeur ajoutée (en concurrence avec le transport routier ou aérien),
- l'épandage de produit phytosanitaire en simultané avec la détection (en concurrence avec les méthodes agricoles classiques),
- l'imagerie très haute résolution en volant à basse altitude (en concurrence avec l'aviation classique et les satellites haut de gamme),
- la recherche scientifique avec analyse de l'atmosphère (e.g. Météorologie, Volcanologie),
- la surveillance, etc.

Pour se faire une place sur ces nouveaux marchés, les performances en termes de vitesse de croisière et d'endurance doivent être grandement améliorées afin d'augmenter la distance franchissable (actuellement limitée à 50 kms), la surface « survolable » (actuellement limitée à 10km² pour une journée de travail), le temps de vol (actuellement de plus ou moins 2 heures) et la possibilité de voler par tout temps (actuellement le vol par fort vent est limité du fait des faibles vitesses de croisière, autour de 15 m/s).

De telles améliorations ne peuvent être obtenues par de simples optimisations locales des solutions actuellement existantes.

Par ailleurs, sur un plan juridique, les drones doivent respecter depuis le 01/01/2021 la nouvelle réglementation européenne qui instaure différentes classes dépendantes du poids du drone considéré :

- les classes de drones de moins de 0,9 kg, qui sont surtout destinées aux drones de loisir,
- la classe C2 des drones de moins de 4kg, qui est la moins contraignante et qui est susceptible d'intégrer des drones professionnels,
- la classe C3, qui limite la masse totale à 25 kg.

Compte tenu des nombreuses exigences imposées à la classe C3, le drone développé par Cab Innovation entre dans la classe C2 et a suivi le cahier des charges suivant:

- masse maximale au décollage de 4kg,
- plus grande autonomie possible,
- grande modularité

Ce dernier point implique une soute relativement grande afin d'accueillir l'ensemble de la charge utile, les batteries et une configuration propulsive afin de dégager le nez de l'appareil pour d'éventuels capteurs. Ceci a pour conséquence de figer la configuration du fuselage.

L'optimisation portait alors principalement sur les ailes et le système propulsif du drone.

Enfin une propulsion électrique est choisie en raison de sa simplicité, de son adaptabilité et des contraintes réglementaires.

3.2 Formalisation du problème d'optimisation

3.2.1 Les paramètres

Les paramètres d'optimisation considérés sont l'envergure, la corde, l'effilement, la flèche et le calage de l'aile à l'emplanture et au saumon qui sont représentés par des variables réelles, ainsi que les choix des modèles de moteur, d'hélice et des profils d'aile à l'emplanture et au saumon qui sont représentés par des variables alphanumériques.

3.2.2 Limitation due à la modélisation

Si l'optimisation hybride ouvre la plage des solutions envisageables, la formalisation du problème peut la réduire intrinsèquement. De bonnes solutions peuvent ainsi être ignorées parce qu'elles ne sont pas accessibles par la modélisation retenue.

La bio-inspiration peut aider à s'affranchir de ces limitations en vérifiant que des solutions adoptées par la nature ne sont pas omises par le modèle.

Dans le cas présent, une étude CFD des « ailes en mouette », inspirées par les oiseaux marins homonymes a montré un intérêt double de cette configuration. Elle présente une trainée induite réduite par rapport à une aile de même envergure. L'angle donné aux ailerons améliore la symétrie du vol en virage.

Enfin, dans la configuration retenue d'aile en flèche, cette cassure améliore la stabilité directionnelle. Ceci permet de ce fait de réduire la taille des stabilisateurs verticaux.

Deux paramètres supplémentaires d'optimisation (définis par des variables réelles) sont donc ajoutés : les positions verticale et horizontale de la cassure des ailes présentées figure 3.

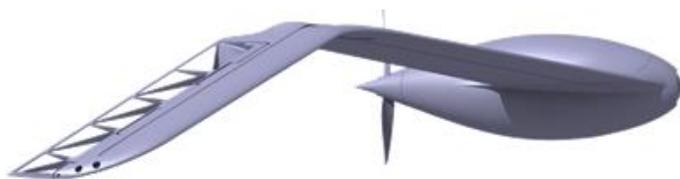


Figure 3. Configuration finale du drone Cabiota

3.2.3 Les contraintes

La seule contrainte considérée est la vitesse de décrochage qui est choisie afin de pouvoir atterrir avec une marge de 10%. La vitesse de décollage est, dans un premier temps, limitée à 15m/s.

3.2.4 Le critère

Le drone étant envisagé pour répondre à un grand nombre d'applications différentes, 2 critères sont considérés.

Le premier est le temps de vol maximal en croisière (endurance). En effet un grand nombre de missions consiste en survoler un objectif le plus longtemps possible en effectuant un circuit d'attente (i.e. mission de surveillance de zone).

Le deuxième est la distance « survolable » par unité de temps (vitesse de croisière). En effet, le coût d'exploitation d'un

drone étant principalement constitué par le salaire du télépilote, ce critère impacte directement la rentabilité de son exploitation (i.e. cartographie, surveillance de ligne).

La Cabiota est donc optimisée pour maximiser ces deux critères au cours de la mission présentée à la figure 4.

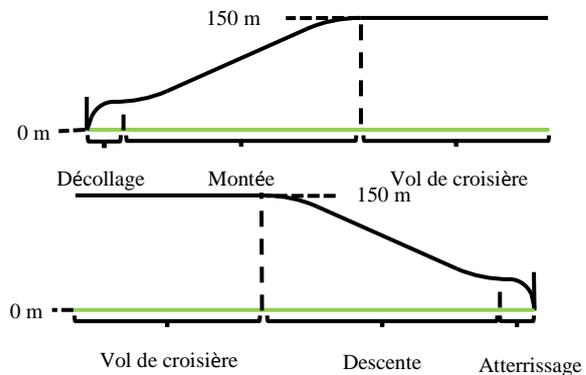


Figure 4. Mission type

Plusieurs calculs intermédiaires sont nécessaires pour estimer ce critère:

- Mécanique

Chaque élément mécanique est dimensionné pour supporter les contraintes propres à la configuration considérée. La masse de chaque élément est ensuite estimée. Ces estimations sont ensuite sommées, donnant ainsi la masse à vide de l'appareil. Cette masse permet d'estimer la masse de batterie admissible restante pour ne pas dépasser les 4 kg.

- Aérodynamique

La vitesse de décrochage de la configuration considérée est estimée afin de vérifier le respect de la contrainte.

Les performances aérodynamiques de l'avion et la puissance mécanique nécessaire résultante sont estimées pour chaque phase de vol.

- Électrique

Une analyse électrique du système motopropulseur permet d'estimer la consommation et enfin le temps de vol en croisière.

A cause des limitations du logiciel employé, ces estimations sont effectuées au moyen de formules semi empiriques et des modèles théoriques simples issues du domaine de la conception amont aéronautique.

3.2.5 Tendance de design préliminaire

Les premiers résultats du calcul par l'algorithme d'optimisation ont été assez différents de ceux qu'attendaient les concepteurs.

Le front de Pareto obtenu suite à la maximisation des 2 critères considérés pour la Cabiota est présenté à la figure 5.

Il apparaît que le meilleur compromis se trouve autour de 30 m/s. Ce qui est 2 fois supérieur à la norme actuellement observée sur le marché. La vitesse de croisière est donc fixée à 30 m/s pour la suite de l'optimisation. Cette vitesse est comparable à celle du transport routier, ce qui rend la Cabiota compétitive pour la livraison de colis.

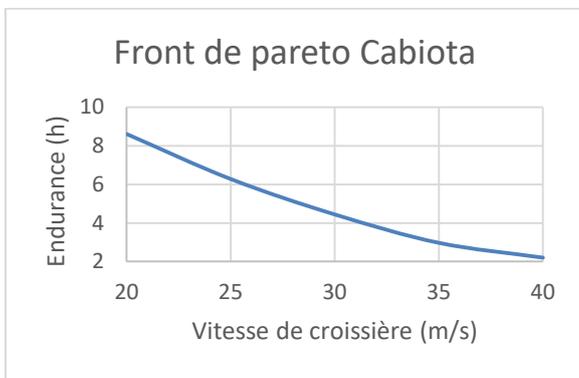


Figure 5. Front de Pareto Cabiota

Le système de propulsion constitué du moteur et de l'hélice donne au drone une puissance maximale près de 3 fois supérieure à sa puissance en vol de croisière. Or ce rapport tourne en général autour de 1.5.

Cette sur-motorisation permet une vitesse de montée très élevée qui réduit le temps de montée en début de mission, et par conséquent, l'énergie consommée au cours de cette phase. Il est à noter que le fort rapport poids/puissance des moteurs électriques n'impacte que peu le devis de masse et donc la capacité de batterie.

Par ailleurs, la position horizontale de la cassure de l'aile obtenue coïncide exactement avec la substitution complète du stabilisateur vertical. Ce paramètre a donc pu être supprimé et sa valeur calculée analytiquement pour alléger les calculs d'optimisation suivants.

Enfin, la contrainte de vitesse de décrochage est très rapidement atteinte au cours de l'optimisation. Une grande vitesse de décrochage est en effet induite par une charge alaire élevée qui est nécessaire pour réduire la surface d'aile mouillée et donc la traînée. Une charge alaire élevée présente, par ailleurs, l'avantage d'être moins sensible à la turbulence.

Il apparait ainsi qu'une augmentation de la vitesse de décrochage soit nécessaire pour améliorer les performances du drone. Mais un système de lancement adapté est alors nécessaire.

3.2.6 Décollage court et atterrissage vertical

Plusieurs systèmes d'assistance au décollage existent, principalement basés sur les principes de la catapulte ou du sandow. Or ces solutions, soit ne permettent pas de transmettre suffisamment d'énergie, soit imposent une accélération initiale beaucoup trop grande pour la centrale inertielle et éventuellement la charge utile. La solution la plus adaptée semble alors être l'utilisation de moteurs fusées (systèmes de type JATO ou RATO) qui présentent malheureusement un risque d'incendie élevé.

Un système de lancement hydropneumatique (basé sur le même principe que les fusées à eau) a donc été développé.

L'optimisation de ce système a été effectuée simultanément avec le reste du drone, en ajoutant comme paramètre supplémentaire, le volume d'air dans la fusée.

En raison des vitesses de vol atteintes, il a été décidé que l'atterrissage serait effectué sous parachute, le drone devenant ainsi de type ADAC / V (décollage court et atterrissage vertical). L'ensemble des phases de vol est présenté sur la figure 6.

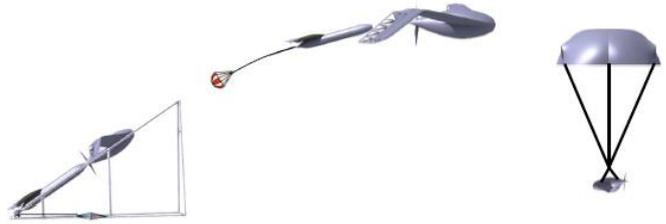


Figure 6. Phases de vol

3.2.7 Résultats de l'optimisation

L'optimisation complète de l'ensemble du système drone conduit à l'architecture présentée à la figure 7 :



Figure 7. Système Cabiota

Une vitesse de libération de 18m/s est atteinte faisant passer la vitesse de décrochage admissible à 16.3 m/s, ce qui est plus du double que celle des drones ordinaires de la classe C2. La surface alaire est ainsi divisée par 4.

Au terme de la conception, les performances du drone CABIOTA sont comparées avec celles des principaux drones de classes C2 actuellement en service, dans le tableau 1. L'autonomie atteinte se révèle très supérieure à celle des produits concurrents. L'optimisation hybride a permis d'obtenir une solution très éloignée des configurations habituelles, correspondant à des optima locaux. Traitant l'ensemble du système de manière globale, la conception du drone a, par ailleurs, été beaucoup plus courte que celle employant des méthodes d'optimisations locales prises séparément.

Ref. produit	UX11	eBee X	Talon	X6	AVEM	Cabiota
Vitesse	15 m/s	15 m/s	15m/s	15 m/s	18 m/s	30 m/s
Masse	1.5 kg	1.4 kg	4 kg	4 kg	2 kg	4 kg
CU	0.250 kg	0.250 kg	0.7 kg	0.5 kg	0.5 kg	De 0.5 à 2.5 kg
Taille	1.1 m	1.16 m	1.72 m	1.5 m	2.14 m	1.8m
Autonomie	1h00	1h00	2h00	2h00	3h00	5h00 (0.5 kg CU) 7h00 à 21m/s

Tableau 1. Tableau comparatif des performances

Toutefois, les contraintes de l'outil d'optimisation utilisé, qui fonctionne dans l'environnement Excel sans réelle possibilité de couplage avec des outils de simulation, ne permettent pas d'exploiter complètement le potentiel de l'optimisation hybride. En effet, si ce calcul était basé sur l'état de l'art de chaque discipline (e.g. CFD, FEM), l'incertitude de l'estimation pourrait être grandement diminuée. Cela permettrait de prendre en compte un beaucoup plus grand nombre de paramètres dont l'impact ne peut pas être évalué par l'estimation actuelle. Ainsi, des solutions inédites pourraient être trouvées sans être basées sur une intuition initiale comme l'« aile en mouette » ou sur un choix parmi des solutions préexistantes comme les profils d'aile qui pourrait facilement être paramétré [Nilesh et al, 2014].

A travers les difficultés et limitations rencontrées durant ce processus de conception innovant, cette application relativement complexe a toutefois permis de préciser les attentes de concepteurs vis-à-vis des outils de MDO.

4 ATTENTES VIS-A-VIS DES METHODES ET OUTILS DE MDO

Pour rendre l'optimisation multidisciplinaire véritablement opérationnelle et pouvoir satisfaire les attentes qu'elle suscite, il apparaît nécessaire de préciser les besoins des concepteurs vis-à-vis des outils de MDO.

4.1 Recherche d'un optimum global

Les méthodes d'optimisation locale, telles que l'algorithme du gradient calculé par la méthode adjointe, sont efficaces pour améliorer une bonne solution mais ne permettent pas d'en trouver de nouvelles en dehors de son voisinage immédiat. Les outils de MDO doivent donc se doter d'une capacité de recherche de l'optimum global pour trouver des solutions innovantes. Cependant, l'hybridation de méthodes globales et locales s'avère indispensable pour résoudre la plupart des problématiques industrielles sans que la durée de convergence devienne rédhibitoire.

4.2 Simplicité et généralité

Les outils de MDO doivent être génériques, afin de pouvoir s'adapter à la diversité des problèmes rencontrés sans imposer de restrictions particulières, tout en étant simples d'utilisation pour des ingénieurs non informaticiens. La formalisation du problème et notamment l'expression des variables et des contraintes doit être suffisamment intuitive pour être facilement compréhensible par l'ensemble des personnes impliquées dans la conception d'un système. Aussi, le codage des variables doit-il être riche, naturel et souple. Il peut être binaire, réel, entier, mais aussi alphanumérique (choix d'un profil aérodynamique standard dans une liste par exemple), ou adapté au dénombrement tel que la permutation (problème du voyageur de commerce), l'arrangement (problème décisionnel) ou la combinaison (problème du sac à dos). D'autres formats plus spécifiques peuvent être également envisagés comme celui décrivant une arborescence.

4.3 Facilité de couplage

Le couplage des outils d'optimisation avec n'importe quel logiciel d'analyse existant doit être aisé, même si ce dernier n'a pas été écrit pour être un jour utilisé à des fins de conception automatisée. Ainsi, le logiciel Xfoil de simulation aérodynamique ne fonctionne que sur des architectures particulières (architecture x86) avec des interfaces machine-machine assez restreintes, voire inexistantes. Aussi, les

évaluations doivent-elles pouvoir s'opérer sur des environnements propres à ces dernières, qui peuvent être différents des environnements d'exécution des algorithmes d'optimisation. Par ailleurs, il faut souligner l'importance d'un couplage efficace à la simulation de Monte-Carlo afin de pouvoir tirer tout l'avantage de la prise en compte de l'aléa dans la résolution de nombreuses problématiques (la diversité des trajets pouvant être effectués par un avion commercial par exemple) [Cabarbaye et Laulheret, 2006].

4.4 Méthodes multiniveaux

Outre la simplicité apportée pour décomposer un problème complexe en problèmes plus restreints, les méthodes multiniveaux se révèlent indispensables quand l'évaluation d'une solution nécessite elle-même une optimisation.

4.5 Recherche multi-objectifs

S'efforçant de trouver le meilleur compromis entre des objectifs parfois contradictoires d'une discipline à l'autre, les outils de MDO doivent avoir la faculté de mener une recherche multi-objectif.

4.6 Calcul parallèle et distribué

La multiplicité des traitements rend nécessaire la parallélisation des calculs et la distribution du MDO sur plusieurs supports traitant des algorithmes d'optimisation et des opérations d'évaluation. Compte tenu de l'existence de *clouds* publics, il apparaît judicieux de n'utiliser des machines de calcul dédiées que quand cela est nécessaire et de s'assurer que les besoins en calculs trouvent des réponses adéquates et adaptatives, afin que les ressources mises à disposition soient utilisées à bon escient en évitant tout phénomène d'engorgement pouvant avoir un impact sur le temps de calcul global.

4.7 Accessibilité

Les outils de MDO doivent être non seulement flexibles et adaptés à un large éventail d'usage mais aussi facilement accessibles pour pouvoir être essayés et évalués par un grand nombre d'utilisateurs différents (experts, chercheurs, concepteur, intégrateurs, etc.). A ce titre, la diffusion en open-source présente un avantage significatif.

4.8 Synthèse

En résumé, l'outil attendu de MDO doit être ergonomique, générique, robuste, efficient (résultat sans effort), multi-objectif, portable (fonctionnant sur des environnements hétéroclites), interopérable, extensible, évolutif, etc. Il est à noter que la performance et la rapidité de traitement ne sont pas mentionnées dans cette liste des qualités requises. Peu importe, en effet, que d'autres outils puissent être plus performants pour un problème donné si celui-ci est traité sans effort, dans un temps raisonnable avec des solutions meilleures que celles trouvées jusqu'alors [Cabarbaye et Cabarbaye, 2019].

4.9 Autres pistes favorisant la MDO

Outre l'utilisation d'outils d'optimisation performants, la mise en œuvre de la MDO peut être facilitée par diverses approches permettant de réduire la durée d'évaluation de chacune des solutions.

Les modèles représentant l'évolution des objectifs et des contraintes, en fonction des variables de conception, peuvent être plus ou moins précis au cours de la recherche. Des modèles empiriques, comme des profils standards en

aérodynamique ou des analyses régressives sur les coûts, peuvent être utilisés pour dégrossir le problème avant que des modèles théoriques plus poussés, tels que les éléments finis ou la CFD (*Computational Fluid Dynamics*), affinent les solutions trouvées.

Des modèles réduits intermédiaires peuvent être également envisagés tels que la surface de réponse polynomiale d'un plan d'expériences dont les solutions résultent d'un outil de simulation aérodynamique, par exemple. De même, des modèles de substitution peuvent être obtenus par des techniques, telles que le krigeage ou la machine à vecteurs de support, quand des variables de conception ne sont pas directement accessibles. Le choix de ces modèles est le fruit d'un compromis entre le temps d'exécution et la précision des résultats, en sachant que la prise en compte de tout détail superflu est inutilement pénalisante (le mieux étant l'ennemi du bien).

Par ailleurs, l'évaluation d'une solution peut être menée de manière plus ou moins grossière, en fonction de l'estimation courante de son potentiel. Rejetant au plus tôt les mauvaises solutions, ce principe permet de réduire d'au moins un ordre de grandeur les temps de calcul lorsque l'évaluation résulte d'une simulation de Monte-Carlo, par exemple.

5 CONCLUSION

Déjà utilisée pour améliorer localement la forme des avions, l'optimisation multidisciplinaire fait actuellement l'objet de recherches en aéronautique, informatique et mathématiques appliquées. Elle ne peut cependant se contenter de rechercher des optima locaux si elle veut faire émerger des solutions réellement innovantes et contribuer, par exemple, à la conception des avions respectueux de l'environnement et moins gourmands en énergie qu'attendent les usagers de l'aéronautique civile.

Comme l'illustre l'application dans le domaine des drones présentée ici, les gains apportés par l'optimisation hybride aux performances d'un avion ne se chiffrent pas en pourcents mais en dizaine de pourcents résultant notamment de la suppression de marges cachées et d'opportunités offertes par une moindre sensibilité aux effets de seuil. Outre l'aéronautique, ces techniques sont susceptibles d'apporter des gains similaires dans de nombreux domaines d'application, dont la conception automobile, navale, ferroviaire, informatique, etc.

D'importants travaux restent néanmoins à faire pour que des outils d'optimisation hybride multidisciplinaire véritablement efficaces voient le jour et contribuent à la conception des produits de demain.

Disposant d'une double compétence tant dans le domaine des algorithmes d'optimisation que de la conception de systèmes complexes, les auteurs de cet article sont prêts à s'associer à des chercheurs en mathématiques appliquées, informaticiens et

autres concepteurs, dans le cadre de projets collaboratifs de recherche et développement en MDO. Un projet d'une telle ambition peut difficilement se concevoir au sein d'une petite entreprise, même si des travaux exploratoires y ont été déjà menés jusqu'à la rédaction de la spécification préliminaire d'un nouvel outil (Gencab Indra) [Cabarbaye, 2016] [Cabarbaye, 2019].

Par ailleurs, le drone Cabiota devrait être utilisé comme démonstrateur de briques technologiques pour améliorer les performances des drones à aile fixe, dans le cadre d'un partenariat avec un laboratoire scientifique. Il pourrait, lui-même, bénéficier des fruits d'un tel projet de R&D.

6 REFERENCES

- Vanderplaats, G. N. (2007) *Multidiscipline Design Optimization*, Vanderplaat R&D, Inc.
- Cramer, E.J., Dennis Jr., J.E., Frank, P.D., Lewis, R.M., and Shubin, G.R. (1994) *Problem Formulation for Multidisciplinary Optimization*, *SIAM J. Optim.*, 4 (4): 754-776.
- Braun R., Gage P., Kroo I., and Sobieski I. (1996) *Implementation and performance issues in collaborative optimization*, Technical Report NASA-AIAA-96-4017, AIAA.
- Martins J. R. R. A., Hwang J. T. (2013) Review and unification of methods for computing derivatives of multidisciplinary computational models. *AIAA Journal*, 51(11):2582–2599
- Tedford N. P., Martin J. R.R.A. (2010), *Benchmarking multidisciplinary design optimization algorithms*, *Optim Eng* 11: 159–183, DOI 10.1007/s11081-009-9082-6.
- Zhoujie Lyu, Gaetan K. W. Kenway, and Joaquim R. R. A. Martins (1994), *Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark*, Helicopter Theory, Wayne Johnson, Courier Corporation.
- Nilesh P., Salunke, Juned Ahamad R. A., Channiwala S.A. (2014) *Airfoil Parameterization Techniques: A Review*, Science and Education Publishing.
- Cabarbaye A., Verdu T., Garcia F., Gorraz M., Bustico A., Bronz M., & Hattenberger G. (2019) Design of a high performance MAV for atmospheric research, IMAV
- Cabarbaye A., Cabarbaye A., Cabarbaye A. (2016) Une plateforme open source d'optimisation générique hautement distribuée, ROADEF.
- Cabarbaye A., Lautheret R. (2006) Couplage entre optimisation et simulation stochastique, ROADEF.
- Cabarbaye A., Cabarbaye A., (2019) L'optimisation multidisciplinaire l'autre enjeu du big data, ISBN : 979-10-97287-05-4.

Improve idea generation process by using ideas graph representation compared to Linkography method

ABDELLATIF DKHIL^{1,2}, HOUCINE DAMMAK¹, MAROUANE MOUATASSIM¹, MICKAEL GARDONI^{1,3}, REMY HOUSSIN³

¹ÉCOLE DE TECHNOLOGIE SUPÉRIEURE.
1100 Rue Notre-Dame Ouest, Montréal, QC H3C 1K3, Québec, Canada.
First name.last name.1@ens.etsmtl.ca, mickael.gardoni@etsmtl.ca

²INSTITUT SUPÉRIEUR DES SCIENCES APPLIQUÉES ET DE TECHNOLOGIE DE GAFSA
Campus Universitaire Sidi Ahmed Zarrouk BP 116 - 2112 Gafsa, TUNISIE.
abdellatif.dkhil@issatgf.rnu.tn

³INSTITUT NATIONAL DES SCIENCES APPLIQUÉES,
24 Boulevard de la Victoire, 67000 Strasbourg, France
First name.last name@insa-strasbourg.fr

Abstract – Linkography has become an established method for studying the idea generation step in the innovation process. In this method, the progress of the ideation session is being observed through the analysis of the graphical representation system called Linkograph. The main objective of this paper is to improve the use of the Linkography method to study the idea generation step. We propose to improve the graphical representation system by using graph theory for visualizing the ideation process. The use of ideas graph allows to improve the analysis possibilities, the communicative aspects and favoring the expansion and development of new ideas.

Keywords –Ideation, ideas graph, Linkography, linkograph

1 INTRODUCTION

It has become widely accepted that business survival and prosperity are strongly related to the ability to innovate. Creative ideas are essential for the generation of original ideas ultimately leading to innovation. Ideas generation is the fundamental step of the innovation process (Sozo & Ogluari, 2019) (Glier et al., 2011). Ideators teams from different areas of expertise are invited to work together during ideation sessions to exchange and create knowledge, usually for a specific goal. An ideation session purpose is to set an environment and implement creativity methods that will help teams to produce, express and combine ideas.

In the literature, research works concerning ideas generation step can be classified into two approaches: Results Oriented Approach and Process Oriented Approach (R. van der Lugt, 2005).

According to Van-de-Lugt (R. van der Lugt, 2005), the Results Oriented Approach is popular thanks to its robustness and its academic credibility. A Substantial research has been conducted to study the result of ideas generation session (J. Shah et al., 2000), (Howard et al., 2011). In this Results Oriented Approach, the majority of methods available for assessing ideas generation session compare the outcomes of different experimental treatments, usually the quantity and/or quality of the generated ideas. Numerous proposals have been made for measuring quantity and/or quality of the ideation step. This research work present different metrics for evaluating the results or outcomes from ideas generation techniques. Shah et al.(J. J. Shah et al., 2003), proposed four separate effectiveness measures: novelty, variety, quality and quantity.

The Process Oriented Approach is not interested in the results of the ideas generation step but in the process allowing their

development. This process is often presented as a Black Box. If we are interested in the ideation step, it is more informative to investigate the ideation process itself, rather than investigating its results (R. van der Lugt, 2005). The results-oriented approach leaves a rich source of information unutilized by discarding the ideation process itself as an object of research. Compared to the results-oriented approach, the study of Ideation process is least studied in the literature. According to Van der Lugt, this is explained by the complexity of the process study. All research works adopting this approach process are based on controlled laboratory experiments. These experiments operate according to a predefined protocol (idea generation method, participant, time, etc.) while monitoring the execution of the tasks. The research work proposed in the literature, relating to the oriented approach process, can be classified into two families (R. van der Lugt, 2005): (1) Work from cognitive science and (2) those derived from design theory.

In the field of design theory, the Linkography has become an established method for visualizing and analyzing ideation processes (Blom & Bogaers, 2018), (Hatcher et al., 2018). Originally developed by Gabriela Goldschmidt to evaluate the creative processes of individual designers (Goldschmidt, 1990), (Goldschmidt, 2016). Many research's work adapted the Linkography method was further developed as a method for assessing the idea generation process by means of investigating the connections between produced ideas (R. van der Lugt, 2005). Van-der- Lugt research work, (R. van der Lugt, 2005), (V. Lugt, 2003) have enabled Linkography to be effectively applied to the evaluation of group ideation in a product design and design engineering context through links that emerge as ideas are developed.

Several authors discuss the limits of the Linkography as developed by Goldschmidt's. The analysis and the interpretations of the Linkography representation, called linkograph, must be improved. The authors study the possibility of using additional tools to acquire more information from the Linkograph. They propose different metrics for Linkography analysis. There are studies which used sophisticated mathematical tools to Linkography analysis, such as Kan and Greo (Kan & Gero, 2020), (Kan & Gero, 2005) .

The main objective of this paper is to improve the Linkography method for studying the idea generation step. We propose the use of graph theory for the ideation process representation. This paper is organized into four sections. Following this introduction, we briefly introduce the Linkography method. Then, the disadvantages of using Linkograph representation are given in the second section. In the third section, we discuss the advantage of using ideas graph representation system. Finally, the fourth section presents our closing remarks, concluding in possible directions for future research.

2 LINKOGRAPHY METHOD FOR IDEATION PROCESS STUDY

2.1 Introduction to Linkography method

Linkography was first introduced to protocol analysis by Goldschmidt (Goldschmidt, 1990) to assess design productivity of designers. It is a method for recording, visualizing and analyzing connections between actions, ideas or events. The design protocol is decomposed into small units called "design moves". Goldschmidt, (Goldschmidt, 1995), (Goldschmidt & Tatsa, 2005), (Goldschmidt, 2003) (Goldschmidt, 1992), defines a design move as "a step, an act, an operation, which transforms the design situation relative to the state in which it was prior to that move" (Goldschmidt, 1996). By linking the related moves, a graphical network, called Linkograph, is then constructed. A linkograph can be seen as of associated moves that represents the design session. As Linkography gains recognition and popularity, it is increasingly extended in a variety of research contexts. For describing ideation process, Linkography was adapted and developed by studying links between ideas rather than design moves. Now, in representing the ideation process with Linkography, the variables must therefore be not the ideas, as in most studies, but the links among them. In this context, this method is not interested in the quality of the outcome, but by analysing the linking among them it purports to: '...be instrumental in comprehending structural patterns of design reasoning'(Goldschmidt, 1990). The progress of an ideation session is made observable through the analysis of connection between ideas. **Figure 1** gives a description of application of Linkography method steps.

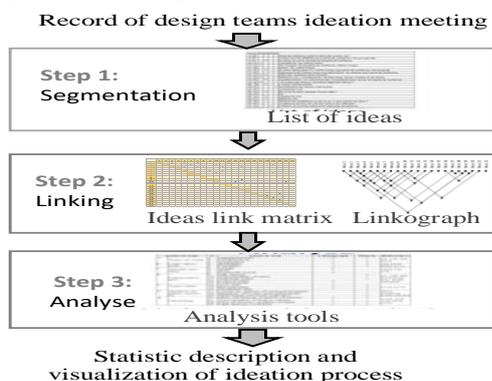


Figure 1- Application steps of Linkography method

The first step allows the identification of ideas listed by segmentation of ideation protocol. The ideation session is decomposed by parsing the recorded ideation meeting protocol into small units corresponding to the generated ideas. After segmentation step, the links between ideas are determined, based in ideation meeting recording. For each identified idea in the step 1, the existence, or non-existence, of links with each of earlier ideas is determined. Then the link between ideas will be generated in a matrix format. Ideas link matrix permits the generation of Linkography representation, called linkograph. These representations enter links among ideas as an input and displays structural idea reasoning patterns as an output. The ideation process can then be looked in terms of the patterns in the linkograph, which display the ideation reasoning and the structure of process. Finally, in the third step, the connections between ideas are analyzed to provide information, investigate the structure, analyze ideation session and compare productivity (Dkhil et al., 2018).

2.2 Limits of Linkography representation system

Linkograph is a graphical representation system of an ideation session which uses ideas links (matrix) as input and displays structural ideas reasoning patterns as output (see **Figure 1** step1). The process of ideation session can be seen as a linkograph that traces the associations of every generated ideas. Mapping these links should reveal something about the nature of interconnectivity among ideas. By using this graphical representation, Goldschmidt wanted to derive at an instrument that would reveal structural patterns of ideation (or design) reasoning (Goldschmidt, 1990), (Goldschmidt, 1992). **Figure 2** shows an example of Linkograph obtained from the ideas link matrix. In this representation, the ideas list is showed on a horizontal line and the links are nodes at the intersections of diagonal network lines connecting related ideas. As shown in **Figure 2** for example, the ideas number 2 has links with idea 6, idea 10 and idea 19, while idea 15 has no links at all. According to Goldschmidt, the Linkograph is a graphic system where a horizontal ideas-line note in sequential order. Link-lines departing from ideas or symbols are draw diagonally underneath the ideas-line at 45° to it. Each idea generates tow linklines: A leftward line for backlinks and a rightward line for forlinks. Only the backlinks are identified in the analysis of the idea's session protocols.

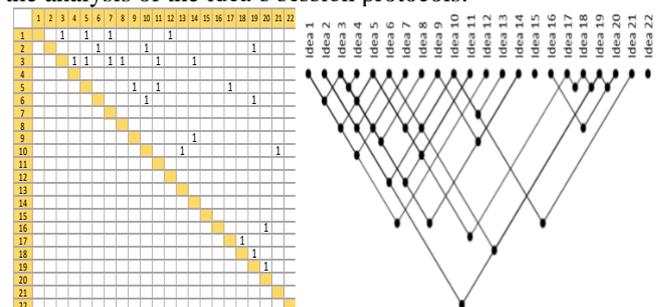


Figure 2- ideas links matrix and Linkograph

In the Linkograph, the connections between ideas are analyzed to provide information about the ideation session, investigate the structure of idea generation processes and compare ideation session productivity (R. van der Lugt, 2005), (Goldschmidt & Tatsa, 2005).

The distribution of links among ideas reflects the structure of ideation session and the status of ideation reasoning. Goldschmidt and Van-der-Lugt suggested that the Linkograph pattern of productive ideation creative team will be different from that of less productive group. Productive team will

produce ideas that have a high potential for connectivity to other ideas. Less productive team will have more random paths with ideas that did not have a high potential for contribution to problem solving.

As we mentioned above, Linkography has become an established method for visualising and analysing ideation and design processes but some limits exist (Roozenburg, 2016). These limits can be classified in three main categories: the first concerns the limited analysis possibilities, the second is related to the communicative limitation and the last concerns the valorization and the reusing of analysis results.

a. **Limit 1** - The analysis possibilities: The availability of tools and metrics for the structural ideation process analysis is the biggest limitation of Linkography approach: this is due to the representation way. Several authors study the possibility of using additional tools to acquire more information from the Linkograph, (Kan & Gero, 2017), (Hatcher et al., 2018).

b. **Limit 2**- Communicative limitations: Using linkograph to describe the structure of idea generation processes and the result of analysis present communicative limitations for ideation teams: for example, ideas generation sequence is not clearly apparent. This representation system has a limitation in describing the entire structure of ideas with multiple links. The relationship between ideas are sometimes difficult to interpret by novice users. If the linkograph are generated wrongly, it many distort the important relationship and the linkograph can be difficult to read. Using this visual support to communicate analysis result is difficult.

c. **Limit 3** - Valorisation and reusing: the reuse of linkograph in other ideation session as work support is difficult. For example, expansion and introducing of new ideas is difficult.

3 APPROACH

As we concluded in last section, the limitations of Linkography method are mainly due to the representation system. If we improve the representation system, we will be able to improve (limit 1) The analysis possibilities, (limit 2) The communicative aspects and (limit 3) The valorisation and the reusing.

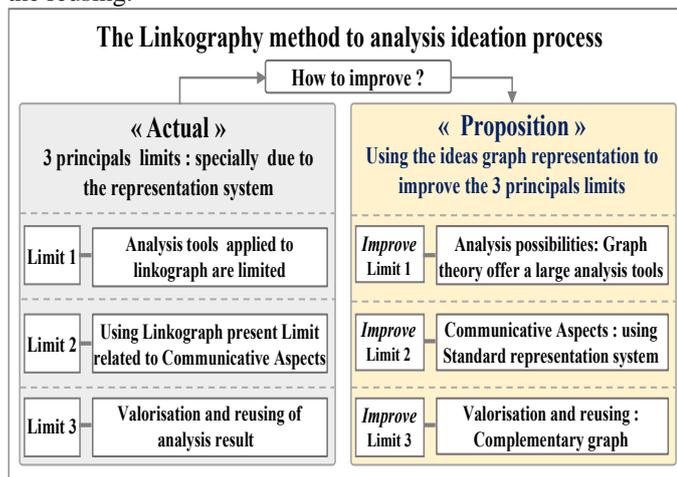


Figure 3 - proposed approach to improve Linkography

The representation system can improve the information extraction and interpretation; can help visualizing complex relationships and resolving ambiguities. Providing interaction not only increases the satisfaction for the user, but also supports the understanding of relationships. Yi et al. (2007) (Yi et al., 2007) identify, how humans interact with a visualization in order to extract information and classified seven interaction categories: select, explore, reconfigure,

encode, abstract, filter, connect. The authors conclude, on the one hand, that interaction supports user perception and overcomes limiting elements of certain forms of representation. On the other hand, they describe that representation and the interaction with the representation cannot be considered separately.

In order to improve the representation system of Linkography we propose to use graph theory. Figure 3 shows the actual limits of Linkography method and our proposition to improve these limits by using ideas graph representation.

3.1 Motivation and some basic concept of graph theory

In all fields where the connectivity between objects (structures, entities, ...) plays a major role, the graph theory finds a privileged place. The drawing and the analysis of graph is an extremely active area of research given the challenges it represents and the variety of application fields it covers: web, knowledge maps, communication, ontologies, artificial intelligence, etc.

The interest of graphs theory emerges when we interest to the problem of a topological nature, such as problems of paths, connectivity or structural reliability, as well as in relationship theory and data analysis (Diestel, 2012). We also find graphs in the study of problems related to the relationships between different entities (dependency or influence graphs) and also in the study of possible evolutions of a system. This explains our proposed approach to use the graph theory to support the study of the ideation process in the Linkography approach.

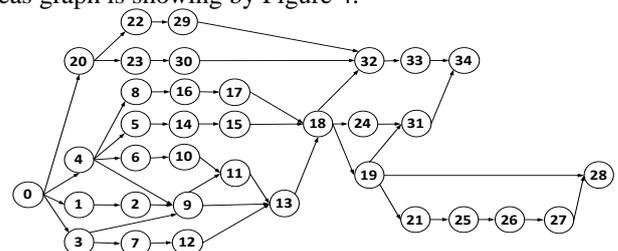
Node-link graphs are simple, powerful, and elegant abstractions that have broad applicability in many areas. Any domain that can be modelled as a collection of linked nodes can be represented as a graph. Table 1 Error! Reference source not found. gives the definition of some basic concepts of the graph theory.

Table 1- Basic concept of graph theory ((Gross & Yellen, 2000))

A graph $G = (V, E)$: a set of vertices V and a set of edges E . Each edge is a node pair (u, v) where $u \in E$ and $v \in E$.
Diagraph: a directed graph called Diagraph is oriented a graph $G (V, E)$ contains edges (u, v) where u is the start point and where v is the ending point of the edge
A path: In diagraph, A path is a sequence of vertices (terminal vertex of an edge in the sequence is the initial vertex of the next edge).
Cycle: is a path that starts and ends at the same vertex.
DAG: A Directed Acyclic Graph called a DAG is A directed graph with no cycles. E.g. All trees are DAG.
Order of graph: is the number n of nodes or vertex of graph
Adjacency matrix $A (n, n)$: n is the order of the graph and rows and columns are assigned to the set of vertices V . the presence of an edge between tow vertices is symbolized by a value (weighted, binar, ...).

3.2 ideas graph representation system

As we mentioned above, an ideas graph is a graph $G = (V, E)$ in which the nodes or vertexes (V) represent ideas, the edges (u, v) represent the relations, and sometime we can use weight or labels to represent the nature of the relations. An example of ideas graph is showing by Figure 4.



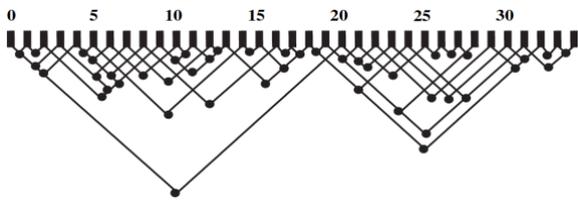


Figure 4- Ideas graph and corresponding Linkograph

In the next section we discuss, the use of ideas graph representation system to improve the Linkography approach.

4 IMPROVE LIMIT 1: ANALYSIS TOOLS FROM GRAPH THEORY AND THEIR UTILITY.

Several authors discuss the limits of the Linkography method as developed by Goldschmidt's, and study the possibility of using additional tools to acquire information (R. van der Lugt, 2005). They propose different metrics for improving analysis. There are studies that used sophisticated mathematical tools to analyze Linkography (Kan & Gero, 2017), (Kan & Gero, 2005), such as who used the x and y coordinates of the linking network of Linkography as the input for a cluster analysis to infer the underlying design activity. According to the literature, different tools of Linkography analysis has been proposed. These tools of Linkography analysis can be used to analyse the idea generation session. In our previous research work (Dkhil et al 2018,) we presented a bibliographic review of these analysis tools. According to Goldschmidt's, some specific patterns in linkograph have a low frequency of occurrence. This decreases the utility and the use of some proposed tools in analysis. In fact, the representation system used in the Linkography method had a great impact on the analysis possibility: limits the possibilities of analysis and the acquiring of information.

Graph theory offers a wide range of structural analysis tools in addition on its pictorial representation. These tools allow a quantitative analysis of connectivity and relationships between ideas. The graph-based representation will permit us to use classical graph theory tools and concepts. Our proposition is to use ideas graph representation in order to benefit from the wide range of analysis possibilities offered by graph theory.

This section aimed to present some examples of analysis tools issues from graph theory and their utility in the analysis of ideation process. In the following, we propose a representative example of the applications and the utility of some analysis tools from graph theory. To illustrate this example, we use the result of a real case study from an ideation session. The used case study corresponds to a brainstorming session to resolve a specific design problem kept anonymous for confidentiality reasons. The design team is composed from four participants. We use the Linkography method to identify the generated ideas and the link between them. The Linkograph given by Figure 5 correspond to generated ideas and the link between them correspond to the result of ideation process.

4.1 Linkograph, Topological sorted graph and linear ordering

In the graph theory, Topological sorting problem consist to find in digraph $G = (V, E)$ a linear or sequential ordering of vertices V such that: for all edges (u, v) in E , u precedes v in the ordering (Gross & Yellen, 2000) . A topological sort on a Graph $G (V, E)$ is an arrangement of its vertex so that all the edges point from left to right. Only graphs without cycles can be topologically sorted.

If we have a graph with cycle, the linear ordering tool allow to

find a sequential ordering of vertex with the minimum number of cycles or Feedback Arc Set (FAS). A diagraph without FAS is a directed acyclic graph (DAG) can to topological sorted. The linear ordering permit to identifying the minimum FAS.

Utility: The linkograph as proposed by Goldsmith correspond to a topological sorted directed acyclic graph. In the Linkograph, the ideas are presented in sequential linear ordering according to the generation time in the transcript. The Figure 5 shows the equivalence between the linkograph and the topological sorted diagraph.

The Linkograph is a graphic system where a horizontal ideas-line note in sequential order. According to Goldschmidt, Other limits of Linkograph representation that the linkograph cannot represent the feedback in the ideation process (see section 2.2). In a real ideation process a feedback link to previous ideas can exist. The great advantage of he using of Topological sorted graph is allow to shows the feedback link. We can represent the graph according to the transcript time line. Also, we can use the Total ordering tools to finding, the ideas flow with the minimum feedback. The linkograph with feedback arc set correspond to total ordering graph.

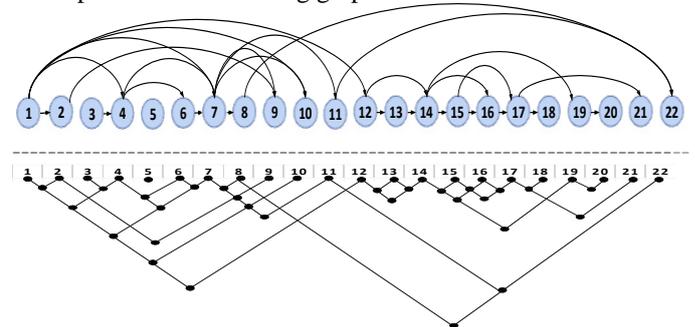


Figure 5- Example of topological sorted graph and corresponding Linkograph

4.2 Levels

As we mentioned above, the topological sorted graph is DAG where the vertex is in a sequential order: each idea correspond exactly to a specific one rank or level (number of levels correspond to the number of vertex). This DAG can be partially ordered by using Level tool:

Let $G (E, V)$ is a directed acyclic graph where V is the set of vertices and E is the set of links, the vertices can be partially ordered. For each vertex v , a rank $R(v)$ has to be calculated that specifies the level that vertex v belongs to. The level 1 is the top most level and correspond to vertex without incoming edges. To identify the level of vertex, we remove, at each iteration, all the vertices which have no incoming edges.

Figure 6 clearly shows the ranking by level of each vertex. The graph representation is established from the topological sorted graph given by Figure 5.

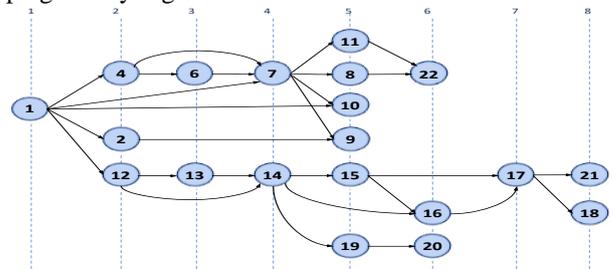


Figure 6 - Ranking by level of DAG

Utility: The decomposition of directed graph into level is very useful for understating the structure of ideas generation session. It highlights a “main” flow in the idea generation session. Decomposition of DAG into levels give information

about the main ideas flows. This representation showed notifiable differences in exploration of precedents, and generation and development of new ideas. Precedence relationship among ideas is clearly visible. The ability to distinguish and identify the ideation process.

This analysis tool can be measured by the vertices level corresponding to the number of vertices per level

$$\text{Level rate} = \frac{\text{Number of level}}{\text{total number of ideas}} (\%)$$

This is interesting during the ideation session analysis;

- If the level ratio is high: this informs that the rate of divergence of ideas is low. The creative team builds one idea on top of other in sequential order: the number of ideas produced from an initial idea is small.
- If the level ratio is low: this informs that the rate of divergence of ideas is high. The ideation process corresponds to a divergence structure with a low a sequential build order.

In the example given by Figure 6, Level rate is equal to 36.36%. this level rate is low and this informs about the ideation process structure: a height level rate informs about a divergence structure of ideation process.

4.3 The degree of a vertex and adjacency list

A graph consists of vertices and a set of edges connecting them. The Degree of vertex V denoted by $d(v)$ is the sum of the incoming or incident edges to vertex V, denoted by $d^-(v)$, and the outgoing edges from V denoted by $d^+(v)$. $d(v) = d^-(v) + d^+(v)$. The degree of the vertex v is calculated from the adjacency list or the neighborhood of vertex. Two vertices u and v are called adjacent or neighbors if u and v are endpoints of an edge e of $G = (V, E)$.

- Utility: To analyze the ideation process, it is important to look at the degree of a vertex. The degree of vertex informs about the criticality and the connectivity of the ideas in the ideation process. Critical ideas are those that have a significant number of links to other ideas. The Table 2 give the degree of each vertex in the graph.

Table 2- degree of vertex in the graph

IDEAS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
$d^-(v)$	0	1	0	2	0	1	3	1	2	2	1	1	1	2	1	2	2	1	1	1	1	2
$d^+(v)$	5	1	1	2	0	1	4	1	0	0	1	2	1	3	2	1	2	0	1	0	0	0
$d(v)$	5	2	1	4	0	2	7	2	2	2	2	3	2	5	3	3	4	1	2	1	1	2

In this example, ideas number 7, 1 and 14 are critical ideas which have the Hight number of links to the other ideas. The critical can be identified in the graph. Figure 7 shows an example of a graph and the degree of each vertex.

In any case study example, we can classify the ideas by the degree of connectivity.

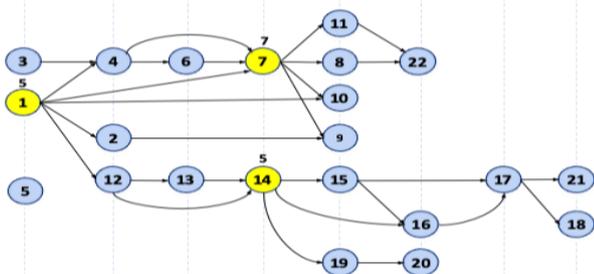


Figure 7 – the critical ideas in the DAG

4.4 Source

A source is a vertex V without incoming edges or predecessor: $d^-(v) = 0$. As show in Figure 8, a directed graph can have one single source (see Figure 8 (a)) or a multiple sources vertex (

see Figure 8 (b)). One single source vertex in directed acyclic graph (DAG) is called root and this special kind of DAG is called an arborescence. In the arborescence, for a vertex u called the root and any other vertex v, there is exactly one directed path from u to v. A tree is a very general structure that can have multiple sources vertex or roots. We can have a directed tree without requiring a single source or root. Every arborescence is a directed acyclic graph (DAG), but not every DAG is an arborescence. A multi-root DAG can be identified.

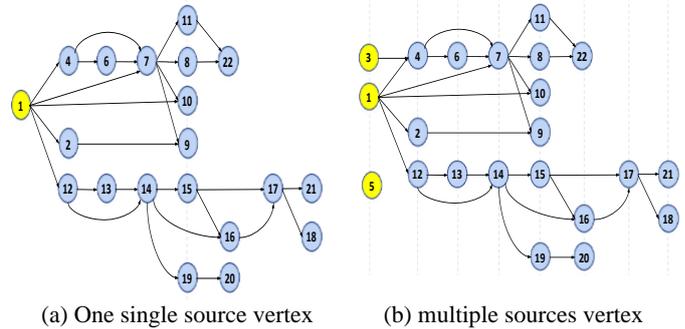


Figure 8- Source vertex in Diagram

- Utility: The root or the single source vertex is the ideas how serves as point of reference for other ideas in the tree or arborescence. In ideation process, we usually keep the root at the top and list other ideas below it. It is the initial idea that gives birth to all other ideas.

There is a difference between a graph with one single source or a multiple sources vertex. The one single source is the parent of all other ideas. From This parent ideas can converge multiple child's with a Hight level of similarity: this ideation process is more homogeneous than the process with process with a multiple source's vertex. This last process is more heterogeneous in terms of ideas variety.

A depth search traversal of the tree starts at the sources, plunges down the leftmost path, and backtracks only when it gets stuck, returning to the source ideas at the end. The depth gives information about the development of the ideation process from the sources. The depth of an ideas x in the tree (or arborescence) is the length (number of ideas) of the (unique) path from the source to that ideas. The Longest common subsequence of ideas in graph dominant forward ideation flows can be identified.

4.5 Sink

A vertex V is called a sink vertex if no outgoing edge exists: $d^+(v) = 0$. Is also called a terminal vertex. In the Figure 9, ideas number 22,10, 9, 21, 18 are a sink

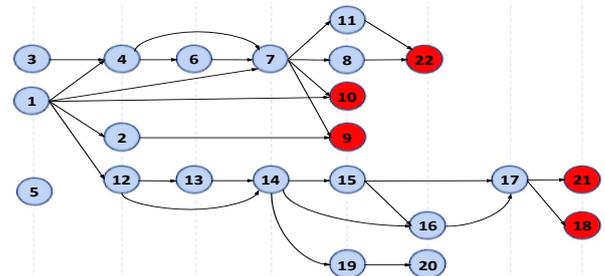


Figure 9- Sinks vertex in the graph

- Utility: Finding the sink in the graph allow us to identify the abandoned ideation path and this for different reasons. In future ideation sessions, sink can be explored to produce and develop new ideas.

4.6 Singleton

The singleton graph is the graph consisting of a single isolated vertex with no edges. It is, therefore, the empty graph on one

node. In Figure 9, the idea number 5 is a singleton.

- **Utility:** In the ideation process, a singleton is an idea which was generated but too quickly abandoned and this for lack of technology, knowledge or expertise, other reasons. The rate of singleton in the graph inform about the Ideational Fluency: ability to rapidly produce a series of related ideas. This rate allows the analysis of ideas flow: difficulties to get a fluid reasoning.

4.7 Cut vertex and cut edge

A node in the graph G is called cut point (or articulation point) if its removal disconnects a graph, i.e. increases the number of components. Also, it makes some points unreachable from some other (Diestel, 2012). Articulation Point (or Cut Vertex) correspond to any vertex whose removal (together with the removal of any incident edges) results in a disconnected graph. In graph G , a Bridge or cut edge is an edge whose removal results in a disconnected graph. An edge is a bridge if its removal results in disconnected sub-graph. A bridge is an edge such that the graph containing the edge has fewer components than the subgraph that is obtained after the edge is removed. Figure 10 shows an example of cut vertex and cut edge.

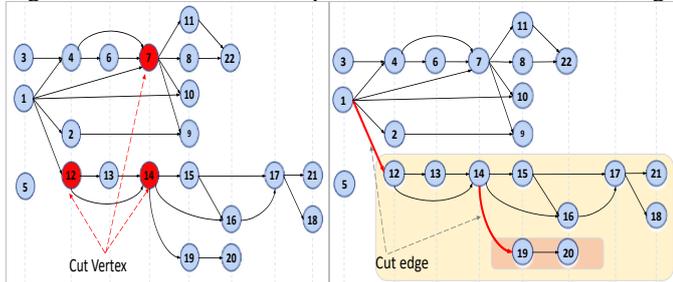


Figure 10 - Example of cut edge and cut vertex

- **Utility:** A cut edge (or cut vertex) is such that its deletion increases the number of connected ideas.

A graph is biconnected if it contains no articulation points. Biconnected graphs, articulation points, bridge are of great interest in the analysis of ideas graph because these are the “critical” points, whose failure will result in the network becoming disconnected. Also, the biconnected components of a graph are the equivalence classes or cluster (see section 4.1). On the other hand, if a connected graph does not contain cut vertices (or cut edge), it means that the graph is “dense” and its decomposition into cluster will be difficult.

4.8 Graph node groups (Collapse/ contraction)

Collapsing graph is an alternative way to reduce visual complexity. Collapsing means removing from the visualization the vertexes that are connected to one node or to a group of vertexes. Any number of vertexes can be collapsed into a single synthetic node: collapse set of vertexes and expand it when needed. Such a synthetic node contains a user-provided text instead of normal disassembly listing.

- **Utility:** If a graph is too large to fit on the screen, groups of related vertexes are (clustered) collapsed into super-vertexes. The users see a “summary” of the graph, namely the super-vertexes and super-edges between the super-vertexes. Some clusters may be shown in more detail than others. The process collapsing involves discovering groups in the data. In the case of graph visualizing collapsing vertexes Groups can be uncollapsed to display the original node content.

5 IMPROVE LIMIT 2: USING IDEAS GRAPH TO IMPROVE THE COMMUNICATIVE ASPECT

As mentioned above, using linkograph has some communicative limitations because the relationships between ideas are sometimes difficult to interpret by novice users. For a novice user these very complex interconnected systems are a challenge. In the linkograph, the sequence of ideas generation is not clearly apparent compared to the graph representation. With linkograph it is difficult to retrace the path of ideas construction. This representation way has a limitation in describing the entire structure of ideas with multiple relationships. Compared to a linkograph, the ideas graph has the advantage to visualize the information in its simplest form. We propose to use graph representation in order to improve the communicative aspects. In the following section, we will discuss the different advantages:

5.1 Ideas graph is a simple and comprehensive visual representation language

Graph representation has long provided visual languages and have been widely used in many different disciplines as formal representation system. A directed graph is a highly abstract way of visualizing complexity in various applications. By letting the nodes represent ideas and the edges the links between them, one obtains the ideas graph, which is a very general and abstract representation for inter-connected systems. Ideas graph is a network defined as a composition of nodes and linkages which indicate the relationships among nodes. This representation provides a more lucid representation to understand the characteristics of each ideation process and identify the differences intuitively. The Ideas graph simplifies the representation and the relationship between ideas by using standard representation allowing a comprehensive structure of ideation processes. The ideas graph described in the best ways the richness and complexity of the ideation process. better understand the structure of the link between ideas.

5.2 Ideas graph improve communication and comprehension between ideation team and users

Ideas graph can serve the ideation team in understanding their own ideas better and simplify the representation and the relations of ideas by using standard representation. This standard visualization system should help to support the transdisciplinary communication and comprehension. Nevertheless, working in transdisciplinary teams within the ideation process requires a common understanding between all teams. Ideas graph representation reduce difficulties in communication and comprehension and conflicts between the respective disciplines arise.

Ideas graph are effective tool for making the structure of ideas generation process explicit because is essential for the ideation team to clarifying their thinking and verifying the establish links. Organizing and prioritizing new ideas, thus stimulating creative thinking allowing for understanding. By improving the representation of ideation process, working together in transdisciplinary teams can save time and therefore improve the generation of new ideas. The visualization has to be comprehensible for interdisciplinary teams, with easily understandable representation system.

5.3 Ideas graph offer a large visualization option:

The field of graph drawing has been flourishing very much in the last two decades. Recent progress in computational geometry, topological graph theory, and order theory have considerably affected the evolution of this field, and have widened the range of issues being investigated. For decades, researchers in information visualisation and graph drawing

have focused on developing techniques for the layout and display of very large and complex networks. The field of graph drawing offers a Different visualization options and functionality: Table 3 shows some examples of representation options and functionality.

Table 3 : examples of visualization options and functionality

<u>Reduce complexity, choosing the detail level in representation:</u> different abstract level: filtering, collapsing, grouping nodes, reduce link.
<u>Large Graph-drawing algorithm:</u> we can choose an efficient layout algorithm offering legible drawings. These algorithms support well-arranged representations of nodes and edges within the graph (drawing quality with minimal edge-crossings, overlapping edges, ...). Many different classification can be used: according to the drawing types (dynamic or static graph), drawing support (2D or 3D), Size of graph (Small, large, hypergraph), data Format (numeric or text format)
<u>Navigation and interaction:</u> user can interact with the graph and navigate through the visualization thus gain detailed information: Rotating, translating, and zooming into a visualization leads to changing viewpoints and causes to a different focus in information retrieval.

6 IMPROVE LIMIT 3: USING IDEAS GRAPH TO IMPROVE THE VALORIZATION AND THE REUSE OF RESULTS

As we mentioned above, Ideas graph is an effective tool for making the structure of ideas generation process explicit because is essential for the ideation team to clarify their thinking and verifying the established links. Ideas graph can serve the ideation team to understand their own ideas better and simplify the representation and the relationship of ideas by using standard representation. Ideas graph can also be reused in other ideas generation meeting to stimulate the process of ideas generation.

6.1 Ability to more easily receive and connect new emerging ideas.

In new ideas generation session, the ideas graph representation can be used for several purposes such as to generate ideas or brainstorming, to develop complex structures, to communicate complex ideas. All these purposes correspond to a Visual learning for the ideation team (Dkhil et al., 2018). In fact, the ideas generated by others serve as external stimuli. The purpose of idea expansion is to expand a set of ideas by modifying or transforming existing ideas. Creative imagination is often structured by prior knowledge.

The major utility of this visual support is the ability to receive and connect new emerging ideas. Compared to the linkograph, in this representation system a new idea can be easily added to the developed (previous) or the under-development ideas graph. With the ideas graph representation, it is more easily to receive and put in relation the new emerging ideas: this is explained by a more understandable structure and by the labelling of ideas. Thus, the Ideas graph is presented as a real creative platform on which it selects and visually groups the latest ideas. Creative teams will be able to organize and prioritize additional ideas, thus stimulating creative thinking allowing better understanding.

Compared to linkograph, ideas graph showed notifiable differences in exploration of precedents, generation and development of new emerged ideas.

6.2 Valorization of previous ideation session result by using the complementary graph:

By using ideas graph it is more easily to retrace and understand

the result of previous ideation process, even if the creative team changes. This allowing the capitalisation of the previous result in order to reuse them for Supporting Divergent Thinking in another ideation session.

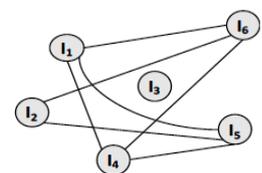
In addition, from previous ideas graph a complementary graph can be generated.

In the graph theory, the complement of a graph is sometimes called inverse of a graph or the edge-complement, (Sampathkumar & Pushpalatha, 1998), (Gross & Yellen, 2000), and denoted G^c (Clark & Entringer, 1983). The complement of graph $G(V, E)$ is a graph $G^c(V, E)$ on the same vertices V such that two distinct vertices (V_i, V_j) of G are adjacent if and only if they are not adjacent in G^c . In the matrix representation of graph $G(V, E)$, the presence of a link between two vertices V is given by a numerical value (Gross & Yellen, 2000). The matrix of complement graph $G^c(V, E)$ is matrix with the same number of vertex but two vertex is adjacent in this matrix if and only if they are not adjacent in ideas link matrix of initial graph.

The complementary graph represents, here, the tool which will force the connection between the ideas as in the Bissociation method. In the ideation process, The Bissociation method puts in relation ideas identified randomly to create new ideas (Dubitzky et al., 2012)). Complementary graph allows the use the same principle of Bissociation but the combined ideas are the previous proposed ones (Dkhil et al., 2018). The Bissociation or the Forced connection between the previous ideas will be established by using complement graph or complement ideas link matrix. This complement of graph makes it possible to associate or combine two or more ideas, sometimes not compatible, in order to create a third, original one. This complement graph can be seen as a particular form of Bissociation. This particular form cannot be applicable using the linkograph method.

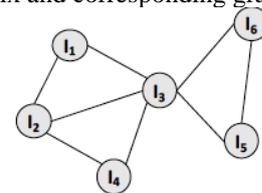
The analysis of this complementary graph should bring out endogenous information by making it possible to identify different relational aspects, favoring the expansion and development of new ideas. The forced relationship is established between two ideas through complementary matrix in a Bissociation process, which gives birth to a creative idea. By “forcing” connections between two unrelated ideas, new ideas could emerge. The complementary matrix, given by the Figure 11(b), is a double entry table allowing the crossing of two ideas from the first ideation session. The boxes with “1” aim to engage the imagination through forced encounters between different ideas. All allow us to force our brain to get out of ready-made thought patterns built since our childhood and to explore cross-roads that will force us to think differently.

	I_1	I_2	I_3	I_4	I_5	I_6
I_1		1	1	0	0	0
I_2			1	1	0	0
I_3				1	1	1
I_4					0	0
I_5						1
I_6						



(a) Initial Ideas link matrix and corresponding graph

	I_1	I_2	I_3	I_4	I_5	I_6
I_1		0	0	1	1	1
I_2			0	0	1	1
I_3				0	0	0
I_4					1	1
I_5						0
I_6						



(b) Complement of Ideas link matrix and complement graph

Figure 11- Example of initial graph and its Complement

7 CONCLUSION

In this paper, we propose the use of graph theory to improve the Linkography method for studying the idea generation step. The proposed approach presented solutions for three limits of the Linkography method: (1) The analysis possibilities, (2) The communicative aspects, and (3) The valorization and the reusing. Our main goal is to help ideators in their creativity session. This approach could also serve as a framework in commercial innovation solutions and for industrial applications as well. Further work will consist of the rationalization of a method for analysis and visualization of the ideation process based on the theory of graph.

8 REFERENCES

- Blom, N., & Bogaers, A. (2018). Using Linkography to investigate students' thinking and information use during a STEM task. *International Journal of Technology and Design Education*. <https://doi.org/10.1007/s10798-018-9489-5>
- Clark, L., & Entringer, R. (1983). Smallest maximally nonhamiltonian graphs. *Periodica Mathematica Hungarica*, 14(1), 57–68. <https://doi.org/10.1007/BF02023582>
- Diestel, R. (2012). *Graph Theory*, 4th Edition (Vol. 173). Springer.
- Dkhil, A., Gardoni, M., Belgacem, L., & Houssin, R. (2018). Linkographic Analysis of Design Ideation Session: Idea Graph Representation and Additional Tools for Analysis. In P. Chiabert, A. Bouras, F. Noël, & J. Ríos (Eds.), *Product Lifecycle Management to Support Industry 4.0* (pp. 715–725). Springer International Publishing.
- Dubitzky, W., Kötter, T., Schmidt, O., & Berthold, M. (2012). Towards Creative Information Exploration Based on Koestler's Concept of Bisociation. In *Bisociative Knowledge Discovery* (pp. 11–32). https://doi.org/10.1007/978-3-642-31830-6_2
- Glier, M., Schmidt, S., Linsey, J., & Mcadams, D. (2011). Distributed Ideation: Idea Generation in Distributed Capstone Engineering Design Teams. *International Journal of Engineering Education*, 27, 1281–1294.
- Goldschmidt, G. (1990). Linkography: Assessing design productivity. *Cybernetics and System '90*, 291–298.
- Goldschmidt, G. (1992). *Criteria for Design Evaluation: A Process-Oriented Paradigm*. New York: John Wiley & Sons, 1992. Pp. 67-79. Includes Bibliography.
- Goldschmidt, G. (1995). The designer as a team of one. *Design Studies - DESIGN STUD*, 16, 189–209. [https://doi.org/10.1016/0142-694X\(94\)00009-3](https://doi.org/10.1016/0142-694X(94)00009-3)
- Goldschmidt, G. (2003). Cognitive economy in design reasoning (pp. 53–62). https://doi.org/10.1007/978-3-662-07811-2_6
- Goldschmidt, G. (2016). Linkographic Evidence for Concurrent Divergent and Convergent Thinking in Creative Design. *Creativity Research Journal*, 28, 115–122. <https://doi.org/10.1080/10400419.2016.1162497>
- Goldschmidt, G., & Tansa, D. (2005). How good are good ideas? Correlates of design creativity. *Design Studies*, 26(6), 593–611. <https://doi.org/10.1016/j.destud.2005.02.004>
- Gross, J., & Yellen, J. (2000). *Graph Theory and Its Application*. *The Mathematical Gazette*, 84. <https://doi.org/10.2307/3621555>
- Hatcher, G., Ion, W., Maclachlan, R., Marlow, M., Simpson, B., Wilson, N., & Wodehouse, A. (2018). Using linkography to compare creative methods for group ideation. *Design Studies*, 58. <https://doi.org/10.1016/j.destud.2018.05.002>
- Howard, T. J., Culley, S., & Dekoninck, E. A. (2011). Reuse of ideas and concepts for creative stimuli in engineering design. *Journal of Engineering Design*, 22(8), 565–581. <https://doi.org/10.1080/09544821003598573>
- Kan, J., & Gero, J. (2005). Can entropy indicate the richness of idea generation in team designing? *CAADRIA 2005 - The Association for Computer-Aided Architectural Design Research in Asia: Digital Opportunities*, 451–457.
- Kan, J., & Gero, J. (2017). Quantitative Methods for Studying Design Protocols. In *Quantitative Methods for Studying Design Protocols*. <https://doi.org/10.1007/978-94-024-0984-0>
- Kan, J., & Gero, J. (2020). DESIGN BEHAVIOUR MEASUREMENT BY QUANTIFYING LINKOGRAPHY IN PROTOCOL STUDIES OF DESIGNING.
- Lugt, V. (2003). Relating the quality of the idea generation process to the quality of the resulting design ideas.
- Roozenburg, N. (2016). Linkography: Unfolding the Design Process. *Design Studies*, 42, 137–139. <https://doi.org/10.1016/j.destud.2015.10.003>
- Sampathkumar, E., & Pushpalatha, L. (1998). Complement of a Graph: A Generalization. *Graphs and Combinatorics*, 14(4), 377–392. <https://doi.org/10.1007/PL00021185>
- Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Design Studies*, 24(2), 111–134. [https://doi.org/10.1016/S0142-694X\(02\)00034-0](https://doi.org/10.1016/S0142-694X(02)00034-0)
- Shah, J., Kulkarni, S., & Vargas Hernandez, N. (2000). Evaluation of Idea Generation Methods for Conceptual Design: Effectiveness Metrics and Design of Experiments. *Journal of Mechanical Design - J MECH DESIGN*, 122. <https://doi.org/10.1115/1.1315592>
- Sozo, V., & Ogliari, A. (2019). Stimulating design team creativity based on emotional values: A study on idea generation in the early stages of new product development processes. *International Journal of Industrial Ergonomics*, 70, 38–50. <https://doi.org/10.1016/j.ergon.2019.01.003>
- van der Lugt, R. (2005). How sketching can affect the idea generation process in design group meetings. *Design Studies*, 26, 101–122. <https://doi.org/10.1016/j.destud.2004.08.003>
- Yi, J. S., Kang, Y., Stasko, J., & Jacko, J. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *Visualization and Computer Graphics, IEEE Transactions On*, 13, 1224–1231. <https://doi.org/10.1109/TVCG.2007.70515>

Session 3.2 - Session spéciale 10 :
Mesure de performance pour le
développement durable et la
responsabilité sociétale des
organisations

Chairman : Francois Vernadat - Lamia Berrah

CIGI 2021

BSC: a literature review and a roadmap towards its implementation in a healthcare network in the province of Québec

ANES BEN FRADJ^{1,2}, TASSEDA BOUKHERROUB^{1,2}, CLAUDE OLIVIER¹

¹ DEPARTMENT OF SYSTEMS ENGINEERING & NUMERIX LABORATORY, ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
1100, NOTRE-DAME WEST, H3C 1K3, MONTRÉAL, CANADA

² INTERUNIVERSITY RESEARCH CENTRE ON ENTERPRISE NETWORKS, LOGISTICS AND TRANSPORTATION (CIRRELT)
UNIVERSITÉ DE MONTRÉAL, PAVILLON ANDRÉ AISENSTADT, BUREAU 3520, 2920, CHEMIN DE LA TOUR, MONTRÉAL, QC, H3T 1J4, CANADA

ANES.BEN-FRADJ.1@ETSMTL.NET, TASSEDA.BOUKHERROUB@ETSMTL.CA,
CLAUDE.OLIVIER@ETSMTL.CA

Abstract - Performance management in the healthcare sector is a major challenge. Healthcare organizations might deal with many challenges related to services accessibility and quality as well as human and material resources management. This paper presents a roadmap for implementing the Balanced Score Card (BSC) in a large healthcare network in the region of Montreal, Canada (CIUSSS Centre-Sud-de-l'île de Montréal – CCSMTL). Our approach is inspired from the literature and the analysis of data collected in the organization by using the SWOT method. We show how the BSC can be integrated to the performance management model currently implemented in the organization, and how the BSC can contribute to improve its performance management.

Résumé - La gestion de la performance dans le secteur de la santé est un défi majeur. Les organisations de santé peuvent faire face à plusieurs défis liés à l'accessibilité et à la qualité des services ainsi qu'à la gestion des ressources humaines et matérielles. Cet article présente une feuille de route pour l'implantation du tableau de bord équilibré (TBÉ) dans un grand réseau de santé dans la région de Montréal, Canada (CIUSSS Centre-Sud-de-l'île de Montréal - CCSMTL). Notre démarche s'inspire de la littérature et de l'analyse des données collectées dans l'organisation à l'aide de l'outil SWOT. Nous montrons comment le TBÉ peut être intégré au modèle de gestion de la performance actuellement mis en place dans l'organisation et comment le TBÉ peut contribuer à améliorer la gestion de sa performance.

Keywords – BSC, Healthcare, Performance management, SWOT.

Mot clés – BSC, Secteur de la santé, Gestion de la performance, FFPM.

1 INTRODUCTION

Healthcare and social services in the province of Quebec, Canada, are provided by networks of institutions called CI(U)SSSs which refer to integrated (University) health and social services centres. CI(U)SSSs are public organizations under the supervision of the Ministry of Health and Social Services (MSSS). This project is conducted in Montreal region, where there are five CIUSSSs, which offer services in the West, West-Centre, South-Centre, North and East. The mission of these organizations is to deliver general and specialized health and social services to the whole population in the Island of Montreal specifically and in the province of Quebec. They also ensure that these services are accessible, effective, efficient and of high quality.

In the healthcare and social services sector, engineers are needed to contribute to the improvement of service delivery and client satisfaction as well as the implementation of an organizational culture based on innovation and efficiency. In this context, this project is conducted in collaboration with the South-Centre CIUSSS (CIUSSS Centre-Sud-de-l'île de Montréal) referred to as CCSMTL. In April 2015, 52 establishments (four hospitals and clinics, 17 residential centres for the elderly, 14 social rehabilitation centres, 10 community service centres, five centres for the physically handicapped people and two administrative offices) were merged and combined to create the CCSMTL. In short, the CCSMTL is deployed in 150 facilities and employs more than 17,000 employees and nearly 800 physicians. It manages a budget of more than CA\$1.7 billion to provide care to a territorial population of 311,000 people.

For performance management, the CCSMTL uses the Quality Performance Model (QPM). This model contains performance dimensions that provide the basis for proposing indicators and allows performance monitoring. This model is currently being implemented in the organization's network, but requires further improvements. The implementation of the QPM model is quite challenging for CCSMTL managers. First, the CCSMTL is a very large organization, providing 28 various services to a population of 311 000 people in 150 facilities. The implementation of the QPM must take place at different levels of decision-making (strategic, tactical and operational) and align the operational practices with the QPM strategic objectives. Due to the wide range of services offered, this deployment is not easy to achieve. Second, it is complex to translate the strategy into performance indicators because some sub-dimensions of the QPM are difficult to measure (e.g. mobilization, which is related to employees' performance). Third, in order to measure the overall performance of the organisation and make sound decisions, the managers need aggregated information that is easy to visualize and to analyse. Currently, there is no well-structured information system that supports this objective.

In this paper, we present how the Balanced Scorecard (BSC) approach can be used to support the QPM model and therefore improve performance management. Our approach is inspired from the literature, analysis of data collected in the CCSMTL, and our knowledge in the healthcare sector. We present the theoretical basis of our approach and a roadmap towards its implementation in the CCSMTL. The remainder of the paper is organized as follows: the next section presents the QPM model and challenges related to its implementation. Section 3 presents our literature review on the BSC. Section 4 presents our approach and roadmap for BSC implementation in the CCSMTL. Finally, section 5 presents our conclusions.

2 QUALITY PERFORMANCE MODEL (QPM)

The CCSMTL aims at implementing its "integrated Quality-Performance Model" (QPM) to meet MSSS's strategic plan (MSSS Strategic plan 2019-2023) (MSSS 2019) regarding healthcare system performance evaluation and to improve the performance of its network. The QPM includes four performance dimensions (or axes): customer, accessibility / quality, mobilization, and optimization. Each axis presents an aspect of the CCSMTL performance. The customer dimension reflects the organization's performance in terms of the satisfaction of its customers (people receiving the service or care, their loved ones, the community and the population). Accessibility / quality dimension is the organization's ability to safely meet the needs and expectations of the customers while ensuring accessible and continuous services.



Figure 1. Quality Performance Model of CCSMTL (CCSMTL 2017)

The two main dimensions in the QPM are accessibility / quality and customer. They represent the "True North" as called in the CCSMTL. In other words, these two dimensions represent a compass needle which indicates the right direction toward a better performance. Optimization and mobilization support the True North objectives. Mobilization is the use of skills and talents of all the human resources to collectively fulfil the organization mission. Optimization is the search for continuous improvement in the use of available resources to ensure, over time, services adapted to customers' needs. In addition, the QPM model presents a number of sub-dimensions within each dimension. We mention for example security and efficacy within accessibility / quality dimension; partnership and communication within customer dimension; collaboration and learning within mobilization dimension; efficiency and innovation within optimization dimension.

In order to implement the QPM model in the organization, the CCSMTL has installed a visual performance management system (i.e., a network of control rooms) including a strategic room, tactical rooms and visual stations. By definition, a control room is a dedicated space where decision-makers and employees meet regularly to take a look on the current situation and initiate discussions to improve future performance (Lagacé and Landry 2016).

The CCSMTL is driven by organizational values that animate the management philosophy and guide daily actions: first of all, the respect to all people (employees and customers), regardless of their skills and their differences. It encourages recognizing and celebrating the efforts, strengths and contributions of each individual and sets the rules for growth and productivity at work. Second, the commitment of the employees and the desire to persevere in their actions and fulfil their tasks. It encourages continuous improvement and the recognition of the responsibilities entrusted to each individual. Third, passion and the willingness to conceive new possibilities, to deploy them with energy and conviction in order to realize them with the primary concern of creating benefits for customers. Fourth, collaboration that is based on being open to others and sharing information. It supports creating positive and productive relationships that encourage the active participation of individuals, other teams within the organization or external partners. The QPM model presents a translation of these values and the vision into strategic dimensions. In fact, the model adopts a customer-focused approach and relies on operational excellence (accessibility / quality). Respect, commitment, passion and collaboration are integrated into the employees' skills (mobilization dimension) with which the strategic objectives should be achieved.

3 LITERATURE REVIEW

The BSC is a method developed in 1992 by Robert S. Kaplan and David Norton. It is the result of a re-examination of performance measurement systems (PMSs) that have placed too much focus on financial results solely. Based on this finding, Kaplan and Norton incorporated the term "Balanced" which emphasizes the concept of balance between all performance pillars (Kaplan and Norton 1992).

BSC aims to measure a company's activities regarding four main perspectives: Financial, Customer, Internal Business Processes and Learning & Growth. The first axis not only measures the financial performance but also ensures the efficient use of financial resources. The second axis addresses

the impact of the business on the market in relation with the customers' satisfaction. The third axis focuses on the internal processes' performance from the angle of efficiency and quality. The last axis is concerned with the global view of the company, business's culture and vision, human resources and information systems.

The next paragraphs first present the causal links in the BSC which allows us to create our own strategic map. Second, we present how the BSC has evolved from a strategic alignment tool to a strategy construction tool. Third, we discuss how the BSC promotes cultural change, and finally, we present some applications of the BSC in healthcare. We also provide a figure (Figure 2) that shows the evolution of BSC development over time.

3.1 Causal relationships in BSC and strategy map

An evolution of the BSC is the emergence of the concept of causality and strategic objectives in the implementation of the BSC in an underwater engineering and construction firm (Kaplan and Norton 1993). In this study, the founders of the BSC show how to create cause and effect relationships between the perspectives of the BSC model and how BSC helps to select relevant and consistent indicators. According to Kaplan and Norton (1998), strategy is a set of hypotheses about causes and their effect on the organization. In practical terms, a well-developed BSC should contain a set of strategic objectives linked to each axis of the model. These objectives are linked to each other by so-called causal relationships. The links between the strategic objectives involve, at a second level, links between the performance indicators.

The article of Malina, Nørreklit, and Selto (2007) discusses the necessity of cause-effect links for efficient management control. This article describes the evolution of the investigation of cause-effect relationships for a PMS. First, the cause-effect relationship predicts the reliability, suitability and capacity of expected outcomes. For example, the analysis of non-financial measures can be used to predict future financial performance. This prediction leads to better decision-making. Second, the cause-effect relationship enables more fluid communication for the achievement of strategic plans and higher performance. For Kaplan and Norton, the key to achieve better performance is easy strategy communication based on cause-effect relationships. Scorecard's causal links that show the best results and trade-offs, support the learning, creation and communication of knowledge. Third, causal links help to deploy the strategy. Since PMSs based on cause-and-effect relationships allow to visualize the impact of actions on future performance, it aligns managers' objectives with those of their employees, which allows the alignment of the operational level with the strategic level.

Kaplan and Norton not only identified the presence of "cause and effect relationships" as a fundamental characteristic of any BSC, but they also highlighted the need to correctly measure the impact of these relationships. In other words, they mentioned the need for a complementary analytical approach that can define, quantify and rigorously test cause and effect relationships throughout the organization (Barnabè, Hoque, and Busco 2012). It is not obvious to analytically verify the validity of all cause-effect relationships in a scorecard. In this context, Lorino (2001) discussed the subjectivity of establishing causal links. According to him, the subjective

character comes from actors' interpretations and grows with the complexity of the situation. As Kaplan and Norton (Kaplan and Norton (1996) suggest, we can try to reduce the degree of uncertainty and subjectivity by attempting to scientifically justify the cause-effect relationships between performances. Lorino (2001) stresses that "this type of attempt can only have a limited scope because technical, commercial and human conditions are changing and correlations observed in the past cannot be extrapolated to the future."

One of the most fundamental critiques of the BSC concerns causal relationships, their relevance, the way managers use them and the ability to use mathematical and computer models to test and simulate their impact on strategy implementation. Ittner and Larcker (2003) discuss the various errors and limits that many companies are unaware of during the implementation of the BSC and that directly and indirectly impact their strategies. The article highlights two aspects related to causal links: 1) the lack of links between the measures considered and 2) the strategy and the non-conformity of the links. First, some organizations do not benefit from the BSC because they make decisions based on measures that do not reflect the evolution of their strategy or the linkages between financial and non-financial measures. This leads to misdirected investments and missing targets. Indeed, many organizations rely on a causal model to evaluate their strategies that do not link correctly financial and non-financial measures. Second, the majority of companies never check the validity of the links between strategic objectives and performance measures. For example, what is the link between employee satisfaction and customer satisfaction? Managers often reply that the links were self-evident (without deep verification to validate the links). Ittner and Larcker consider that such assumptions are "half-baked" or wrong. It often ends-up by measuring too many things that do not meet the strategic objectives.

A continuation of BSC development was presented in 2001, by Kaplan and Norton, in their book "The strategy-focused organisation" (Kaplan et al. 2001). This book outlines how to build a visual map, called strategy map, that shows the goals and causal links necessary to execute a strategy. A strategy map is a visual representation designed to communicate and validate the "scenario" of a company's strategy to enable its stakeholders to understand, control and contribute to the achievement of that strategy. Scholey (2005) gives a well structured 6-step process to adequately create a strategy map. The first step is to determine the overriding objective. An overriding objective presents the final goal that the company aims to achieve. The second step is to select the appropriate value proposition, in this phase the managers choose the strategic vision with which the company intends to achieve the overarching objective. The third step is to determine the general financial strategies to follow. The fourth step is to determine customer-focused strategies. The fifth step is to decide how internal processes will support execution of the chosen strategies. The last step is to implement the skills/capabilities and employee programs that are required to achieve strategy.

3.2 BSC as a strategic alignment tool and a constructivist approach

The BSC from 1992 up to date has undergone major changes from its normative form to improved versions. The BSC has moved from a strategic alignment tool to a strategy

construction tool. When Kaplan and Norton created BSC, they wanted to provide an effective tool to help companies define their strategy and monitor its execution. The execution of this strategy requires (necessarily) the involvement of all actors in the company. For this purpose, each company must adopt a deployment approach for its multidimensional strategy model. The normative version of the BSC proposes a “top-down” deployment approach to operationalize the strategy and align departmental and personal goals with the overall strategy. In this concept, the strategy and objectives are defined exclusively by the management team and are then transmitted to all members of the company to align their actions with the decisions of the managers. However, Kaplan and Norton were aware that new strategy ideas can come from lower levels of the company. There is a need for managers at all levels to review and judge the relevance of the strategy (Norreklit 2000). The BSC can also be interpreted and understood differently by the company’s employees.

Some authors consider that the BSC can contribute to an emerging strategy in the organization. Emerging strategy is the concept that new strategies appear over time when objectives are confronted and allows to adjust the overall strategy with a changing reality. According to Mintzberg and Waters (1985) the strategy achieved at last is different from the strategy desired at the beginning because there are some plans we fail to carry out (and there are new strategies that arise over time. In this case, it would be more favourable to opt for a “bottom-up” approach, where the strategic objectives would result from collaborative and participatory management. The “bottom-up” approach allows the participation of all the actors, which promotes the emergence of new objectives and the construction of the strategy over time. Simons (1994) developed the idea that for a better balance of the organization, a co-construction of the strategy and its implementation is highly recommended.

As an example of the constructivist approach, we can cite the evolution of a PMS in a British aerospace company (BAE systems). Jazayeri and Scapens (2008) describes the “Business Values Scorecard” (BSV), which is the PMS used to measure the non-financial performance of the company and to support a cultural change project focused on five dimensions: performance, people, customers, partnerships and innovation and technology. Each of these dimensions is measured on the basis of a limited number of key indicators. Through interviews and document review, the authors described how the PMS had evolved, rather than being imposed, and how the PMS was used to enable the emergence of the strategy for

change. This performance measurement tool (as it was), was not designed as a complete PMS that satisfies all the needs of the company, but rather has evolved and changed over a long period of time to adapt to the situation and the objectives of BAE system. For the deployment of this PMS, an ERP (Enterprise Resources Planning) system is connected to the BSV for the day-to-day sharing of a dashboard in a visually attractive way. This information is shared with all employees and managers, which allows for strategic alignment and fosters fluid communication between all stakeholders that leads to strategy construction.

3.3 BSC a tool promoting cultural change

As the BSC matures over time, the literature on the relationship between organizational culture and the BSC grows (Deem 2009). Kaplan and Norton (2004) note that successful BSC implementation leads companies to have an organisational culture in which “people were deeply aware of and internalized the mission, vision, and core values needed to execute the company’s strategy”. Also, The Learning and Growth dimension in the BSC is used for “shaping the culture” of the organization. Henri (2006) proposes a model to test the relationships between organizational culture and the components of the PMS. For him, culture is the most important factor affecting every organizational aspect. A well understanding of this conditional factor is required in order to examine and understand the PMS to be implemented. Henri explained the relationship between organizational culture and PMSs developed by organizations using a modelling tool called Structural Equation Modelling (SEM). This tool enables the simultaneous study of several causal relationships. This model considers four dimensions of culture: institutional character, institutional leader, institutional cohesion and institutional strengthening. In short, this tool allows managers to judge whether a PMS meets their strategic needs.

3.4 BSC uses in the public and healthcare sectors

Healthcare institutions have a reputation as rigorous and hard to administrate (Chang, Chiao, and Tsai 2017). A BSC may be beneficial to hospital administration by contributing in performance assessment, policy implementation, monitoring, accountability and the development of strategies (Gao et al. 2018). Canada leads in developing the BSC in healthcare. According to Porporato, Tsisis, and Vinuesa (2017), the use of scorecards in hospitals dates back to the 90's. In Toronto, Women's College Hospital began using a scorecard organized around key objectives. Since then, Ontario hospitals have gained a great experience in designing and using the BSC. The authors mention other early experiences of BSC use in Canada

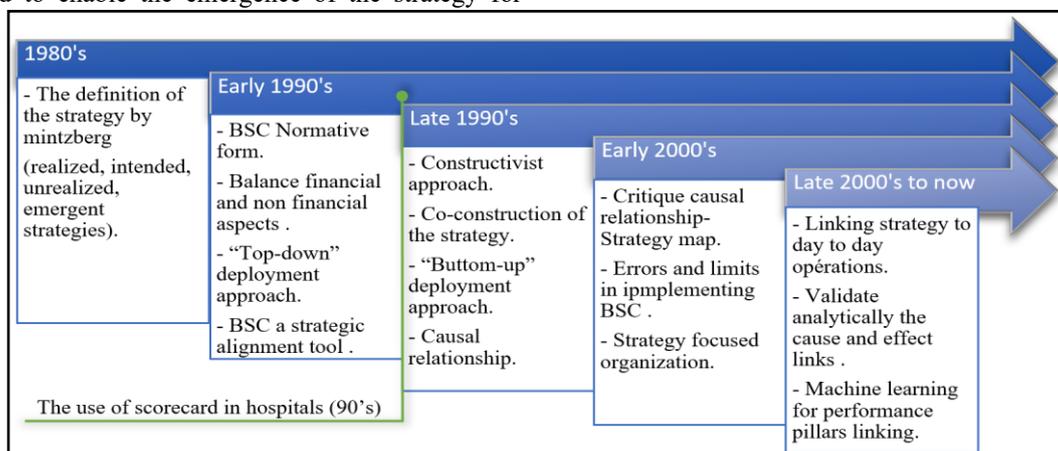


Figure 2. BSC evolution over

such as in the province of Alberta.

The results of the analytical research of Porporato, Tsasis, and Vinuesa (2017) indicate that it is very difficult to demonstrate a causal relationship between outcome (lag) measures and prospective (lead) measures in healthcare organizations. In addition, some statistical techniques can identify correlations between variables in the best case, but they cannot determine precisely what the cause and effect are. This study shows how difficult it is to establish cause-effect relationships in the healthcare sector. But on the other hand, a clear correlation appears between process and outcome measures. In another Canadian experience, L. Chan (Lilian Chan 2009) illustrated in his article the evolution of health systems in implementing the BSC and strategy map in the Ontario healthcare organisations. The paper provides decision-makers with choices on how to implement the strategic management tool in their establishments. The priority of a non-profit organisation is different to the other which seeks to have a better financial balance sheet. In this case, the financial dimension takes the last rank compared to other perspectives of the model.

In the USA, Kaplan (2020) used BSC to promote better emerging and collaboration between major healthcare providers. Many public healthcare organizations require the practices of independent physicians and the services of independent organizations. BSC helps this merging organizations to express the new strategic plan into a balanced set of performance measures, to better communicate the strategy between collaborators and to align the decision with day to day actions. Kaplan presents the merge experience between St. Mary's Duluth Clinic with top-down culture and Duluth Clinic with decentralized bottom-up culture and how they co-create a new strategy, strategy map and a common BSC. Also, the show how BSC helps aligning doctors' vision with the new strategy and culture of the organization. According to Kaplan, managers of merging entities enhance the chances of success by implementing a BSC to establish a convergence point to discuss and debate priorities within the strategy.

In Europe, Montalan and Vincent (2011) propose a methodology to develop a BSC for strategic control in a hospital in France. . The authors have chosen not to follow the original normative approach of the BSC, but to adopt an approach allowing a co-construction of the strategy by referring to the works of (Mooraj, Oyon, and Hostettler 1999), (Lorino 2001), (Choffel and Meyssonier 2005). By planning meetings with the hospital teams, the authors developed a scorecard made up of four axes: customer, internal process, organizational learning and economic. The economic axis (financial axis of the BSC) is moved to the fourth position as it presents a constraint in healthcare organisations. Then, the authors defined four primary strategic objectives (one for each axis). A strategy map was developed and validated by a multidisciplinary working group, from which they proposed a number of indicators for each strategic objective.

4 METHODOLOGY

This section is divided into four parts. First, we present our data collection process. Second, we present the data analysis that we performed by using the SWOT method. Third, based on these results, we propose an approach combining the QPM model and the BSC principles to support the efficient implementation of the QPM model. Finally, we show how the

BSC can contribute to strengthen the QPM an performance management in the CCSMTL.

4.1 Data gathering

We collected data in order to develop a well understanding of the CCSMTL's internal processes and its current situation regarding the QPM implementation and use (models used, indicators selected, etc.). Two methods were used to gather data, observation/interviews and review of documents provided by the CCSMTL.

4.1.1 Observations and interviews

In the beginning of the project, in order to understand the issues encountered by the CCSMTL, we participated in the QPM Coordination Committee and the QPM Advisory Committee meetings. The members of the Coordination Committee are a restricted number of senior managers from the ethics, quality and organisational performance department. Their mission is to carry out the QPM implementation throughout the CCSMTL in compliance with the strategic plan of the MSSS. On a regular basis, the Coordination Committee invites the managers of all CCSMTL's departments, who represents the QPM Advisory Committee, to inform them on decisions to be made and ask their feedback. additionally, we visited two hospitals, two rehabilitation centres and several control rooms. We participated in a trainee on performance development led by two CCSMTL managers as well. The main purpose of the observations was to collect data and identify practices and behaviours that might not be gathered with interviews alone or other techniques.

To collect more information, we interviewed CCSMTL experts. In total, we conducted 10 interviews with 12 CCSMTL's stakeholders and senior managers, from May 2019 to February 2020. We proceeded through the following five steps in each meeting. First, we specify the objective of the meeting. The main objective is to check the QPM's implementation progress within the departments in the CCSMTL. Second, we choose the meeting method. DiCicco-Bloom and Crabtree (2006) mention three types of meetings: structured, semi-structured and in-depth individual interviews also called unstructured interviews. We have chosen the structured interview method. Under this method, the interviewer asks a number of structured questions and the interviewee reply in their own way. Third, we recruit the interviewees. We select the participants based on our progress in the project. To identify relevant participants, we collaborated with a manager engaged in the QPM implementation. Fourth, we conduct the meeting, and if necessary, we plan the next meeting. Fifth, we analyse information and formulate the results. Most of the interviews were one-to-one meetings.

4.1.2 Document review

MSSS's Laws and guidelines, Accreditation Canada, and PLANETREE represent the main references on which the CCSMTL relied for the creation of its QPM model. First of all, the CCSMTL follows the recommendations and guidelines of the MSSS's Law on Health and Social Services (MSSS 2012). Referring to the ministerial reference framework for performance evaluation (MSSS 2012), the Law provides a model for managing the human, material and financial resources that promotes the contribution of all actors and encourages responsibility distribution among all participants. Furthermore, the MSSS announced that an accreditation in

health and social services must be adopted by all healthcare organizations. In 2017, the MSSS chose, for all organizations in the healthcare network, the services of “Accreditation Canada” which evaluates institutions against Health Standards Organization (HSO) norms. Moreover, Accreditation Canada suggests a set of performance management practices (Agrément-Canada 2014). The francophone PLANETREE (Cosgrove 1994) of which the CCSMTL is a member, is a group of health and social services organisation looking at developing healthcare and social services. PLANETREE offers a model for implementing management and healthcare practices focused on the customer.

4.2 Data analysis

Through documents reviewed, meetings, field visits, participation in committees and events, we gathered a large amount of information. We used the SWOT method in order to structure the data and to identify the strengths and weaknesses related to the current implementation of the QPM model. SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis is a tool from which strategies can be developed and which supports improving organisational performance (Pickton and Wright 1998). It helps to clarify the strengths to be maintained, the weaknesses to be addressed, the opportunities to benefit from, and the constraints to be handled. Our results and finding were validated by experts from the CCSMTL.

Table 1. SWOT analysis for the CCSMTL

	Strengths	Weaknesses
Internal	<ul style="list-style-type: none"> - QPM implemented in strategic level - good communication (regular meetings) - enough indicators in Accessibility/Quality - Several practices ensuring service quality are implemented. - Department stakeholders can suggest other indicators 	<ul style="list-style-type: none"> - imbalance between the QPM dimensions (in indicator perspective) - lack of customer indicators - The need to embed a customer-focused organizational culture. - lack of linkage between indicators and organisational objectives - Not possible to ensure the strategic tracking appropriately with the current PMS. - The need to clarify the organizational vision. - QPM not well deployed in the tactical and operational level. - Lack of strategic alignment with operations.
	Opportunities	Threats
External	<ul style="list-style-type: none"> - Planetree, Accreditation Canada, MSSS models. 	<ul style="list-style-type: none"> - Ministerial constraints - dependent and conflictual indicators

We observe that the majority of practices and indicators address administrative objectives that satisfy ministerial constraints (Threats). In fact, the CCSMTL needs to translate the ministerial strategy into performance measures appropriate to its circumstances. Added to that, the ministerial strategic plan is generic for all Quebec's healthcare systems, therefore the CCSMTL needs to plan and set its own strategic objectives that derive from the ministry's overall strategy. We noticed an

imbalance between the QPM dimensions. More precisely, there is a lack of indicators that measure the optimization and mobilization dimensions (Weaknesses) while, there are enough indicators in accessibility / quality dimension (Strength). Also, we noticed a lack of indicators within the customer dimension (Weaknesses), which is the most important component in the QPM. Yet, there are several practices to ensure service quality and check customer satisfaction. The CCSMTL managers stress the needs to integrate a customer-focused organizational culture and clarify the organisational vision.

Some of the CCSMTL indicators are dependent with each other while others may be conflictual (threats). We consider this as a threat because, in some cases, trade-offs are inevitable, and this might be difficult to handle by managers. For example, it is necessary both to reduce waiting time and to improve direct service time (time with customers), an improvement in one indicator reduces the other and inversely, it is therefore necessary to make trade-offs between the two aspects. We also observed a lack of links between some indicators and the organisational objectives (weakness). It is not possible for senior managers, with the current PMS, to ensure strategic tracking and check the overall performance without analysing a large number of indicators (weakness). In addition, they are not able to examine the performance of a specific department in an easy way (weakness). Each department/service measure its performance indicators independently from the higher hierarchical level and other departments/services (weakness). Despite this, department's managers are free to suggest other indicators than those proposed by senior managers and MSSS (strength). An interesting finding is that the QPM is well implemented at the strategic level (Strengths). However, it is not used enough at the tactical and operational levels (Weaknesses). In fact, there is a big interest of senior managers in aligning the employees' practices with the QPM objectives and there are regular meetings in every department (Strengths), but the dashboard results are not well discussed (weakness).

4.3 Adapting the BSC within the CCSMTL framework

It is clear that there exist resemblances between the QPM and the BSC model. First, there is a strong similarity between the QPM dimensions and the BSC axes, in a way that they present similar performance pillars (customer, internal processes, resources management, etc.). Second, both models allow the derivation of performance indicators to create a performance measurement system. Third, QPM and BSC provides a platform for stakeholders to share and discuss strategic objectives. In addition, the BSC is often recommended in healthcare because it gives a more complete picture of healthcare activities (Inamdar, Kaplan, and Reynolds 2002), (Nobre 2001). Inspired by Canadian experiences regarding BSC implementation in healthcare organizations (Porporato, Tsisis, and Vinuesa 2017) and in an effort to address the weaknesses and threats identified in the SWOT analysis, we propose the following approach to integrate BSC principles in the CCSMTL performance management model. More precisely, our objectives are to:

- 1) Plan and set strategic objectives.
- 2) Translate the strategy into performance measures.
- 3) Balance the QPM dimensions.
- 4) Support cultural change within the organization
- 5) Clarify the organization's vision.
- 6) Ensure strategic tracking and feedback.
- 7) Lead to strategic alignment

- **Plan and set strategic objectives**

Despite the fact that the CCSMTL is constrained by the MSSS's strategic plan, which sets out the strategic objectives that the health and social services network is committed to achieve over the years (2019-2023), it also has a degree of freedom to choose its strategic plan to meet the ministry's objectives. Our aim is to support the CCSMTL in defining this strategic plan.

- **Translate the strategy into performance measures**

Many organizations build their strategies based on erroneous measures. They may track several indicators without being significant for their strategy and sometimes end up with conflicting results. To address this issue, we rely on selecting proper key indicators. The literature on BSC provides a good starting point for identifying performance indicators that are consistent with the CCSMTL's strategy and enables the achievement of the organizational objectives.

- **Balance the QPM dimensions**

For improving the QPM model, it is essential to dedicate sufficient interest to each pillar of the performance and not neglect any dimension of the model. In our case, it can be seen that the CCSMTL gives more attention to the Accessibility / Quality dimension compared to the other aspects, which creates an unbalanced model and a single lever for performance driving.

- **Support cultural change within the organization**

One main objective of the CCSMTL is to share the organizational values with all members and stakeholders of the organization. Values are a set of collectively shared beliefs that describe the employees' behaviours that are essential to achieving the organization's vision. Furthermore, the CCSMTL adopts a customer-focused approach. We intend to create a scorecard that leads to increased customer satisfaction and that shapes employees' practices and drives cultural change.

- **Clarify the organization's vision**

The vision is not the strategy, but a guideline for making the right strategic decisions. It presents a clear view of the future and defines where the CCSMTL wants to go and the general goals it wants to achieve. As mentioned in section 1, the QPM model presents a translation of the CCSMTL's vision and values into strategic objectives. A successful implementation of this model leads to a successful deployment of the organizational vision.

- **Ensure strategic tracking and feedback**

Defining performance indicators is not enough. Setting realistic, measurable and achievable objectives is also essential to evaluate the gap between the outcomes and the targets and monitor the evolution of the strategy. One of our objectives is to provide the required tools to enable the tracking of performance indicators throughout the CCSMTL departments.

- **Lead to strategic alignment**

This consists in aligning day-to-day operations with the organization's strategy. The QPM model, as a strategic tool, must also be implemented in the tactical and operational levels. To do so, we rely on the BSC's foundation to enhance communication between all stakeholders and employees that will lead to a better understanding of the strategy.

In the following, we propose to compare the BSC axes with the dimensions of the QPM. To this end, we relied on the literature, mainly on the model proposed by Kaplan (2006), that he adapted to the public sector (figure 3).

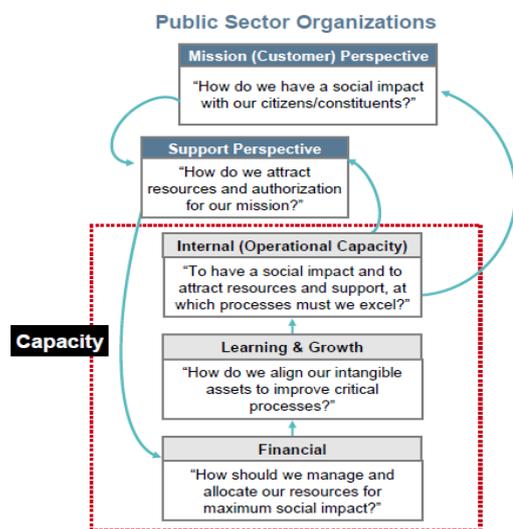


Figure 3. BSC for public sector, adapted of Kaplan (2006)

4.3.1 QPM-BSC Model axis definition

This sub-section presents the analogy between QPM and BSC axes. Based on this analysis, we identify the axes of our QPM-BSC model.

- **Customer**

This axis corresponds to both the "customer" dimension of the QPM and the "customer" or "mission" perspective of the Kaplan and Norton model. In healthcare, the customer represents the person receiving the service or care, their family, the community and the population mainly in the south-central territory of Montréal, but also in all Montreal region and the province of Quebec.

- **Accessibility / Quality**

This axis corresponds to the "internal process" or "operational capacity" of Kaplan and Norton model. As mentioned in Section 1, accessibility / Quality is the ability of the organization to safely meet the needs and expectations of customers according to best practices through offering more accessible and continuous service.

- **Mobilisation**

There are three main components of learning and growth axis: human capital, information capital, and organizational capital (Scholey 2005). The "Human capital" includes the management of all the skills and behaviour of human resources, which is part of the "mobilization" dimension of the QPM. The "mobilization" means to use the skills and talents of everyone in the organization, partners, customers and their families, inspiring personal growth and commitment to accomplish the mission of CCSMTL. The "information capital" is the way in which organizations use their information systems, networks and databases to carry out their strategy. This aspect of the BSC seems to be part of the "Optimization" dimension of QPM. The "organisational capacity" is defined as the organization's ability to align employees' goals with the strategy. This requires committed and passionate employees and efficient, agile and innovative management. Therefore, this coincides with both the "mobilization" and the "optimization" dimensions, which is the search for continuous improvement and innovation for an efficient use of the resources to provide, over time, services that are adapted to the needs of customers. In short, learn and growth axis is included in both the "mobilization" and the "optimization" dimensions.

- **Optimization**

This axis corresponds the financial axis and a part of learning and growth axis of the BSC as mentioned above. According to

Aidemark (2001), in the healthcare sector, the financial aspect is a constraint and not a goal. These aspects cannot be guided by the customer (patient) as proposed in the BSC normative model. Which coincides with what mention Kaplan and Norton (1999): “The financial perspective provides a clear long-run objective for profit seeking corporations; however, it serves as a constraint, not an objective, for non-profit organisations.” Therefore, the “financial” or “economic” axis as indicated by Montalan and Vincent (2011) during the last position in the model and corresponds to a part of the “optimization” dimension of the QPM.

4.3.2 Strategy map development

Based on the guidelines suggested by Scholey (2005), we adapted his six-step based method to create our strategy map. The first step is determining the overriding objective which is the final result that is desired after implementing the strategy. For a profit-making organization, this goal should include a financial target and a time dimension. In healthcare sector, the primary objective is generally disinterested of the financial aspect. It can be related to customer experience or accessibility outcomes and it should be a certain result in a certain timeframe. For example, increase the rate of access deadline compliance to 95% in one year. The second step is to select the appropriate value proposition. According to Treacy and Wiersema (2007), there are three value positions for an organisation: operational excellence (operational effectiveness in healthcare), product leadership (service quality in healthcare) and customer intimacy (customer experience in healthcare). The CCSMTL follows a customer-centred approach. Therefore, all the strategy should focus on a better customer experience. The third step is to determine strategies related to customer dimension. We need to clearly define the organization’s strategies for dealing with clients and identify the objectives related to the customer dimension. The fourth step is to determine the objectives related to the internal processes. In this step, we move from what we want to achieve to how we intend to achieve it. The fifth step, we focus on learning and growth. Three main elements must be considered in this step: human capital (e.g., knowledge, collaboration, employee motivation, etc.), information capital (i.e. how the organization use its information systems, measures, networks, databases, etc.), and organizational capital, i.e., the alignment with the organisation’s values and vision. In short, these three aspects must match the value the chosen proposition. The last step is to determine the global financial strategies to track. In this phase, instead of improving revenues, we focus on reducing costs because the financial excellence is not the most important goal to reach in public healthcare organizations.

4.3.3 Cause-effect links

It is difficult to establish cause-effect relationships between the dimensions of the QPM model. First, we need a mature model to establish links between its dimensions, however, the QPM is not fully deployed throughout the organization. Second, some dimensions do not match strategic objectives and indicators, which prevents the establishment of causal links. Third, the historical data regarding some indicators are poor, so we could not judge which indicator make an effect on others. Finally, from our literature review, it is observed that the majority of organizations do not prove the validity of their causal model (Ittner and Larcker 2003) and also it is very difficult to establish cause-effect relationships between indicators in the healthcare sector (Porporato, Tsisis, and Vinuesa 2017). For these reasons, it is extremely difficult to create and validate

analytically the cause and effect links between all the pillars in our strategy map. However, we could rely on the expertise of managers to validate causal links between strategic objectives. After a period of a few years and after having collected the necessary information on all the indicators, it will be possible to verify analytically (at least partially) causal links validity and to make strategic adjustments following a constructivist approach.

Figure 4 shows the axes of the proposed QPM-BSC for the CCSMTL. This model also shows the overall causal links between the axes: the optimization axis presents the technical support for accessibility / quality axes and also allows to manage the financial resources necessary for the functioning of the organization. Also, there is a mutual relationship between the optimisation axis and the mobilisation axis. The mobilization and commitment of employees have an impact on the internal processes of the CCSMTL which subsequently has an effect on customer satisfaction which refers to the organizational mission. The requirements and needs of the customers put a stress on the Ministry (MSSS). In its turn, the MSSS requires the CCSMTL’s “capacity” to follow its strategic plan and it gives the necessary financial funds for that.

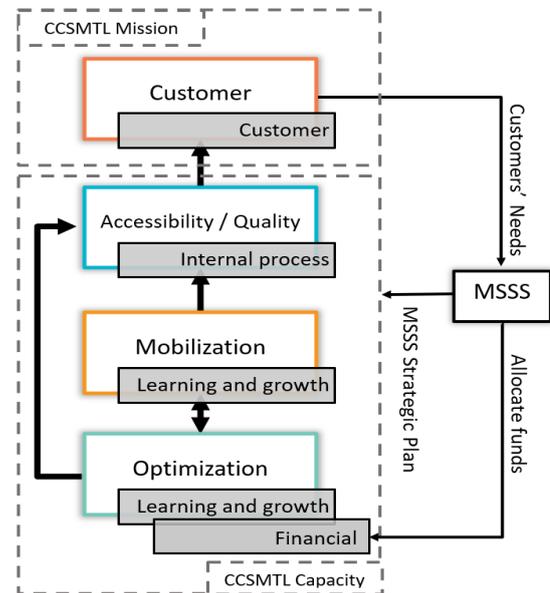


Figure 4. QPM-BSC Model

4.3.4 QPM-BSC operationalization

To operationalize the proposed QPM-BSC model, developed a five phase-based approach to set up a PMS in the CCSMTL. This approach is discussed in details in (Ben fradj, Boukherroub, and Olivier 2020). The first phase aims to shape the overall structure of the indicator system. To this end, we considered the CCSMTL organisation chart. In the second phase, we identify with the decision-makers, performance indicators that meet eight criteria. These criteria are the outcome of an in-depth review of the literature on performance measurement and performance indicators. The managers in each department use the strategy map to draw the organizational objectives associated with that department. Then, they choose performance indicators based on the eight criteria mentioned earlier. In this way, managers become aware of choosing performance indicators that are linked to all aspects of performance, especially for the customer dimension to implement a customer-focused culture. The third phase allows to put on the same scale indicators that have different units, that are calculated differently or that have different

targets (normalisation). To this end, we used “Min-Max Normalization” equation (Goyal et al. 2014). In the fourth phase, we use the AHP (Analytic Hierarchy Process) method for weighting the indicators. The last phase aims at calculating the overall indicators (performance indices) for each dimension of the QPM-BSC model. In this phase indicators are aggregated from operational level to tactical level, and then tactical to strategic level, in a “bottom-up” approach, which promotes the emergence of new departmental objectives and the construction of the strategy over time. The “bottom-up” culture encourages collaboration with all the CCSMTL actors to contribute to building, over time, the strategy of the organization in a constructivist way.

Another aspect that will support the implementation of the QPM-BSC model is the deployment of the control room network. The control rooms are hierarchically structured and allow decision-makers to delegate tasks to a lower decision-making level or to request support from a higher level. Moreover, the periodic meetings between the advisory and coordination committees ensure the coherence and the “bottom-up” deployment of the QPM-BSC. On the other hand, governmental organizations such as the CCSMTL have a centralized hierarchical structure where ministerial constraints have to be respected. Therefore, in some cases, some strategic objectives are imposed by policy-makers, who imposes a “top-down” approach. For this reason, the QPM-BSC is embedded in a mixed top-down and bottom-up culture that allows both the strategic alignment with the ministerial strategic plan and collaboration in the rest of the organizational objectives.

4.4 How the proposed QPM-BSC can contribute to improve performance management in the CCSMTL?

The QPM-BSC model is a means to address the weaknesses and threats identified in the SWOT analysis (Section 4.2). First, the way in which we have integrated the BSC's aspects with the QPM model allows managers to reconstruct their strategic plan in a way that respects the ministerial constraints and meets their specific organizational needs. By using a constructivist approach, the CCSMTL will be able to build the strategy over time by allowing the emergence of new perspectives. This principle promotes the integration of all stakeholders' views. Second, a rich literature on BSC provides managers with the necessary information to translate the organizational strategy into relevant performance measures that meet the four dimensions of the QPM. Third, the BSC, as its name indicates, helps to ensure a balance between the pillars of performance. It alerts managers that they should not forget to measure the performance related to one or more dimensions of the QPM. This will lead them to address the lack of indicators in the mobilization, optimization and customer pillars and adequately balance the four QPM dimensions. Fourth, the BSC responds to the CCSMTL's need to build a customer-focused culture: by creating the strategy map, all resources and performance pillars will be directed towards the goals of customer dimension, which is in the top of the strategy map. The operationalization of the QPM-BSC model ensures the deployment of this culture in all CCSMTL institutions. This enables also the implementation of the organizational vision, which leads employee to adopt practices aligned with the organization's values. Fifth, the proposed PMS generates performance indexes that measure each dimension of the QPM-BSC model for each department of the CCSMTL. This allows managers to efficiently monitor the performance. Finally, the proposed PMS is a tool that

aggregates performance indicators from the operational to the strategic level in a “bottom-up” approach. At the same time, it helps aligning the strategy with employee practices in a “top-down” approach. In fact, by supporting senior managers in the indicator identification process, this tool helps in setting up operational indicators that meet strategic needs. This encourages employees to align their practices in order to improve these indicators.

5 CONCLUSION

Performance in the CCSMTL is driven by its Quality Performance Model (QPM) that encompasses four dimensions (Customer, Accessibility / Quality, Mobilization, Optimization). The objective of this study is to support the implementation of this performance model by using the BSC principles. First, we performed a SWOT analysis based on large data gathered in the CCSMTL. This led us to identify a number of aspects of the QPM that need to be improved. To this end, we proposed an approach that combines BSC principles and the QPM model. The approach supports senior managers in planning and setting strategic objectives starting from the QPM dimensions and MSSS's requirements. Second, it facilitates translating the chosen strategies and objectives into performance measures. Third, it helps managers balancing the QPM dimensions and it enables them to identify performance indicators in all performance pillars. Fourth, the approach supports cultural change within the organization (customer-centred culture). Fifth, it helps in clarifying the organization's vision. Sixth, it ensures strategic tracking and feedback, and, finally, day-to-day operations alignment with the organizational strategy. This study, which is rather qualitative, is complemented by a quantitative study in which we propose a five-phase based approach to operationalize the QPM-BSC by providing and implementing PMS (Ben fradj, Boukherroub, and Olivier 2020). This is a long-term project that requires the participation of all members of the CCSMTL and that will bear fruits only years after the implementation of our proposed approach. We intend to carry on this project in the future.

6 ACKNOWLEDGEMENT

We acknowledge the MITACS Globallink program, the CCSMTL and École de technologie supérieure (ÉTS) for financing this research. We are grateful to Mr. Jacques Drôlet and all our collaborators from the CCSMTL.

7 REFERENCES

- Agrément-Canada. 2014. Guide sur l'évaluation de l'amélioration de la qualité et les indicateurs de qualité.
- Aidemark, Lars-Göran. 2001. "The meaning of balanced scorecards in the health care organisation." *Financial accountability & management* 17 (1):23-40.
- Barnabè, Federico, Zahirul Hoque, and Cristiano Busco. 2012. "The causal relationships between performance drivers and outcomes." *Journal of Accounting & Organizational Change* 8 (4):528-538. doi: 10.1108/18325911211273518.
- Ben fradj, Anes, Tasseda Boukherroub, and Claude Olivier. 2020. "A framework for developing a performance measurement system in healthcare: a case study in montreal." accepted in the 13th International Conference on Modeling, Optimisation and

- Simulation - MOSIM'20, Agadir – Morocco, November 12-14, 2020.
- CCSMTL. 2017. Le Rapport annuel de gestion 2016-2017 « dévoués à LA QUALITÉ ET à LA PERFORMANCE ». Centre intégré universitaire de santé et de services sociaux (CIUSSS) du Centre-Sud-de-l'Île-de-Montréal.
- Chang, Chuan-Hui, Yu-Ching Chiao, and Yafang Tsai. 2017. "Identifying competitive strategies to improve the performance of hospitals in a competitive environment." *BMC health services research* 17 (1):756.
- Choffel, Denis, and François Meyssonier. 2005. "Dix ans de débats autour du Balanced Scorecard." *Comptabilité-Contrôle-Audit* 11 (2):61-81.
- Cosgrove, Tracey L. 1994. "Planetree health information services: public access to the health information people want." *Bulletin of the Medical Library Association* 82 (1):57.
- Deem, Jackie W. 2009. *The relationship of organizational culture to balanced scorecard effectiveness*: Nova Southeastern University.
- DiCicco-Bloom, Barbara, and Benjamin F Crabtree. 2006. "The qualitative research interview." *Medical education* 40 (4):314-321.
- Gao, Hongda, He Chen, Jun Feng, Xianjing Qin, Xuan Wang, Shenglin Liang, Jinmin Zhao, and Qiming Feng. 2018. "Balanced scorecard-based performance evaluation of Chinese county hospitals in underdeveloped areas." *Journal of International Medical Research* 46 (5):1947-1962.
- Goyal, Himani, Raghavendra Pokuri, Sandeep Kathula, and Naveen Battula. 2014. "Normalization of data in data mining."
- Henri, Jean-François. 2006. "Organizational culture and performance measurement systems." *Accounting, Organizations and Society* 31 (1):77-103. doi: 10.1016/j.aos.2004.10.003.
- Inamdar, Noorein, Robert S Kaplan, and Kimberly Reynolds. 2002. "Applying the balanced scorecard in healthcare provider organizations/Practitioner's Response." *Journal of healthcare management* 47 (3):179.
- Ittner, Christopher D, and David F Larcker. 2003. "Coming up short on nonfinancial performance measurement." *Harvard business review* 81 (11):88-95.
- Jazayeri, Mostafa, and Robert W. Scapens. 2008. "The Business Values Scorecard within BAE Systems: The evolution of a performance measurement system." *The British Accounting Review* 40 (1):48-70. doi: 10.1016/j.bar.2007.10.007.
- Kaplan, Robert S. 2006. "Using the balanced scorecard in the public sector." *Government Summit*.
- Kaplan, Robert S. 2020. "Using the Balanced Scorecard for Successful Health Care M&A Integration."
- Kaplan, Robert S, Thomas H Davenport, Norton P David Kaplan S Robert, Robert Steven Kaplan, and David P Norton. 2001. *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*: Harvard Business Press.
- Kaplan, Robert S, and David P Norton. 1992. "The Balanced Scorecard œ Measures That Drive Performance—, in: Harvard Business Review, January-February 1992." *Putting the balanced scorecard to work*.
- Kaplan, Robert S, and David P Norton. 1996. Using the balanced scorecard as a strategic management system.
- Kaplan, Robert S, and David P Norton. 1998. "Putting the balanced scorecard to work." *The economic impact of knowledge* 27 (4):315-324.
- Kaplan, Robert S, and David P Norton. 1999. "The balanced scorecard for public-sector organizations." *Balanced Scorecard Report* 15 (11):1999.
- Kaplan, Robert S, and David P Norton. 2004. "Measuring the strategic readiness of intangible assets." *Harvard business review* 82 (2):52-63.
- Kaplan, RS, and DP Norton. 1993. "Putting The Balanced Scorecard to Work." *Harvard Business Review*.
- Lagacé, Denis, and Sylvain Landry. 2016. "Salles de pilotage: un nouveau mode de gestion de la performance." *Gestion* 41 (3):90-93.
- Lilian Chan, Yee-Ching. 2009. "How strategy map works for Ontario's health system." *International Journal of Public Sector Management* 22 (4):349-363. doi: 10.1108/09513550910961628.
- Lorino, Philippe. 2001. "Le balanced Scorecard Revisite: dynamique stratégique et pilotage de performance exemple d'une entreprise énergétique."
- Malina, Mary A, Hanne SO Nørreklit, and Frank H Selto. 2007. "Relations among measures, climate of control, and performance measurement models." *Contemporary Accounting Research* 24 (3):935-982.
- Mintzberg, Henry, and James A Waters. 1985. "Of strategies, deliberate and emergent." *Strategic management journal* 6 (3):257-272.
- Montalan, Marie-Annick, and Béatrice Vincent. 2011. "Élaboration d'un balanced scorecard en milieu hospitalier." *Revue française de gestion* (2):93-102.
- Mooraj, Stella, Daniel Oyon, and Didier Hostettler. 1999. "The balanced scorecard: a necessary good or an unnecessary evil?" *European Management Journal* 17 (5):481-491.
- MSSS. 2012. Cadre de référence ministériel d'évaluation de la performance du système public de santé et de services sociaux à des fins de gestion.
- MSSS. 2019. Plan stratégique du ministère de la Santé et des Services sociaux 2019-2023.
- Nobre, Thierry. 2001. "Management hospitalier: du contrôle externe au pilotage, apport et adaptabilité du tableau de bord prospectif." *Comptabilité Contrôle Audit* 7 (2).
- Norreklit, Hanne. 2000. "The balance on the balanced scorecard a critical analysis of some of its assumptions." *Management accounting research* 11 (1):65-88.
- Pickton, David W, and Sheila Wright. 1998. "What's swot in strategic analysis?" *Strategic change* 7 (2):101-109.
- Porporato, Marcela, Peter Tsisis, and Luz Maria Marin Vinuesa. 2017. "Do hospital balanced scorecard measures reflect cause-effect relationships?" *International Journal of Productivity and Performance Management*.
- Scholey, Cam. 2005. "Strategy maps: a step-by-step guide to measuring, managing and communicating the plan." *Journal of Business Strategy*.
- Simons, R. 1994. "How managers use innovative control systems to drive strategic renewal." *Strategic Management Journal* 15 (3):169-189.
- Treacy, Michael, and Fred Wiersema. 2007. *The discipline of market leaders: Choose your customers, narrow your focus, dominate your market*: Basic Books.

CIGI 2021

Toward an Innovative Risk- and Opportunity-Oriented System for SMEs' Decision-Makers

ROMAIN BEN TALEB¹, MATHIEU DAHAN³, AURELIE MONTARNAL¹, MATTHIEU LAURAS^{1,3}, ROMAIN MICLO²

¹ Center of Industrial Engineering, University of Toulouse, IMT Mines Albi
Albi, France

romain.bentaleb@gmail.com; aurelie.montarnal@mines-albi.fr; matthieu.lauras@mines-albi.fr

² AGILEA

Toulouse, France

romain.miclo@agilea-group.com

³ School of Industrial and Systems Engineering, Georgia Institute of Technology
Atlanta, GA, USA

mathieu.dahan@isye.gatech.edu

Résumé – La crise de la COVID-19 a mis en exergue de nombreuses faiblesses dans les entreprises, notamment les petites et moyennes. Parmi ces faiblesses, il en est une symptomatique : l'incapacité des décideurs à prendre des décisions stratégiques robustes dans un contexte incertain et fortement perturbé. Pourtant, un tel contexte, peuplé de risques et d'opportunités, est à considérer comme une nouvelle normalité. Une analyse plus fine de la situation montre que les responsables d'entreprises utilisent principalement des outils issus de la Stratégie et du Contrôle de Gestion pour soutenir leurs décisions. Or, ces outils présentent de nombreuses limites telles que leur ancrage déterministe, leurs analyses ex post, ou encore leur dimension souvent qualitative de la gestion du risque. Le présent travail de recherche propose les bases d'un premier modèle original centré sur la gestion des actifs des entreprises et inspiré du Génie Industriel et de l'Intelligence Artificielle. Ce modèle vise à mieux appréhender les risques et les opportunités dans les mécanismes de prise de décision stratégique. La proposition est illustrée sur un cas d'école permettant de mettre en perspective son potentiel. Des axes de développement futurs sont finalement proposés pour alimenter la suite de ce travail de recherche débuté récemment.

Abstract – The COVID-19 crisis has demonstrated numerous weaknesses in Small- and Medium-sized Enterprises (SMEs). Among them, the incapability of top management to make robust strategic decisions in an uncertain environment composed of risks and opportunities, is probably one of the most critical. Unfortunately, such a disrupted context is now the norm. By analyzing the situation, we notice that most of these top management decision-makers use traditional Strategy and Management Accounting methods to make decisions for their business. However, these methods have numerous limits regarding the management of uncertainties. Notably, most of them are deterministic and reactive (i.e., a posteriori analysis) and they often manage risks through pure qualitative approaches. Based on these limitations, this article develops the fundamentals of a first innovative model focused on the management of assets and inspired by Industrial Engineering and Artificial Intelligence risk- and opportunity-oriented tools. The proposed approach intends to support more robust strategic decisions in disrupted contexts. An illustrative case is then developed in order to highlight the potentiality of our approach. Finally, some avenues for future research are exposed regarding the current work which is currently in its infancy.

Mots clés – Système d'Aide à la Décision, Risque, Opportunité, Contrôle de Gestion, PME.

Keywords – Decision Support System, Risk, Opportunity, Management Accounting, SME.

1 INTRODUCTION

The COVID-19 crisis revealed to the whole world how fragile the economy is with respect to uncertainties and disruptions. In particular, the economic consequences of this crisis are catastrophic for many Small- and Medium-sized Enterprises (SMEs). In France for instance, the economic growth of SMEs should drop by 15% for the year 2020 [INSEE, 3rd trimester 2020]. As explained by [Schumpeter, 1943], managing a company requires managing the impacts of uncertainties it

faces, whether they are negative (risks) or positive (opportunities). Unfortunately, catastrophes such as the financial crisis of 2007-2008, or the pandemic crisis of 2020 and later, seem to demonstrate that companies, and notably SMEs, are not properly equipped to efficiently cope with such uncertainties.

Essentially, the SMEs' decision-makers define and implement a strategy to ensure the proper functioning of their company. As shown by [Mintzberg, 1987], a strategy is composed of a plan,

a positioning, a perspective, a pattern, and a ploy (5P's). Currently, the SMEs' top management teams mainly use strategy and management accounting tools to make decisions, particularly with respect to the strategic time frame. However, most of these tools are either purely qualitative, or only provide limited quantitative insights derived from analytical and deterministic analyses. These tools mainly utilize descriptive analytics, i.e., they first analyze past information to determine the cause of bad performances, and then make corrective decisions for the future.

Nevertheless, each component of a company's strategy is potentially impacted by hazards, changes, and variabilities during their implementation. As a result, the current decision tools do not enable SME's managers to make robust decisions in what is henceforth known as the VUCA worlds (Volatile, Uncertain, Complex and Ambiguous) [Du and al. 2018.]. In addition, most of the decision-makers lack a culture of risk and opportunity, and have solely focused on maximizing efficiency [Peretti-Watel, 2005]. This results in decisions that are not sufficient for staying economically safe in case of disruption and for being sustainable, as demonstrated by the current COVID-19 crisis. Thus, it is of prime importance to consider the VUCA environment as a new normal for governance and decision-making.

Considering this, our research problem consists in studying how the SMEs' decision-makers tools can be improved to better manage the new normality of the economic world. Our main research claim is that a cross learning opportunity between various disciplines such as Strategy and Management Accounting on one hand, and Industrial Engineering and Artificial Intelligence on the other hand, can significantly improve the decision-making process. In particular, we aim to leverage the maturity of the Industrial Engineering and Artificial Intelligence disciplines in terms of uncertainty management to create prescriptive analytics tools. It is a potential avenue for enhancing the capabilities of SMEs to perform well and to become sustainable regardless of the risks and opportunities they have to cope with. Finally, our research objective is to develop an innovative risk- and opportunity-oriented decision-support system dedicated to SMEs' strategic decisions.

In this article, we intend to (i) position this contribution within the existing literature; (ii) describe the potential design of such a system; (iii) share a first illustrative example; and (iv) draft a roadmap for future research on this subject.

2 BACKGROUND

2.1 Major Strategic Decisions for SMEs

Among the resources available to decision-makers, environment analysis patterns can be used to identify risks, opportunities, strengths and weaknesses of their firm. [Knuston, 2018] reviews several of these tools. First, Business Model Canvas can help define value creation by taking into account stakeholders and available resources of the firm. Once this process is defined, analyzing and qualifying the environment is required. For example, the PESTEL pattern triggers a particular environment analysis based on Political, Economic, Social, Technological, Environmental and Legal dimensions and enables to identify opportunities.

[Aguilar, 1967] and [Learned et al., 1969] combined the PESTEL pattern with the SWOT (Strengths, Weaknesses, Opportunities, Threats) matrix. This matrix analyzes risks and opportunities regarding strengths and weaknesses of the firm. Another combinable analysis pattern is given by Porter's 5 (+1) competitive forces. They consider competitors on a firm's market or strategic activity domain and also help decision-makers to set up the firm's position, as one of Mintzberg 5P's described in the introduction section.

However, these analysis patterns mainly use qualitative inputs. One of the first tools using quantitative dimensions is the BCG (Boston Consulting Group) Matrix [Hofer et Schendel, 1978]. It consists in positioning each of the firm's strategic activity domains on a graph with two axes: relative market share and potential growth rate. Still, the BCG matrix can be combined with the above-mentioned analysis patterns as well others, such as the Technology and Product Life Cycle analyses.

Other analytical tools are available for top managers to base their decisions upon. For example, generic strategies defined by [Porter, 1997] set up three different types of strategies in a strategic activity domain: focus, differentiation, and cost domination. Each strategy type triggers specific decisions and actions in the firm. Similarly, different types of decisions are based on the Ansoff matrix [Ansoff, 1957] and depend on product and market innovation.

Thus, we have identified two steps for decision-making. First, an analysis of the environment is required to identify risks and opportunities. Then, the decision is made by taking into account the firm's strengths and weaknesses, and by considering the resulting impact from the risks and opportunities. For this reason, our proposed model will implement these two steps of ex-ante decision assessment.

2.2 Basics of Strategy and Accounting Management

At the beginning of the last century, [Fayol, 1917] defined top management work through 5 tasks: Predict, Organize, Command, Coordinate and Control.

Management Accounting is an ally of this vision as explained by [Anthony, 1965], who describes it as the process that ensures the efficient use of resources to reach the firm's objectives. Moreover, each management function of organizational structure that assists top management (Human Resources, Purchase Department, Marketing, etc.) can be controlled by management accounting with specific methods [Mintzberg, 1993].

Plans in the organization have been managed by Accounting through the budgeting process. Each function of the firm has to prepare its budgets and only top management is able to sum up the predicted resources engaged in the following year(s). However, the most common methods for budgeting are limited by this specialization and lack of interdependence between the firm's departments. Moreover, a large subjective part of estimations is etched in these documents.

For this reason, a significant body of work in management accounting has been developed following the Balanced Scorecard [Kaplan and Norton, 1997]. This method enables a

thorough evaluation of a firm's performance by combining decisions and four types of objectives in the value creation process: learning, processes, customers and finance.

Further research in management accounting opens another gate in systemic approach. For example, combining Simon's levers of control and Mintzberg's 5P's shows that a firm is a system where management accounting is an enabler and a controller. This enables and controls culture, performance, position and processes.

Management Accounting is thus an interface between strategic decisions and their implementation. However, the current work lacks a quantitative systemic approach, which drives our current research.

2.3 How Strategy and Accounting Management Consider Risk and Opportunities?

As explained in previous paragraphs, usual decision tools take primarily qualitative inputs to produce qualitative outputs. Quantifying the impact of a risk or an opportunity is the final step of a quantitative process that carries uncertainty and subjective information at each step. One example is the risk matrix [Thomas et al., 2013] in which the identified risks are quantified. The quantitative impact of a risk is the product of its potential occurrence and gravity. Although the occurrence of a risk can be statistically estimated, no quantitative methods that evaluate its impact on a firm's performance have been developed, to the best of our knowledge. At best, a firm can simply refer to past risks faced by comparable firms, and estimate the corresponding impacts.

In the context of risk management in the strategic and accounting dimensions, we aim to propose a model to predict the impact of a risk on a firm's performance and the cascading effects on the firm's assets.

2.4 How Industrial Engineering and Artificial Intelligence Consider Risk and Opportunities?

[Zeng et al., 2017] define a risk, in its usual understanding in the Industrial Engineering domain, through three main components: (i) a driver that triggers the risk, (ii) an event with a probability of occurrence, and (iii) the resulting consequences. The authors go further by defining the Risk with Casual Relationship (RCR) model in which the drivers can be both triggering factors, defined as "uncontrollable or immeasurable" causes but also risks themselves, defined with a probability. The RCR model also considers the consequence as either a direct financially quantifiable impact or another subsequent risk (with a potentially bigger consequence). In line with this, research works such as [Fang et al., 2012] aim to model and evaluate the interactions between risks and their cascading effects, using network theory in this particular case, and provide decision-makers with an enhanced visibility about the trajectory of a project.

Essentially, [Benaben et al., 2019] came up with two statements: (i) the usual risk representation above-mentioned remains simple and does not fully support the cascading effect and (ii) the step of risk identification remains a laborious human task based on reviews, interviews or questionnaires.

In the field of supply chain risk management where risk analysis is a cornerstone activity, [Baryannis et al., 2019] suggest that

the use of a proactive approach to predict the occurrence and impact of risks is necessary and that, considering the large datasets it should utilize, AI-based methods are quite relevant. The authors define several categories of candidate AI solutions such as: continuous or scenario-based stochastic programming, fuzzy programming, robust optimization, hybrid mathematical programming, machine learning and big data, and network-based and agent-based reasoning. [Baryannis et al., 2019] observe that only 8% of the articles studied in their survey define a decision support system, and 4% combine risk identification and assessment. Thus, risk identification remains a hardly covered topic that requires solutions using for instance network-based models, automated reasoning and machine learning/big data analytics. Interestingly, the authors describe a lessened interest of managers in techniques that do not provide any decision-making capability while researchers are inclined to use proven mathematical programming solutions rather than other AI techniques.

In the field of accounting, [Sutton et al., 2016] urge AI researchers to deepen the efforts: while knowledge-based systems did not survive in the field, there are new promises enabled by the recent machine learning techniques – especially natural language processing, considering that text has become a very large source of data in accounting activities – that have to be considered.

This overview of the state-of-the-art shows that decision-making should be based on a proper risk analysis including their identification and impact assessment, which can only be relevant when considered in the entire firm's environment. In addition, accounting management is at the interface between decision-making and proper decision implementation, and therefore takes an important place in the risk management process that has to be led. AI comes with big promises and expectations in this regard.

3 PROPOSAL

3.1 The Process

Following a traditional risk management lifecycle [ISO 31000, 2009], we suggest a seven-step approach to support our ambition as shown in Figure 1.



Figure 1. The risk and opportunity management process (inspired by ISO 31000)

Basically, the first step consists in identifying the trigger events which are in our case, risks or opportunities. These events might be endogenous (internal processes) or exogeneous (external environment).

The second step is about assessing the potential impacts and consequences of each individual or combination of risks and/or opportunities on the one hand, and the probability of occurrence on the other hand. Usually this is called a criticality evaluation even if in our case, this criticality can be negative or positive.

The third step consists in identifying concrete alternatives (actions, options, etc.) which can bring a solution to mitigate the risks and/or to catch the opportunities if needed.

The fourth step leads to the selection phase. During this phase, evaluation of all alternatives (scenarios) will be made in order to assess the potential benefits that each of them can bring to the SME, along with the associated costs (in cash, resources, etc.).

During the fifth step, the decision-maker will select the best alternative for them, i.e., the scenario that provides the optimal tradeoff between the costs and the benefits among all the identified alternatives.

The sixth and seventh steps represent the implementation of the decision in the field and the gathering of associated feedbacks to close the loop with the first step.

3.2 The Model

Given our summary in Section 2.1, our proposal is based on the paradigm that all SMEs' strategic decisions are driven by the assets of the company. Specifically, each strategic decision of a manager can be categorized as:

- Purchasing a new asset,
- Removing an existing asset,
- Updating an existing asset to optimally exploit it.

According to management accounting research, a company manages at least 7 main categories of assets if we consider capitals identified in EFQM framework [Trebucq and al., 2017]:

- Cashflow,
- Human resources,
- Material resources,
- Information Technologies,
- Business processes and abilities,
- Brand and image,
- Customers' portfolio.

For each of them, the challenge for the decision-maker will be to apply the process described in Figure 1 to objectively assess the positive or negative impacts of an internal or external trigger event on an asset or a group of assets. In particular, risks and opportunities might have different features for each category of assets and must be managed in a tailored manner.

Based on this approach, we developed the model presented in Figure 2. This model shows how an endogenous or exogeneous trigger event (e.g., a change in the demand, a delay in a supply, new bank conditions, a human resource application, etc.) can generate a risk or opportunity for one or several assets of the SME. This risk or opportunity will imply potential changes in the asset's status and its associated performance level. Therefore, a specific decision could be deduced to benefit from this new status or to avoid potential performance decrease. This decision relates to one of the above-mentioned three generic

behaviors that a decision-maker may have on an asset [IFRS, 2020].

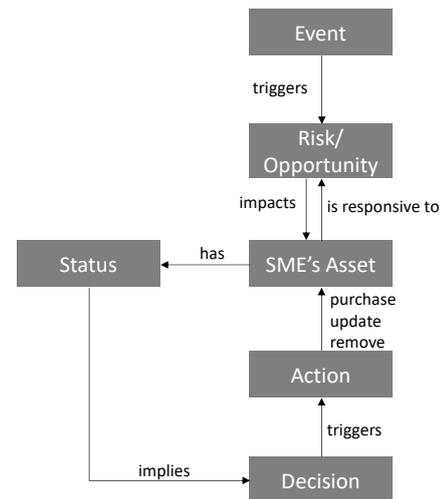


Figure 2. The proposed model

All the associations (verbs) described in Figure 2 must be applied to past, current, and potential situations. For instance, this model can be applied in reaction to an existing situation or a projected one.

As an illustration, Table 1 develops a set of potential strategic decisions that an SME can activate to mitigate some risks or leverage opportunities according to the model described in Figure 2.

Considering this model, two main stakes can be identified. The first one consists in providing capabilities to detect and evaluate as soon as possible current or future potential risks and opportunities that can have an impact on the SME's assets. The second one consists in making a robust decision in such a situation to maximize efficiency while avoiding potential failure or degradation of the SME's performances.

Regarding the detection stake, and as indicated in Section 2, existing decision support systems used by the majority of SMEs at the strategic level gather historical information to analyze past situations and make corrective actions to go toward a better position. In other words, this kind of approach is mainly reactive and descriptive, and entails managing the SMEs' business with ex-post Key Performance Indicators (KPIs). Our ambition with the current proposal is to move toward a more proactive and predictive step implying ex-ante KPIs and management.

Regarding the decision-making stake, our proposal consists in exploiting the causal relationships described in Figure 2 to better prepare the future of SMEs. Specifically, we propose to secure long-term perspectives with respect to a large set of potential scenarios. In practice, this consists in predicting a set of potential futures regarding ongoing or potential risks or opportunities in order to support a large, qualitative and quantitative "what-if" perspective. Ultimately, we aim to derive a prescriptive approach that provides dynamic, wide, and effective contingency plans.

3.3 The Tools

Regarding the suggested process and framework, two main obstacles should be considered for selecting the most appropriate decision-support tools from existing literature and practices in Industrial Engineering and Artificial Intelligence. The first one relates to the core capabilities that the tools are

Table 1. Examples of decisions

ASSETS	EXAMPLES OF DECISION VARIABLES				
	Purchase	Update			Handover
		Availability	Ability	Capacity	
Cashflow	Get a loan	Transfer money to checking account	Repurchase some shares	Reduce customer payment period term	Invest
Human Resources	Hire someone	Accept or Decline a day-off	Train someone	Train someone	Fire someone
Material Resources	Buy an equipment machine	Procure raw material	Change the machine adjustments	Open a new shift	Sell an equipment
Information Technologies	Buy a software	Contract with a third-party provider	Adapt the parameters	Buy additional tokens	Do not renew a licence
Business Processes and Abilities	Design a new process	Plan the activities	Adapt the means of an activity	Adapt the means of an activity	Stop an activity
Brand and Image	Buy an existing brand	Buy a domain name (Internet)	Change a logo	Do a social network campaign	Sell a brand
Customer's Portfolio	Buy a new company	Create key account manager roles	Develop a Customer Relationship Management Project	Start a promotional campaign	Sell a customer portfolio to another company

supposed to have and the second one is about their abilities to estimate the future. Let us start with the identification of the main abilities that are required to execute the process described in Section 3.1. Basically, such an instantiation needs to be able to:

- Identify, detect or predict the information pertaining to the risks and opportunities to be assessed on one hand, and to the data and parameters of the evaluation models (i.e., models able to assess the past/current or potential future status and performance of each asset) on the other hand. To achieve this goal, we can use traditional tools such as statistics or more advanced ones such as data mining or machine learning. Naturally, more qualitative techniques such as Natural Language Processing or Knowledge-Based Engineering might also be useful.
- Assess quantitatively and/or qualitatively the potential consequences of an occurrence of a risk or an opportunity on a given asset or group of assets. This ability could be activated through the use of quantitative techniques such as Heuristics or Simulation on one hand, and of qualitative techniques such as Expert Systems or Model Driven Engineering on the other hand.
- Design alternative scenarios able to mitigate the risks or to benefit from opportunities. This task can be performed manually or automatically depending on the context, with the help of modeling tools from Operations Research (Simulation and Optimization models mainly). However, learning from the past through Machine Learning techniques or Expert Systems (rule-based) might also be a good option. Finally, using a Model-Driven Engineering approach could be an interesting option to support the design step from a qualitative perspective.
- Evaluate the different alternative scenarios to assess a priori the expected performance of each scenario. Here, quantitative techniques should be preferred such as Optimization, Simulation and Heuristics. It is important to notice that this evaluation can be carried out with respect to both benefits (i.e., the stakes) and costs (i.e., accessibility). Several performance dimensions might be considered and should be driven by the typology of the associated assets.
- Compare the results of the different scenarios in order to support the final choice for the decision-maker. This ability relates to the overall performance of each scenario in terms of stakes and accessibility, or in terms of Return-On-Investment. Decision-support tools such as Multi-Criteria Decision Making, Decision Trees, Data Mining could be used accurately to support this ability.
- Monitor and control the execution of the plan in the field. This ability could be supported through Data Mining or Online-Flow Simulation systems.

Table 2 summarizes our vision of potential decision support tools extracted from the Industrial Engineering and Artificial Intelligence disciplines that can be used to concretize our proposal. Basically, techniques such as Machine Learning (reinforcement, supervised, unsupervised), Flow Simulation (discrete event simulation, multi-agent, system dynamics) and Optimization models are excellent candidates as they have the ability to proact ahead instead of “only” learning from the past.

One important justification of this model comes from the international accounting rules (IFRS) and the key concept of assets. This work could be a bridge between Industrial Engineering and Accounting research. Using different forms of capitals identified in Management Accounting research and converting them into relevant objects both in the Accounting and Industrial Engineering framework is an illustration of this

Table 2. The potential toolbox to solve the proposed model

Process Step	Toolbox											
	Data Mining	Machine Learning	Statistics	Natural Language Processing	Knowledge Based Engineering	Model Driven Engineering	Expert Systems	Heuristics	Optimization	Simulation	Decision Trees	Multi Criteria Decision Making
Identify Risks and Opportunities	X	X	X	X	X	X				X		
Assess Impact and Occurrence		X				X		X		X		
Define Alternatives		X			X	X	X	X	X	X		
Evaluate Benefits vs Costs							X	X	X	X		
Make Decision	X		X				X		X		X	X
Implementation												
Control	X									X		

purpose. We claim that this first concept of assets is necessary to achieve this goal. However, a more complex model involving contracts between assets may be able to propose a more sophisticated system in order to interface the planning and simulation methods from Industrial Engineering with this first model.

The purpose of this proposed model is to develop a tool able to connect firms and their environment and assess a decision implementation regarding the performance criteria chosen by top management. Integrating common externalities in the assets' attributes and firms' processes seems to be an opportunity to enable decision-makers to assess multi-impact decisions and multi-objective achievements.

4 ILLUSTRATIVE EXAMPLE

In this section, we illustrate our approach for the ex-ante evaluation of decisions using two examples involving a risk and an opportunity. We aim to highlight the impacts on different assets that can be possibly quantified, predicted and optimized with Industrial Engineering and/or Artificial Intelligence common processes.

We first consider the example of a firm that produces computers and suppose that a new component provider is entering the market. To make a decision regarding this event, the decision-maker must evaluate its impact on the firm's assets. Since this event can positively impact the firm's ability to get components for its production process (Material Resources), purchase costs (Cashflow) and production quality (Business Processes and Abilities), it can be qualified as an opportunity.

Making a decision in regard to this opportunity depends on the status of the firm's different assets. The status is defined by the availability, capacity and ability of the affected assets.

Moreover, the status is monitored through a performance measurement. In the case of a new supplier, the information provided in the offer can be leveraged to evaluate the impact on the assets of a decision regarding this event. Particularly, opportunities can be evaluated on their impact on Cashflow. The identification of this new supplier triggers alternatives to the current purchase options that should be evaluated.

Designing the flow of impacts on the firm's assets status triggered by the event as a system can enable the choice between not purchasing new components, including the new components and keeping the current suppliers, or exclusively purchasing from the new supplier.

Basically, the costs associated with a decision are evaluated on the main impacted asset. For example, getting components from a new supplier might be evaluated in terms of their purchasing costs. An ex-ante evaluation of a decision entails taking into account all of the decision's consequences. In this case, buying components from a new supplier can require updating the processes for delivery reception, can impact components storage in the plant, and may need to update the information system procedures for command and invoicing.

More complex impacts can also be evaluated by taking into account lead-time and quality. An increase in component quality can reduce non-quality costs and thus production costs. Regarding the lead-time and quality improvement of components, the firm might also increase its own products' lead-time and quality and increase its prices, its sales and thus its turnover. Moreover, a new supplying contract can be a new asset for the firm depending on its ability to generate revenue or savings and to reduce risks of shortages. An ex-ante evaluation can predict the complete cost of the decision and value creation for the firm.

The flow of impacts on different assets has to be processed ex-ante to make the best decision triggered by the event qualified as risk or opportunity.

Evaluating a risk is driven by the same ex-ante evaluation to decision-making. For example, the risk of a fire at a provider's plant has a probability to affect buyers purchase ability and thus

its production process, its turnover, and its brand. An ex-ante evaluation requires to set up the flow of impacts on different assets and their features. This model also predicts the complete cost of the risk and the associated decisions. Stopping the contract with the supplier, purchasing from another supplier, or updating the contract in

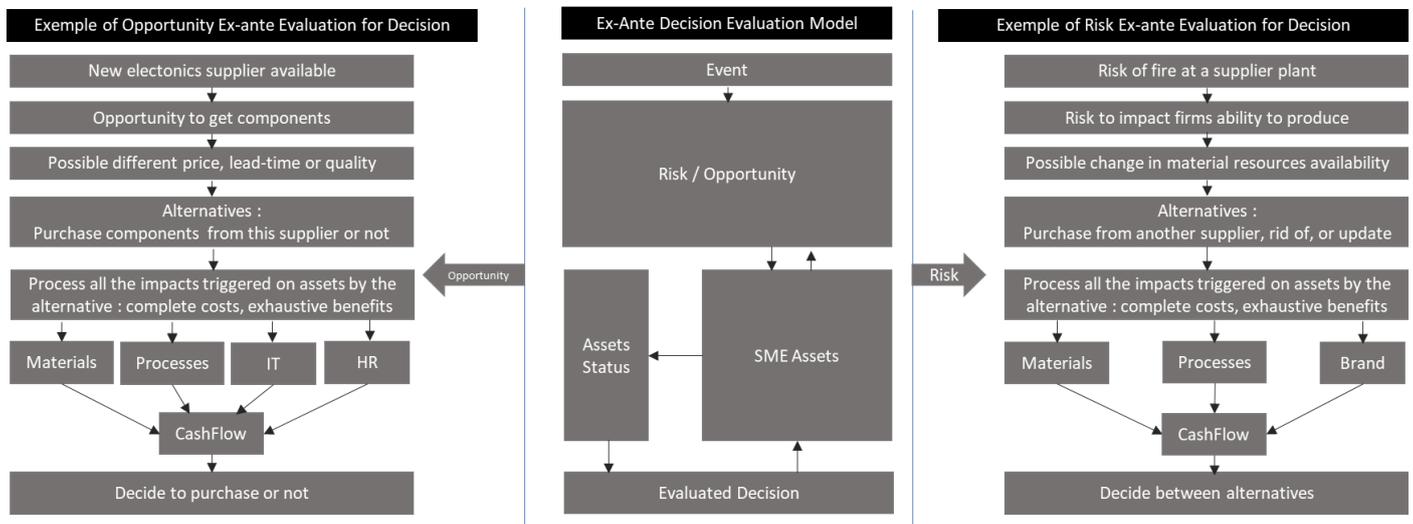


Figure 3. Examples of risk and opportunity decision process

a supplier development strategy can be evaluated ex-ante. This last option could lead to more research because contracts are assets for the firm and increasing their value could be a lever of value creation. As in opportunity evaluation, the best decision regarding future cashflows triggers choice.

To conclude this example, our model highlights a parallel process for risks and opportunities and thus an ex-ante evaluation of decisions as it is evaluated on its impact on the firm's assets, their features, and their ability to trigger cashflow.

5 CONCLUSION AND PERSPECTIVES

In this article, we have argued that the current top manager decision support tools are not appropriate for the management of risks and uncertainties that SMEs have to cope with. Particularly, the methods issued from Strategy and Management Accounting disciplines present some critical limitations such as their deterministic and *a posteriori* features. Based on this diagnosis, we have presented the first characteristics of an innovative model able to serve as a basis for future SME's top manager decision support systems developments. The main assumption of this research work consists in changing the decision paradigm of top managers by using both a risk-management process and a set of predictive and prescriptive techniques capable of better exploiting uncertainties. Industrial Engineering and Artificial Intelligence backgrounds should be relevant sources of inspiration to go further.

To sum up, this work is still in its infancy. As a consequence, there are numerous avenues for research. First, the asset-oriented approach should be tested and validated with several

real cases. Second, a finite set of typologies of risks and opportunities must be defined to enable the implementation of the approach in the field. Third, additional investigations

regarding the decision categories associated with each asset of a company should be conducted. Fourth, each association of the suggested model presented in Figure 2 should be supported by a specific decision-support tool able to deliver a quick, relevant and accurate decision-support support mechanism. Further research work should consequently be done to both define the technical requirements and select the appropriate decision-support tools for each association of the model. For instance, some specific online predictive algorithms might be derived to detect in real time the risks and opportunities that top managers should account for regarding their business.

6 REFERENCES

- Aguilar, F. J. (1967). *Scanning the business environment*. Macmillan.
- Alcouffe, S. (2004). *La diffusion et l'adoption des innovations managériales en comptabilité et contrôle de gestion : le cas de l'AFC en France* (Doctoral dissertation).
- Ansoff, H. I. (1957). Strategies for diversification. *Harvard business review*, 35(5), 113-124.
- Anthony, R. N. (1965). *Planning and control systems: A framework for analysis [by]*. Division of Research, Graduate School of Business Administration, Harvard University.
- Baryannis, G., Validi, S., Dani, S., & Antoniou, G. (2019). Supply chain risk management and artificial intelligence: state of the art and future research directions. *International Journal of Production Research*, 57(7), 2179-2202.
- Benaben, F., Montreuil, B., Gou, J., Li, J., Lauras, M., Koura, I., & Mu, W. (2019, January). A tentative framework for risk and opportunity detection in a collaborative environment based on data interpretation.
- David, A., Hatchuel, A., & Laufer, R. (2012). *Les nouvelles fondations des sciences de gestion: éléments d'épistémologie de la recherche en management*. Presses des MINES.

- Du, J., & Chen, Z. (2018). Applying Organizational Ambidexterity in strategic management under a “VUCA” environment: Evidence from high tech companies in China. *International Journal of Innovation Studies*, 2(1), 42-52.
- Fang, C., Marle, F., Zio, E., & Bocquet, J. C. (2012). Network theory-based analysis of risk interactions in large engineering projects. *Reliability Engineering & System Safety*, 106, 1-10.
- Fayol, H. (1917). *Administration industrielle et générale: Prévoyance, organisation, commandement, coordination, controle*. Paris: H. Dunod et E. Pinat.
- Hofer, C. W., & Schendel, D. (1978). *Strategy formulation: Analytical concepts*. West Publ..
- IFRS, Updating a Reference to the Conceptual Framework (Amendments to IFRS 3), <https://www.ifrs.org/projects/2020/updating-a-reference-to-the-conceptual-framework-ifrs-3/>
- ISO, *ISO 31000:2009 Management du risque - Principes et lignes directrices*, <https://www.iso.org/fr/standard/43170.html>
- Kaplan, R. S., & Norton, D. P. (2005). The balanced scorecard: measures that drive performance. *Harvard business review*, 83(7), 172.
- Max, K. (2018). *Three Business Plan Models to Help a Start-up with Strategic Analysis: Lessons from using SWOT, Porter's Five Forces, and Price Equalizer to write a business plan (Dissertation)*. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-76712>
- Learned, E. P. (1969). *Business policy: Text and cases*. RD Irwin.
- Le Roy, F., & Pellegrin-Boucher, E. (2005). Bruce Henderson comme fondateur de la pensée stratégique. *Revue française de gestion*, (1), 9-20.
- Max, K. (2018). *Three Business Plan Models to Help a Start-up with Strategic Analysis: Lessons from using SWOT, Porter's Five Forces, and Price Equalizer to write a business plan*.
- Mintzberg, H. (1987). The strategy concept I: Five Ps for strategy. *California management review*, 30(1), 11-24.
- Mintzberg, H. (1993). *Structure in fives: Designing effective organizations*. Prentice-Hall, Inc.
- Nobre, T. (2001). Méthodes et outils du contrôle de gestion dans les PME. *Finance contrôle stratégie*, 4(2), 119-148.
- Peretti-Watel, P. (2005). La culture du risque, ses marqueurs sociaux et ses paradoxes. *Revue économique*, 56(2), 371-392.
- Point de conjoncture du 8 septembre – Présentation – Points de conjoncture 2020 | Insee. (n.d.). Retrieved November 26, 2020, from <https://www.insee.fr/>
- Porter, M. E. (1989). How competitive forces shape strategy. In *Readings in strategic management* (pp. 133-143). Palgrave, London.
- Porter, M. (1982). *Choix stratégiques et concurrence techniques d'analyse des secteurs et de la concurrence dans l'industrie*. Economica.
- Porter, M. E., & de Lavergne, P. (1986). L'avantage concurrentiel.
- Porter, M.E. (1997), "COMPETITIVE STRATEGY", *Measuring Business Excellence*, Vol. 1 No. 2, pp. 12-17. <https://doi.org/10.1108/eb025476>
- Schumpeter, J. A. (1943). *Capitalism, socialism and democracy*. routledge.
- Simons, R. (1994). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Harvard Business Press.
- Sutton, S. G., Holt, M., & Arnold, V. (2016). “The reports of my death are greatly exaggerated”—Artificial intelligence research in accounting. *International Journal of Accounting Information Systems*, 22, 60-73.
- Thomas, Philip & Bratvold, Reidar & Bickel, J.. (2013). The Risk of Using Risk Matrices. *SPE Economics & Management*. 6. 10.2118/166269-MS.
- Trébucq Stéphane, Elisabetta Magnaghi, *Using the EFQM excellence model for integrated reporting: A qualitative exploration and evaluation*, *Research in International Business and Finance*, Volume 42, 2017, Pages 522-531
- Zeng, B., & Yen, B. P. C. (2017). Rethinking the role of partnerships in global supply chains: A risk-based perspective. *International Journal of Production Economics*, 185, 52-62.

Vers un système d'évaluation de la performance globale et multidimensionnelle de l'usine du futur

MARIE-HELENE GENTIL¹, CATHERINE MERLE²

¹ Maître de conférences
IUT de Bordeaux/Université de Bordeaux, France
marie-helene.gentil@u-bordeaux.fr

² Maître de conférences
IUT de Bordeaux/Université de Bordeaux, France
catherine.merle@u-bordeaux.fr

Résumé

Le monde industriel, à l'image de notre monde, est en pleine transformation. Quels seront les impacts sur l'évaluation de la performance dans l'usine du futur ? L'objectif de cette communication est de proposer aux décideurs une méthode d'évaluation de la performance à travers l'exploration des relations croisées entre les multiples approches de la performance et l'évolution de l'entreprise.

Cette méthode permettra de prendre en compte l'impact des technologies sur les performances technologiques, financières, organisationnelles, environnementales, sociales, et sociétales.

Nous proposons un modèle multi-niveaux et multidimensionnel d'évaluation de la performance de l'usine du futur applicable dans toute structure industrielle.

Abstract

The industrial world, like our world, is changing. What will be the impacts on performance assessment in the factory of the future? The objective of this communication is to provide decision-makers with a method for evaluating performance through the exploration of the interrelationships between multiple approaches to performance and the evolution of the company.

This method will take into account the impact of technologies on technological, financial, organizational, environmental, social, and societal performance.

We propose a multi-level and multidimensional model to evaluate the performance of factory of the future applicable in any industrial structure

Mots Clés

Usine du Futur - Evaluation de la Performance - Méthode - Tableau de bord

Keywords

Factory of the Future - Performance Evaluation - Method - Dashboard

1 INTRODUCTION

L'usine du futur est souvent présentée comme une organisation ultra moderne par rapport à l'aspect technologique, numérique, et connectée. Elle est caractérisée par différentes technologies de type robots, cobots, fabrication additive, réalité virtuelle ou augmentée, réseaux de capteurs, logiciels de conception, etc... L'objectif de cette mutation est de permettre aux entreprises de rester compétitives et performantes.

Historiquement, la performance de l'entreprise était centrée sur les aspects économiques et financiers. Aujourd'hui, une vision globale et élargie de l'entreprise est indispensable pour satisfaire toutes les parties intéressées. Logiquement une vision multidimensionnelle permettrait de prendre en compte d'autres types de performance telle que par exemple la performance sociétale (environnementale, sociale, économique) ...

Comment la performance évolue et s'évalue dans l'usine du futur ? Quelle pourrait être la méthode qui permettrait d'évaluer au mieux l'usine dans le futur ? Comment intégrer la performance technologique de plus en plus présente dans l'usine du futur ?

Cette communication est découpée en deux parties. Dans la première partie nous définirons les concepts d'usine du futur et les différentes approches de la performance. Ce qui nous permettra de préciser la typologie de la performance choisie. Puis, nous présenterons dans une deuxième partie notre approche multi-niveaux d'évaluation de la performance dans le cadre de l'usine du futur.

2 CONTEXTE D'EVALUATION DE L'USINE DU FUTUR

2.1- Usine du futur

Le terme « Industrie 4.0 » a vu le jour en 2011, lorsqu'il a été cité pour la première fois au forum d'Hanovre suite à une

discussion entre les représentants de l'industrie, de la recherche, des syndicats et de l'état. Ce mot évoque à la fois l'industrie du futur, la Smart Factory ou bien encore l'usine connectée. [Blanchet et Varin, 2016], [Kohler et Weisz, 2016], [Julien et Martin, 2018], [André, 2019].

Plusieurs initiatives stratégiques sont nées à travers le monde : le « smart manufacturing » aux Etats-Unis, « l'Internet+ » en Chine, « l'industrie du futur » en France et « l'industrie 4.0 » en Allemagne. Le terme industrie 4.0 semble s'imposer au niveau international. L'industrie 4.0 vise à connecter l'ensemble des objets et intervenants tout au long de la chaîne de valeur au sein de l'usine de demain

L'objectif premier de l'initiative allemande n'est pas d'augmenter l'automatisation des productions, mais de rendre plus intelligents les modes de production à travers la mise en réseau des machines et des hommes.

L'usine du futur s'appuie sur de multiples technologies mises en place dans les entreprises. Le **cobot** (Collaborative Robot), ou robot coopératif est un robot destiné à produire une transformation matérielle tout en interagissant physiquement avec des humains dans un espace partagé. La **fabrication additive**, plus communément appelée impression 3D, permet de fabriquer des pièces complexes voire très complexes en un temps record. La **réalité augmentée**, contrairement à la réalité virtuelle, ajoute des informations dans notre champ visuel. En effet, elle affiche des informations en superposition de l'environnement visible (le monde réel). Elle permet d'ouvrir un large horizon aux simulations et reconstitutions. **L'internet des objets** (IoT : internet of things) caractérise des objets physiques connectés ayant leur propre identité numérique et capable de communiquer les uns avec les autres. Les nouvelles technologies de l'information, la miniaturisation des objets informatiques et la baisse des consommations énergétiques permettent d'incorporer des capacités de communication aux différents objets physiques. **L'intelligence artificielle** (deep learning) offre la possibilité d'exploiter intelligemment le « Big Data » (concept qui permet de traduire l'augmentation du volume de données à traiter par les entreprises). Cette technologie a pour objectif de traiter et d'analyser des données en temps réel. Les **jumeaux numériques** sont des copies virtuelles de tout systèmes, machines, processus, lignes de production, site. Ils évoluent en temps réel comme leurs jumeaux ou en accéléré, ce qui permet d'anticiper le comportement fonctionnel ou comportemental et améliore le pilotage, l'optimisation, et donc la performance du système jumelé.

A l'heure actuelle, les consommateurs souhaitent des produits de plus en plus personnalisés voire même uniques. L'industrie 4.0 représente non seulement ces nouvelles technologies mais aussi l'entrée dans l'ère de la personnalisation des produits. Grâce à cette révolution, les entreprises peuvent répondre à l'exigence d'un produit unique et personnalisé, tout en gardant un coût acceptable malgré les faibles volumes de production. L'industrie 4.0 est la représentation du pilotage industriel qui vise la synchronisation en temps réel des flux physiques et d'information et de la production des produits uniques et personnalisés à partir de la demande des clients. La mise en

œuvre de ces technologies participe à la recherche continue de compétitivité, d'amélioration et de performance.

En parallèle des changements technologiques, la place de l'homme et ses missions vont forcément évoluer de façon importante. Globalement, cela devrait entraîner une diminution de la « charge physique » (moins de pénibilité) et une augmentation de la « charge mentale et cognitive » (autonomie, responsabilisation). Le rôle des acteurs sera plus collectif et actif (intelligence collective, entreprise apprenante voire libérée), [Ivanov et al., 2020].

2.2- Le concept de performance

Le terme performance vient du latin et a pour origine le terme « performare » qui signifie « former entièrement ». Ce terme fait référence à la réalisation d'une action.

Selon Philippe Lorino [Lorino, 2003] : « est performance dans l'entreprise tout ce qui, et seulement ce qui contribue à atteindre les objectifs stratégiques ». La performance d'une entreprise s'articule aussi autour de l'amélioration du couple « valeur-coût » et tend ainsi vers la maximisation de la création nette de valeur.

Selon l'ISO 9000-2015, la performance est un « résultat mesurable » qui peut porter sur des éléments quantitatifs ou qualitatifs tout en restant basée sur des mesures [Iso9000, 2015].

De nombreux modèles et démarches de représentation de la performance sont traités dans la littérature [Camara et al., 2013] Aucun ne traite spécifiquement des aspects multidimensionnels de la performance de l'usine du futur.

Il existe également de multiples points de vue de la performance : « globale-locale », « valeur-coût », « interne-externe », « statique-dynamique », « coût-qualité-délai », « efficacité/efficience », productivité, flexibilité, réactivité, etc . [Estampe, 2015], [Hamadmad, 2017], [Moeuf, 2018].

La performance d'un processus ou d'un projet dépend à la fois de la qualité (le projet ou le processus doit répondre aux normes/exigences de qualité), du coût (respect du coût minimum), et du délai (le plus court possible).

L'efficacité est le degré de réalisation des objectifs. L'efficience est le rapport entre les ressources employées et les résultats atteints. L'effectivité est l'adéquation des objectifs, des moyens et des résultats au regard de la finalité du système. [Sénéchal, 2004].

La productivité est le rapport entre la production réalisée et les ressources mises en œuvre pour sa réalisation.

Une entreprise agile est capable de s'adapter rapidement à des changements inattendus de son environnement et recherche en continu la création de valeur. L'entreprise proactive doit être tournée vers le futur. Pour R. Pérez [Perez, 1995], une entreprise proactive est une entreprise qui ne « doit pas seulement s'adapter à des contraintes de marché, mais tenter de modifier ces contraintes en exploitant tout son potentiel dynamique ».

La performance est donc un concept complexe et multidimensionnel qui entraîne une perception différente des acteurs, des niveaux, de la structure, et des objectifs et par voie de conséquence des indicateurs, des variables d'action et des contraintes. C'est dans ce sens que la performance peut se décliner en fonction des différents angles, points de vues, aspects. Chaque décideur doit pouvoir construire son univers de performance et le faire évoluer dans un but ultime de pérennité, de compétitivité et d'amélioration.

Il existe différentes perceptions de la performance à chaque niveau décisionnel : stratégique, tactique et opérationnel. Nous avons choisi de retenir une approche globale de la performance en identifiant les critères déclinables aux trois niveaux décisionnels en termes de points de vues technologique, environnemental, social, économique, sociétal, financier et organisationnel.

Nous avons choisi de reprendre les trois niveaux de représentation de l'entreprise d'un point de vue de l'analyse systémique. Le niveau global prend en compte les objectifs du système, ses composants et leurs interactions, ses liens avec l'environnement, son infrastructure. Le niveau intermédiaire représente le comportement en utilisant le concept de processus ou de projet. Le niveau local met en évidence le comportement détaillé au travers du concept d'activité.

L'approche multidimensionnelle de la performance sera mise en cohérence avec la représentation des trois niveaux de l'usine du futur.

3 PROPOSITION POUR UNE EVALUATION DE LA PERFORMANCE DE L'USINE DU FUTUR

Dans la partie précédente, nous avons expliqué qu'il n'existe pas une vision « unique et universelle » de la performance. Au contraire, suivant le niveau décisionnel de l'entreprise et l'angle que l'on souhaite examiner, cette vision peut être différente. Elle dépend également des valeurs, des contraintes, des opportunités, des choix faits par l'entreprise. Elle peut même être différente suivant le moment ou la période examiné.

C'est pourquoi nous proposons une démarche inspirée des démarches de management des risques [Iso31000, 2018] ou management de la Qualité [Iso9001, 2015]. L'objectif est d'aboutir à la conception de tableaux de bord déclinables aux trois niveaux décisionnels en respectant les multiples points de vue de la performance. Dans le cadre de l'usine du futur, l'utilisation de jumeaux numériques par exemple permettra d'améliorer la traçabilité et le traitement des données, la génération des tableaux de bord, l'analyse des résultats obtenus et le reporting. Le jumeau numérique anticipera les résultats des indicateurs et ainsi, les délais de réponse seront

diminués, favorisant une réaction plus rapide aux variations diverses.

Sur la figure 1 présentant les grandes phases de la démarche, les étapes numérotées 1, 2, et 3 seront développées dans la suite du document.

Sur la base de la vision établie au plus haut niveau décisionnel, la performance sera évaluée au travers des indicateurs de performance déterminés à long terme au niveau stratégique (niveau usine du futur), à moyen terme (niveau tactique, processus ou projet) et court terme (niveau opérationnel, postes de travail). Les dimensions évaluées seront sélectionnées selon les besoins, attentes et valeurs de l'entreprise, son contexte et les parties intéressées pertinentes. La revue de performance sera un bilan régulier pour réévaluer les besoins d'amélioration du système en place. Le jumeau numérique anticipera les résultats du système réel. Tout le système pourra être amélioré si besoin et/ou régulièrement selon le résultat des « revues de performance » et du jumeau numérique. Tout ceci repose sur le principe du cycle PDCA (Plan, Do, Check, Act) [Deming, 1996] qui permet de planifier, mettre en œuvre, vérifier et adapter si nécessaire le système d'évaluation de la performance.

3.1 - Détermination de la vision stratégique, tactique et opérationnelle

Avant tout, il faut que l'entreprise détermine le domaine d'application, c'est-à-dire délimite le champ, l'étendue sur laquelle sera effectuée la future évaluation de la performance. Puis l'entreprise devra définir son contexte interne et externe. La prise en compte du contexte et des parties intéressées pertinentes vont permettre de définir les principaux axes à prendre en compte dans le futur cadre de l'évaluation de la performance. L'usine du futur ne peut se contenter de la vision historique et traditionnelle qui consistait à considérer exclusivement la voix des clients et fournisseurs. On parle désormais de « sphère d'influence » constituée des acteurs à la fois internes (personnel, syndicats...) et externes (clients, fournisseurs, actionnaires...).

Grâce à la transformation digitale de l'usine du futur, la communication et la collaboration avec toutes ces parties intéressées seront facilitées.

De plus, en interne, les salariés devraient constituer une partie intéressée plus active en raison de l'évolution de leurs rôles au sein de l'usine transformée. Leurs besoins et attentes vont évoluer transformant potentiellement ainsi les visions stratégiques, tactiques et opérationnelle de l'usine du futur. Ce seront des éléments constitutifs de la performance globale et de son évaluation.

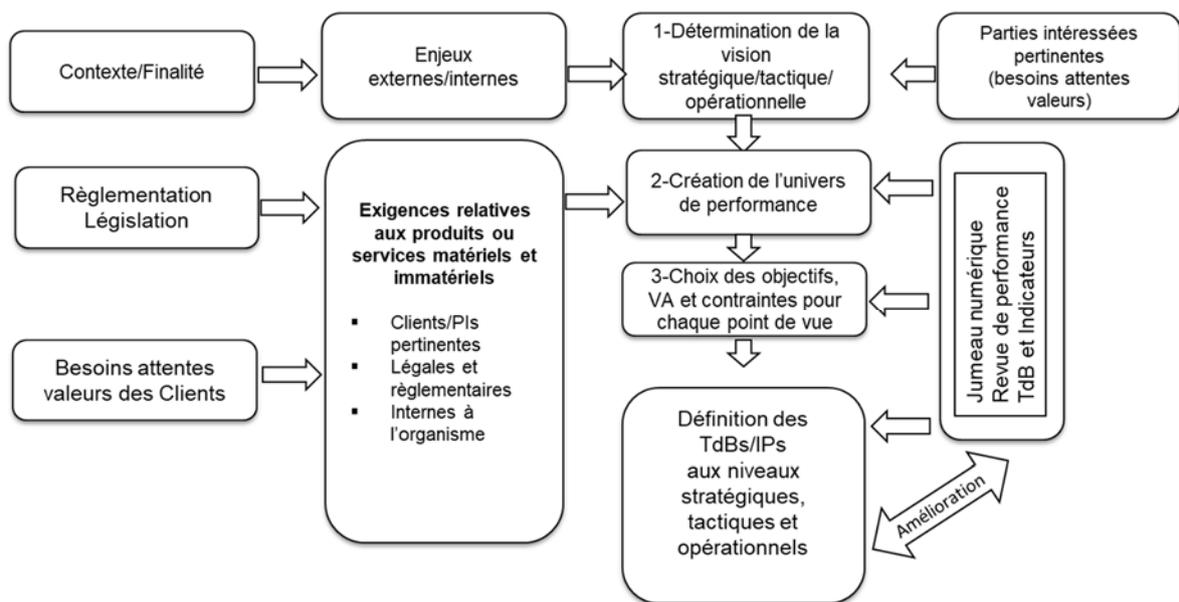


Figure 1 : La démarche de construction des tableaux de bord

3.2 - Création de l'univers de la performance

L'analyse réalisée dans l'étape précédente permettra de prendre en considération les intérêts de l'ensemble des composants de cette sphère d'influence. Quels aspects de la performance est-il important de considérer ?

Nous obtenons alors un « univers de la performance ».

Classiquement l'évaluation de la performance de l'entreprise se fait aux niveaux financier, économique, social, organisationnel, environnemental et sociétal. Pour l'évaluation de l'usine du futur, nous proposons l'ajout de l'aspect technologique décrit dans la partie précédente.

La **performance technologique** de l'usine de futur représente le niveau atteint en matière de nouvelles technologies (digitalisation, data, robots, intelligence artificielle, internet des objets...). La **performance financière** permet de mesurer la rentabilité de l'entreprise et sa capacité de financement des nouveaux projets. La **performance organisationnelle** représente la mesure de la performance de l'entreprise au niveau de la qualité, des délais, de la productivité et de la flexibilité. La **performance économique** permet de mesurer les composantes de la compétitivité, prix et compétitivité hors prix de l'entreprise. La **performance environnementale** représente les résultats mesurables de l'entreprise en relation avec les aspects environnementaux (pollution, déchets, consommation des ressources...). La **performance sociale** représente le niveau atteint par une entreprise d'un point de vue social (salaires, formations, compétences, bien-être, sécurité, autonomie...). En complément, la **performance sociétale** représente le niveau atteint dans les trois domaines

3.3 - Détermination des objectifs, indicateurs, variables d'action et/ou contraintes

Pour chaque niveau décisionnel, pour chaque point de vue de la performance que l'entreprise souhaite évaluer, il faudra déterminer l'objectif à atteindre, les variables à collecter, les indicateurs à calculer, pour ensuite pouvoir, lors de l'exploitation du tableau de bord, analyser les mesures obtenues. Les actions de pilotage de la performance seront

ensuite fixées en fonction des contraintes et suivant les variables d'action définies [Ravelomanantsoa et al., 2018], [Berrah L. et al., 2018].

Les tableaux 1,2 et 3 que nous avons construits permettent de présenter notre proposition de la performance de l'usine du futur selon la déclinaison des trois niveaux décisionnels.

Notre approche est basée sur les principes de l'analyse systémique [Melese, 1991], [Yatchinovsky, 2018]. Cette modélisation de l'entreprise du futur nous donne à la fois une vision globale de l'usine du futur et une décomposition plus détaillée en processus et activités.

Le tableau 1 permet d'illustrer la mise en œuvre de notre méthode pour l'usine du futur au niveau stratégique, dans les domaines de la performance environnementale, sociale, économique, sociétale, technologique, financière et organisationnelle.

Le même raisonnement devra être répété aux niveaux tactique et opérationnel. Le niveau tactique représente les décisions à moyen terme pour un système constitué de processus et de projets (tableau 2). Un processus est une ensemble d'activités corrélés ou en interaction qui utilise des éléments d'entrée pour produire un résultat escompté [ISO9000, 2015].

Le tableau 3 reprend le même raisonnement au niveau opérationnel. Le niveau opérationnel représente les décisions à court terme des systèmes décomposés en activités ou des postes de travail. Une activité est un ensemble de tâches élémentaires réalisées par une personne qui en maîtrise le savoir-faire et produisant un résultat escompté [ISO9000, 2015]. A ce niveau-là, les objectifs seront caractérisés par rapport aux 5 M (Matière, Milieu, Matériel, Main d'œuvre et Méthodes), [Barsalou, 2015].

Ces trois décompositions décisionnelles (Tableaux 1,2,3) vont permettre de construire une liste ouverte d'indicateurs intégrés dans des tableaux de bord pour faciliter l'analyse des situations et anticiper les évolutions souhaitables. Les tableaux de bord ainsi élaborés seront également des instruments de communication pour les différentes parties

intéressées de l'usine de futur. Chacun sera décliné en fonction des objectifs dans le but d'évaluer la performance de l'usine du futur soit dans sa globalité, soit à des niveaux plus détaillés.

Organisation	Points de vue de la performance	Objectifs	Variables d'action	Indicateurs
Usine du futur Approche système <i>Ensemble de processus</i>	Technologique	Intégrer et développer les nouvelles technologies	Innovation Résistance au changement	Digitalisation, Cyber sécurité, Connectivité Cloud computing
	Environnementale	Réduire l'empreinte carbone	Règlementation Pollution/déchets	Niveau de Pollution, Emission de gaz à effet de serre
	Sociale	Construire le bilan social Améliorer la qualité de vie au travail	Droit du travail Bilan social	Effectif, Satisfaction Turnover ...
	Economique	Optimiser le couple Valeur/Coût	Barrières à l'entrée Chaîne de valeur Concurrence	Valeur économique, Compétitivité prix, Compétitivité hors prix
	Sociétale	Mettre en œuvre l'économie circulaire, Contribuer au développement économique et social local	Pollution/Déchets Bilan social Chaîne de valeur	Niveau de Pollution, Type de compétences, Développement des parts de marchés,
	Financière	Garantir la rentabilité financière	Capacité financière, Capacité d'investissement	Cash flow, Chiffre d'affaires Retour sur investissement (ROI)
	Organisationnelle	Maîtriser la structure Assurer la résilience	Apprentissage organisationnel Résistance au changement	Degré de flexibilité Vitesse de mise en œuvre du changement

Tableau 1 : La performance au niveau stratégique (global/long terme/agrégé)

Organisation	Points de vue de la performance	Objectifs	Variables d'action	Indicateurs
Processus et/ou Projet <i>Quels processus ? Décomposition Chaîne de valeur Cartographie des processus ISO 9001-2015</i> <i>Ou</i> <i>Quels projets ? PERT GANNT</i>	Technologique	Planifier et mettre en place les nouvelles technologies	Innovation Ressources Résistance au changement	Robotisation- Cobot Internet des objets Jumeau numérique
	Environnementale	Recycler les produits et les composants	Caractéristiques techniques des ressources et des produits	Consommation d'énergie (kw/h) Quantité de produits recyclés
	Sociale	Gérer la qualité de vie au travail Construire un plan de gestion prévisionnelle des emplois et des compétences (GPEC)	Ressources humaines	Savoir-faire Communauté apprenante Autonomisation des équipes Compétences
	Economique	Améliorer le triptyque coût-qualité-délai	Mode de réponse au marché	Time to Market Satisfaction clients Coût de revient
	Sociétale	Recycler les produits et les composants Construire un plan de gestion prévisionnelle des emplois et des compétences (GPEC) Améliorer le triptyque coût-qualité-délai	Caractéristiques techniques et produits et des ressources	Quantité de produits recyclés Compétences Coût de revient
	Financière	Gérer les investissements et les budgets	Ressources financières Budget	Chiffre d'affaires Résultats Marges
	Organisationnelle	Planifier les flux et ressources	Coût, qualité, délai	Lead time Takt time Cycle Time

Tableau 2 : La performance au niveau tactique (moyen terme/ intermédiaire)

Organisation	Points de vue de la performance	Objectifs	Variables d'action	Indicateurs
Activités <i>Ensemble de tâches ou postes de travail</i>	Technologique	Numériser les postes Améliorer la robustesse	Collecte des données Compétences Résistance au changement	Données traitées-Big data (volume, temps) Fabrication additive Réalité augmentée Impression3D
	Environnementale	Traiter les déchets Traiter les rejets	Caractéristiques techniques des ressources	Quantité de déchets au poste de travail par type Quantité d'eau usée
	Sociale	Assurer la qualité de vie au travail Améliorer l'ergonomie des postes de travail	Droit du travail Absentéisme Législation française	Nombre de jours d'absence, d'arrêts, d'accidents, Taux de fréquence et de gravité Accidents du travail
	Economique	Qualité des produits Respecter les délais de livraison Réduire les Coûts de revient	Satisfaction des exigences clients et réglementaires Fiabilité de la logistique Maîtrise des coûts	Produits non conformes Planning de livraison Coûts
	Sociétale	Traiter des déchets Assurer la qualité de vie au travail Améliorer le triptyque coût-qualité-délai	Type de déchets Satisfaction du personnel Coûts cachés	Quantité de déchets par type Nombre de formations Taux de satisfaction clients Taux de service Taux de productivité
	Financière	Réduire le coût de revient	Type de ressources Type de matières	Coût des matières Coût de transformation des ressources Coût logistique
	Organisationnelle	Piloter le système physique Gérer les flux physiques et d'informations	Cadence de production Capacité (hommes/machines)	Délai d'obtention du produit Taux de rendement synthétique

Tableau 3 : La performance au niveau opérationnel (local/ court terme /détaillé)

Le tableau 4 considère la dimension « technologique » sur les trois niveaux vus précédemment. La détermination des indicateurs est réalisée en utilisant la méthode d'identification des Objectifs/Variables d'action/Indicateurs selon chaque dimension étudiée [Bitton,1990].

Ainsi, l'aspect « innovation » se décline au niveau stratégique via la digitalisation (« La digitalisation représente la transformation d'un objet, outil, processus en un code informatique dans le but d'améliorer la performance et la rentabilité des entreprises »), au niveau tactique via les

nouvelles technologies permettant l'amélioration des flux des processus (robots/cobots, système d'information...) et au niveau opérationnel via les innovations pour faciliter le traitement des données (Big Data, internet, logiciel...).

Par exemple, les indicateurs déterminés permettent de mesurer au niveau stratégique le degré de digitalisation, au niveau tactique, le nombre de robots et de cobots, et au niveau opérationnel, des indicateurs sur le Big Data ou les logiciels...

Points de vue technologique de la performance	Objectifs	Variables d'action	Indicateurs
Stratégique	Intégrer et développer les nouvelles technologies	Innovation/digitalisation	Degré de digitalisation,
Tactique	Planifier et mettre en place les nouvelles technologies	Innovation/flux Innovation/Ressources	Nombre de Robots Nombre de Cobots % d'Informatisation du Système d'information Niveau de développement de l'Internet des objets et des Jumeaux numériques
Opérationnel	Numériser les postes Améliorer la robustesse	Innovation/Collecte des données	Big Data (volume, temps) Réseau Internet Logiciels

Tableau 4 : Déclinaison la performance technologique – exemple de la digitalisation

Le principal intérêt de cette démarche est qu'elle s'appuie sur les transformations technologiques de l'usine du futur pour en tirer tous les bénéfices (rapidité, anticipation, efficacité).

Par contre, elle ne nécessite pas de compétences spécifiques ni de transformations profondes puisque basée sur des connaissances largement répandues dans les entreprises qui ont déjà mis en place une logique risque ou qualité de type ISO.

Ensuite, les tableaux de bord définis correspondent réellement aux attentes, besoins et valeurs des parties prenantes ainsi qu'au contexte interne/externe et à la stratégie. Ainsi nous assurons la réelle intégration de l'usine du futur dans son environnement local et global.

Les différents tableaux de bord définis aux niveaux stratégiques, tactiques et opérationnels seront cohérents et complémentaires. Ainsi nous éviterons de multiplier les données inutiles et les indicateurs qui ne servent à rien. A chaque niveau, on déterminera le tableau de bord optimal dans le contexte du moment, contenant les indicateurs importants, et utiles.

La performance évaluée prendra en compte l'ensemble des performances environnementales, économique, sociétales... suivant les axes prioritaires pour l'entreprise et ses parties intéressées pertinentes.

A chaque niveau décisionnel correspond un tableau de bord adéquat associé aux variables d'action et contraintes connues et disponibles à ce niveau. Ce qui garantit une action possible en cas de résultat non satisfaisant tout en ayant la connaissance des risques et des contraintes.

Et enfin, on voit bien qu'à tous les niveaux de l'entreprise chacun peut agir pour permettre d'améliorer la performance globale de l'usine du futur dans les différents domaines de l'environnement de la technologie de la finance ou du sociétal.

4 CONCLUSION

En conclusion, nous avons mis en évidence la nécessaire évaluation multi-niveaux et multidimensionnelle de la performance de l'usine du futur. En effet, les trois niveaux décisionnels de l'entreprise du futur : stratégique, tactique et opérationnel sont indispensables pour une évaluation cohérente et pertinente.

Dans notre réflexion, la performance de l'usine du futur prend en considération de façon simultanée les aspects technologiques, environnementaux, économiques, sociaux, sociétaux, financiers

Les perspectives de recherche concernent, dans la continuité de cette réflexion, l'étude des relations entre les différents niveaux (stratégiques, tactiques et opérationnels) et les différentes dimensions (technologiques, environnement, social, ...). En effet, il existe des liens et des relations, qu'il sera nécessaire d'étudier (hiérarchisation, causes-effet, corrélation).

Nous proposerons, dans le cadre de nos futurs travaux, l'application de la méthode au sein de plusieurs entreprises de différents secteurs d'activités et de différentes tailles.

En définitive, il n'existe pas de méthode universelle pour évaluer la performance de l'usine du futur. Nous espérons toutefois que notre proposition contribuera à construire une

approche pour construire un système d'évaluation qui intègre les aspects multicritères de la performance de l'usine du futur.

5 REFERENCES

- Andre J.C., (2019), « Industrie 4.0, paradoxes et conflits » *Iste Editions*.
- Barsalou M.A., (2015), « Root Cause Analysis: A Step-By-Step Guide to Using the Right Tool at the Right Time », *CRC Press*.
- Berrah L., Cliville V., Foulloy L., (2018), « Objectifs et performances industriels », *Iste Editions*.
- Bitton M., (1990), « ECOGRAI : méthode de conception et d'implantation de systèmes de mesure de performances pour organisations industrielles », *Thèse, Université Bordeaux*.
- Blanchet M., Varin P. (2016). « Industrie 4.0 nouvelle donne industrielle, nouveau modèle économique ». Gambais (Yvelines): Lignes de repères.
- Camara M. S., Ducq Y., Dupas R. (2013), « A methodology for the evaluation of interoperability improvements in inter-enterprises collaboration based on causal performance measurement models », *International Journal of Computer Integrated Manufacturing*.
- Deming W.E. (1996), « Du nouveau en économie », *Economica*.
- Estampe D. (2015), « Performance de la supply chain et modèles d'évaluation », *Iste Editions*.
- Hamadmad H., (2017), « Définition d'une expression temporelle de la performance des entreprises manufacturières », *Thèse, Université de Grenoble*.
- Iso 9000 (2015) « Systèmes de management de la qualité - Principes essentiels et vocabulaire », *AFNOR*.
- Iso 9001 (2015), « Systèmes de management de la qualité – Exigences », *AFNOR*.
- Iso 31000 (2018), « Management du risque - Lignes directrices », *AFNOR*.
- Ivanov D. et al., (2020) « Researchers perspectives on industry 4.0 : multi-disciplinary analysis and opportunities for operations management », *International journal of production research*.
- Kohler D., Weisz J.D. (2016), « Industrie 4.0, les défis de la transformation numérique du modèle industriel allemand », *Paris, la documentation française*, p.26.
- Julien N., Martin (2018). « L'usine du futur stratégies et déploiement industrie 4.0, de l'IOT aux jumeaux numériques » Malakoff, *Dunod*.
- Lorino P., (2003), « Méthodes et pratiques de la performance », Editions d'organisation.
- Melese J., (1991), « l'analyse modulaire des systèmes, AMS : Une méthode efficace pour appliquer la théorie des systèmes de management », *Editions d'organisation*.
- Moeuf A., (2018), « Identification des risques, opportunités et facteurs critiques de succès de l'industrie 4.0 pour la performance industrielle des PME. », *Thèse, Université de Paris Saclay*.
- Perez R. (1989), « Contraintes stratégiques et logiques d'action des groupes alimentaires », *Economies et Sociétés, série AG, n°20, 23, n°7, pp.9-26*.
- Ravelomanantsoa M.S., Ducq Y., Vallespir B., (2019), « A state of the art and comparison of approaches

for performance measurement systems definition and design “, *International Journal of production research*. Vol. 57 - n°15-16 – pp 5026-5046

Senechal O. (2004), « Proposition d'une méthodologie pour l'aide à l'estimation des performances physico-économiques des systèmes de production dans une approche concourante », *Thèse, Université de Valenciennes*.

Yatchinovsky A. (2018), « L'approche systémique. Pour gérer l'incertitude et la complexité », *ESF*.

Performance industrielle : une évolution intégrant l'éthique dans le contexte de l'Industrie 4.0

LAMIA BERRAH¹, VINCENT CLIVILLE¹, DAMIEN TRENTESAUX², CLAUDE CHAPEL³

¹ LISTIC, Université Savoie Mont Blanc
74940 Annecy France

lamia.berrah@univ-smb.fr, vincent.cliville@univ-smb.fr

² LAMIH UMR CNRS 8201, Université Polytechnique Hauts de France,
59313 Valenciennes France

damien.trentesaux@uphf.fr

³ NTN SNR, 1 rue des Usines,
74010 Annecy France

clau.de.chapel@ntn-snr.fr

Résumé - Cet article traite de la performance industrielle et de l'enrichissement de son triangle (efficacité, efficacité, effectivité) en vue d'y intégrer la dimension éthique. L'apport de l'éthique à la performance est discuté dans le contexte actuel de l'Industrie 4.0 qui, de par la numérisation des activités de production, le développement de Systèmes Autonomes voire Intelligents et la profusion de données qui la caractérisent, offre bon nombre de situations dans lesquelles le modèle de performance traditionnel trouve ses limites pour la garantie de la pérennité des systèmes. Ces limites peuvent être liées aussi bien aux objectifs assignés qu'aux moyens utilisés pour les atteindre. Après un rappel des spécificités de l'Industrie 4.0, une analyse de l'évolution de la notion de performance et de ses outils de déploiement, depuis la période taylorienne jusqu'à aujourd'hui, est présentée. Le triangle (efficacité, efficacité, effectivité) y est progressivement introduit et ses limites sont montrées. Une proposition intégrant, dans un tétraèdre, l'éthique à ce triangle est décrite et illustrée sur les conséquences de la mise en place d'un MES (Manufacturing Execution System) chez un équipementier aéronautique, en termes de traçabilité des produits et d'autonomie des opérateurs. Une discussion et des perspectives à la proposition concluent cette étude.

Abstract - This paper deals with industrial performance and the enrichment of its triangle (efficiency, effectiveness, relevance) with a view to integrating the ethical dimension. The contribution of ethics to performance is discussed in the current context of Industry 4.0 which, due to the digitisation of production activities, the development of Autonomous or even Intelligent Systems and the profusion of data that characterise it, offers situations in which the traditional performance model finds its limits in terms of guaranteeing the durability of systems. These limits can be related both to the assigned objectives and to the way in which the means are used to achieve them. After a reminder of the specificities of Industry 4.0, an analysis of the evolution of the performance notion and its deployment tools, from the Taylorian period to today, is presented. The triangle (efficiency, effectiveness, relevance) is gradually introduced and its limits are shown. A proposal integrating, in a tetrahedron, the ethics to this triangle is described. Such a proposition is illustrated regarding the consequences of the implementation of a MES (Manufacturing Execution System) for an aeronautical supplier, in terms of product traceability and operator autonomy. Discussion and prospects to the proposal conclude this study.

Mots clés - Performance industrielle, Industrie 4.0, Ethique.

Keywords - Industrial performance, Industry 4.0, Ethics.

1 INTRODUCTION

Depuis les trente glorieuses, la performance des entreprises s'est implicitement faite synonyme de rentabilité [Taylor, 1911] [Johnson, 1975]. Pour répondre à cette vision, la performance industrielle, *i.e.* celle des systèmes de production, a commencé par s'exprimer en termes de Coût : rendement des équipements et productivité de la MOD (Main-d'Oeuvre Directe). Le modèle de la performance s'est par la suite étendu pour prendre en compte les évolutions du contexte, en particulier les notions de satisfaction client et de compétitivité. L'expression de la performance a alors intégré respectivement la Qualité : conformité des produits et process et le Délai : de mise à disposition. L'expression de la performance était ainsi fondée sur le déploiement d'un objectif stratégique de rentabilité en objectifs tactiques et opérationnels essentiellement de Coût, Qualité et Délai. Les notions d'Indicateurs de Performance et de

Systèmes d'Indicateurs de Performance ont été mises en avant dès lors que les seuls calculs financiers ne suffisaient plus à l'expression de la performance [Kaplan et Norton, 1992]. Définis pour renseigner de l'atteinte des différents objectifs assignés et par-delà même de l'efficacité des actions associées, ils deviennent les outils d'aide au pilotage des systèmes de production [Popova et Sharpanskykh, 2010].

En ayant identifié la pérennité de l'entreprise à une rentabilité, courttermiste, qu'il suffisait d'atteindre en maximisant les résultats atteints au regard des critères mentionnés, le modèle de performance était présumé valide. Progressivement, le bien-fondé de ces pratiques a commencé néanmoins à être interrogé avec des évolutions telles que la mondialisation accrue et ses besoins de coordination, la limite des ressources naturelles et ses besoins de réduction de consommation, etc. La recherche permanente de maximisation des critères usuels a de ce fait eu de plus en plus besoin d'être bornée par des règles et des normes

prenant en compte les nouvelles hypothèses de production. En particulier, par sa révolution technologique, l'Industrie 4.0 amène à son tour des changements renforçant le besoin de questionner le modèle de performance, en matière d'objectifs à atteindre et d'actions à mener. La « transformation » digitale touche, en effet, par l'intelligence et la profusion de données proposées, les différentes dimensions reliées à l'activité industrielle, *i.e.* aussi bien le fonctionnement (technologique) des systèmes que leur pilotage (organisation et management) [Fettig et al., 2018]. Les opportunités de cette transformation restent encore aujourd'hui mal connues et requièrent ainsi une vigilance et une analyse accrues.

Traitant de l'évolution du modèle de performance industrielle, cet article se place dans le contexte de l'Industrie 4.0. En ce sens, parmi les concepts relevant de questions de bien-fondé autour de la notion de performance vient le concept d'effectivité [Le Moigne, 1990]. Troisième facette de la performance, aux côtés de l'efficacité qui traite du rendement des équipements et de l'efficacéité qui concerne l'atteinte des objectifs assignés, l'effectivité porte sur l'adéquation, pour le système, entre les objectifs assignés et les moyens alloués. Jusqu'à l'Industrie 3.0, l'effectivité s'est traduite par une action sur les moyens, bien maîtrisés, pour maximiser l'atteinte des objectifs et aura été fondue dans la double recherche de l'efficacité et l'efficacéité.

Parmi les pistes pour compléter le triplet (efficacité, efficacéité, effectivité), nous choisissons celle globale de l'éthique, dans sa recherche, par essence, d'« *une vie bonne, avec et pour autrui, dans des institutions justes* » [Ricoeur, 1990]. Vue aujourd'hui comme faisant partie des politiques de RSE (Responsabilité Sociétale des Entreprises), l'éthique représente le paradigme permettant de borner tout objectif et action œuvrant pour la pérennité de l'entreprise, *i.e.* sa rentabilité à la fois à court et long termes. Afin d'illustrer cette idée, nous reviendrons dans un premier temps sur les spécificités de l'Industrie 4.0 ainsi que les conséquences de son implantation. Nous analyserons alors le modèle de la performance puis préciserons la problématique. Nous montrerons en quoi le concept d'éthique et ses paradigmes font nécessairement évoluer le concept de performance. Le principe d'un déploiement de la performance enrichi de cette dimension est proposé puis illustré par un cas d'étude apporté par un équipementier aéronautique. Quelques remarques et perspectives concluront enfin cette étude.

2 L'INDUSTRIE 4.0

De par ses piliers technologiques (Figure 1), l'Industrie 4.0 transforme les systèmes de production qui passent d'Automatisés à Autonomes/Intelligents (Systèmes Cyber-Physiques, Cobots, Robots). Vue comme « *an umbrella term used to describe a group of connected technological advances that provide a foundation for increased digitization of the business* » [Burritt et Christ, 2016], elle représente un changement majeur dans la manière de concevoir et de réaliser l'activité industrielle. Au-delà, il est admis que le concept n'est pas maîtrisé en matière de portée des opportunités qu'il présente [Ivanov et al., 2019]. L'Industrie 4.0 se caractérise en effet par des aspects, synthétisés en les points suivants, qui la rendent source de restriction, complexité et risques [Ghobakhloo, 2018] [Trentesaux et Caillaud, 2020] :

- une digitalisation de plus en plus importante des différentes activités liées à la production ;
- le développement de systèmes autonomes/intelligents, à même d'exécuter ces activités ;
- l'augmentation des interactions entre l'Homme et les Intelligences Artificielles.



Figure 1. Les piliers de l'Industrie 4.0 [Saturno et al., 2017]

L'aspect restrictif de l'Industrie 4.0 concerne les dimensions environnementales et sociétales dans la limitation des ressources utilisées et des impacts sociétaux des technologies utilisées. Quant à la complexité, elle est induite par :

- le grand nombre d'acteurs impliqués à la fois dans la conception des systèmes et l'usage de ces derniers ;
- l'impossibilité d'avoir une vue exhaustive des comportements, en cas de changements, des systèmes (autonomes/intelligents).

Les risques soulignés dans la mise en œuvre de l'Industrie 4.0 concernent les retombées sur les Hommes qui y sont impliqués. Ils sont essentiellement liés [Trentesaux et Caillaud, 2020] :

- à la fiabilité limitée des interactions des Hommes avec les systèmes intelligents ;
- au sacrifice potentiel de leur bien-être, voire leur remplacement, à des fins d'accroissement de la productivité ;
- à la perte de compétence en conséquence de ce remplacement.

Concernant l'expression de la performance, l'Industrie 4.0 offre plusieurs opportunités [Kamble et al., 2018]. Vient celle d'une meilleure atteinte des objectifs, de par la précision des procédés mis en œuvre. D'autre part, de par : une traçabilité accrue, la mise à disposition des informations en temps réel et à tout moment, la réduction des erreurs de *reporting* du fait de leur automatisation, elle permet un pilotage encore plus réactif. L'Industrie 4.0 vient renforcer l'efficacité et l'efficacéité du système de production.

En résumé, nous pouvons considérer que les systèmes de production de l'Industrie 4.0 sont des systèmes par essence digitaux et complexes et dont les comportements et réactions aux changements sont parfois imprévisibles et potentiellement risqués, même si des recherches sont menées pour assurer la correcte intégration des Hommes dans ces systèmes [Pacaux-Lemoine et Trentesaux, 2019]. Maintenir le cap sur une amélioration permanente du triptyque C-Q-D (Coût-Qualité-Délai) pour une recherche d'efficacité-efficacéité-effectivité telle que pratiquée dans l'Industrie 3.0 peut en effet questionner. Afin de répondre à un tel questionnement, rappelons dans un premier temps les définitions et évolutions essentielles de la notion de performance industrielle et de ses postulats.

3 D'UNE VISION HISTORIQUE A UNE VISION MODERNE DE LA PERFORMANCE

3.1 La notion de performance

Dans son sens commun, la performance « *désigne un résultat exceptionnel, hors du commun, optimal. Elle relève d'attentes que l'on peut traduire en objectifs* » [Boisvert, 2019]. Véhiculant la notion d'optimalité, la performance a été associée, au niveau des entreprises, au critère de rentabilité et aux objectifs qui contribuent à cette dernière (Tableau 1). La rentabilité ayant été à son tour identifiée à la finalité de

l'entreprise, elle s'en est devenue assimilée à sa pérennité. Aussi, performance, rentabilité, pérennité sont-elles devenues plus ou moins judicieusement synonymes en milieu industriel. Héritant de cette vision, la performance industrielle a été en tout premier lieu synonyme de productivité de la MOD (Main-d'Oeuvre Directe), reposant sur le critère de Coût. Cette performance s'est alors déployée à la fois en interne et en externe, étendant d'une part la productivité à l'ensemble des facteurs de production, et intégrant d'autre part les notions de compétitivité et de satisfaction client, et ce, par la prise en compte de critères de Qualité et de Délai.

Tableau 1 Déploiement classique de la performance [Jacot et Micaelli, 1996]

niveau de décision	niveau de performance	objectifs et critères usuels	nouveaux objectifs et critères
Métropolitaine finalité de l'entreprise	Sociétal	Performance interne de l'entreprise	Performance externe de l'entreprise
Stratégique objectifs pour le système de production	Financier	Rentabilité des capitaux	Pérennité de l'entreprise
Tactique fonctionnalités pour le système de production	Marchand/Commercial	Compétitivité prix	Compétitivité hors prix approche par processus
Opérationnel solutions techniques et organisationnelles	Physique	Productivité partielle du travail	Productivité globale

Plus précisément, si la période des trente glorieuses avait associé la performance de ces systèmes au concept d'efficacité et avait installé les objectifs de productivité rattachés au Coût, celle qui lui a succédé aura étendu la notion de performance au couple (efficacité, efficience). L'efficacité prend en compte la manière dont tout objectif, au-delà de celui de productivité, est atteint. La dimension multicritère de la performance avait alors été mise en avant, en adjoignant au critère Coût la Qualité puis le Délai. On parlait alors d'efficacité, *i.e.* «atteindre les objectifs de qualité et de délai», sous une contrainte d'efficience, *i.e.* «optimiser le rendement». Aujourd'hui encore, un système de production est conventionnellement jugé performant s'il satisfait au triptyque C-Q-D [Ducq et al., 2001]. Fournir aux moindres Coûts des produits de Qualité et ce, dans les meilleurs Délais reste la stratégie communément adoptée. Le contrôle de gestion, initialement conçu pour les évaluations financières et les vérifications *a posteriori* [Anthony, 1965], avait à son tour été enrichi par des approches de pilotage *a priori*, pour des évaluations technico-économiques, rattachées à des critères autres que le Coût [Brimson, 1991]. On parlait alors de *contrôle* à long terme, par le contrôle de gestion et de *pilotage réactif* à court et moyen termes, par la mise en place de Systèmes d'Indicateurs de Performance et de tableaux de bord opérationnels [Neely et al., 1995] [ISO 22400, 2014]. Naturellement, au fur et à mesure des évolutions de l'offre technologique, des stratégies concurrentielles et des exigences clients, les systèmes de production se sont complexifiés. Les facteurs de performance se sont étendus et les interactions entre eux ont augmenté. Des règles permettant une réaction aux événements survenus ont été intégrées aux règles de pilotage préétablies, laissant place à un besoin d'aide à la décision pour la détermination des solutions adéquates dans un contexte assumé incertain. De nouvelles stratégies ont été mises en avant et le triptyque C-Q-D s'est enrichi, en intégrant de plus en plus de sous-critères. En l'occurrence, la Qualité n'a plus seulement concerné le produit fini, le Coût la production, et le Délai, la mise à disposition. Les Systèmes d'Indicateurs de Performance ont tenté dans ce sens d'embrasser à la fois la complexité des nouveaux systèmes de production et celle des interactions entre

les différents critères, introduisant de nouveaux critères internes, tels que la réactivité, la flexibilité ou la fiabilité [Ghalayini et al., 1997]. Des considérations telles que celles d'appartenir à une chaîne logistique [Mishra et al., 2018], de répondre à des objectifs de Développement Durable [Gasparatos et al., 2009] ou plus récemment de s'inscrire dans la logique de digitalisation actuelle [Shah et al., 2020] [Tobon Valencia et al., 2020] sont des illustrations de telles évolutions. Comme évoqué précédemment, les changements observés dans ce dernier cas sont particulièrement majeurs et parfois non maîtrisés.

Au-delà de cette recherche d'adaptation des cadres d'expression de la performance et de pilotage de systèmes de production aux évolutions du contexte, il est légitime d'avoir une autre forme d'analyse et de réfléchir au bien-fondé de ces cadres, tant dans leur définition que dans leurs évolutions.

3.2 Efficience, Efficacité et Effectivité

Comme la performance s'est déployée en fonction de multiples critères, elle s'est fondée sur le postulat de devoir vérifier trois conditions afin de pouvoir continuer à garantir la pérennité de l'entreprise. Ces conditions sont respectivement [Gilbert et al., 1980] [Neely et al., 1996] (Figure 2) :

- l'efficience du système élargi aux activités périphériques à la production ;
- l'efficacité des processus mis en œuvre pour l'atteinte des objectifs ;
- l'effectivité des objectifs au regard des moyens disponibles.

Bon nombre de réflexions ont été menées autour des notions d'efficacité et d'efficience, confondant parfois l'une et l'autre, au vu de leur anglicisation (*effectiveness* ou *efficiency* pour efficacité) [Kanigel, 2005], et de leur rattachement toutes deux à l'atteinte d'objectifs associés au système de production. Pour sa part l'effectivité connaît une définition unanime de pertinence (autre synonyme trouvé dans la littérature) de ces objectifs [Wisner et Fawcett, 1991], avec une nuance toutefois dans la portée de ces objectifs : opérationnels [Bescos 1995], *i.e.* locaux au système considéré ou stratégiques, *i.e.* globaux à l'ensemble de l'entreprise [Le Moigne 1990].

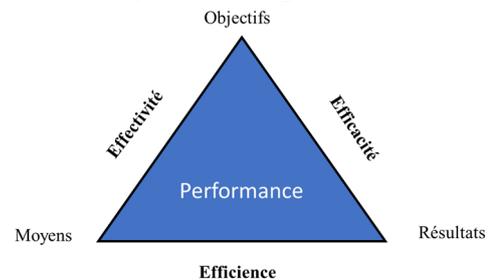


Figure 2. Le triangle de la performance 3.0 [Gilbert et al., 1980]

La performance se définissait ainsi par : de l'efficacité dans l'atteinte des objectifs, dans une utilisation efficiente des moyens disponibles, sous réserve d'effectivité de ces objectifs pour le système en matière de moyens utilisés. Comme évoqué précédemment, si, de par l'Histoire, la performance industrielle a intégré les facettes d'efficacité et d'efficience, la prise en compte de la facette d'effectivité s'est essentiellement inscrite dans une vision d'adaptation des moyens aux objectifs. En effet, préoccupation stratégique, l'effectivité est allée dans le sens d'une maximisation de l'atteinte des objectifs, dans la mesure où ces derniers découlaient de l'unique objectif de rentabilité et que celle-ci garantissait la pérennité de l'entreprise. Les objectifs n'étaient par conséquent pas reconsidérés et les pratiques n'étaient pas appelées à être modifiées, si ce n'est pour

une recherche d'augmentation de la rentabilité. De ce fait, la maximisation de l'efficacité ainsi que celle de l'efficience ont fait l'objet de bon nombre de développements, tant au niveau des modèles de pilotage que des modèles d'évaluation de la performance [Globerson, 1985] [Lorino, 1996] [Nudurupati et al., 2011]. Les plans d'action, inscrits dans une dynamique d'amélioration continue, étaient centrés sur cette maximisation [Imai, 1992]. C'est le sens du *Lean Manufacturing* qui, s'inspirant des principes du *Toyota Production System* [Womack et al., 1990] [Ono, 1988] peut être décrit par le passage d'une logique de rentabilité financière à court terme (que l'on peut décrire sous la forme «*Maximum gain with minimal capital investment*») à une logique d'optimisation de l'efficience à long terme (que l'on peut décrire sous la forme «*Maximum added value with a minimal use of all resources*»). Une performance à tout instant, conformément à ce modèle préétabli, est ainsi, à son tour, implicitement devenue synonyme d'une performance continue. Dès lors que cette équation ne s'est plus avérée, des remises en cause ont commencé à être envisagées et la question de l'effectivité, du point de vue de la remise en cause des objectifs a alors commencé à se poser.

3.3 Prise en compte de l'effectivité : une condition suffisante ?

Le besoin de parer aux urgences, changements, opportunités et incertitudes de l'environnement a ainsi parfois induit une sorte d'identification de la notion de réactivité à des réponses immédiates dont les retombées ne peuvent toujours être prises en compte. Ce constat s'est exacerbé dans le cas du fonctionnement non systématiquement prévisible des systèmes de production de l'Industrie 4.0. L'amélioration à court terme de la performance peut en effet dans de tels cas ne pas garantir la performance, globale et à long terme. Inscrire la performance dès lors sous une contrainte d'effectivité qui amène *stricto sensu* à veiller à ce que les moyens du système aillent dans le sens des objectifs peut ne fournir que des garanties partielles. Cette condition, établie à l'ère de l'Industrie 3.0, ne sera en effet suffisante que si en premier le comportement des moyens est connu et maîtrisé, même si cette question de ne considérer les objectifs qu'en fonction des moyens avait déjà été soulevée [Le Moigne, 1990]. De plus, telle qu'elle a été établie, l'effectivité avait pour objet de continuer à véhiculer la stratégie et les objectifs globaux de l'entreprise. Ne pouvaient être remise en cause que la tactique pour les atteindre, soit leur déploiement en objectifs tactico-opérationnels. Cette condition trouve aujourd'hui ses limites au vu notamment des restrictions liées aux problématiques de limitation des ressources et de la logique d'associer de nouveaux objectifs à de nouveaux systèmes.

La pérennité de l'entreprise ne pouvant donc être restreinte à sa dimension financière, il est important aujourd'hui d'étendre la mise en œuvre de cette notion d'effectivité à des dispositifs qui permettent cette pérennité. La question de l'effectivité doit en effet se poser en termes conjoints d'objectifs et de moyens et la performance doit aller au-delà de la rentabilité. L'effectivité peut alors s'appuyer sur des démarches permettant une approche globale allant au-delà d'un seul critère, quel que soit son importance, et tenant compte des changements possibles des moyens alloués pour satisfaire à ce critère. Certaines approches prônant cette vision ont d'ores et déjà été mises en œuvre, en l'occurrence les démarches de Développement Durable [Sikdar, 2003], les démarches de RSE [ISO 26000, 2008], d'économie circulaire [Rosa et al., 2019] ou de gouvernance globale [GRI, 2020] Les critères considérés dans ces approches viennent rectifier la focalisation sur le critère financier en lui associant les critères environnementaux et de plus en plus les critères

sociétaux. En revanche, la question de la prise en compte du comportement des systèmes reste posée.

3.4 Vers la considération de l'éthique en performance

Dans cette optique, la définition de la performance doit évoluer afin de prendre en compte la pérennité du système, en particulier quand celui-ci intègre des entités autonomes aux comportements non tous prévisibles. Cette démarche prônant une vision juste et globale peut trouver écho dans la notion d'éthique. De par sa définition et son intégration dans les démarches RSE, le cadre que le concept d'éthique propose et ses applications dans des domaines proches de celui du secteur industriel, semble répondre au besoin identifié. Le triangle (efficience, efficacité, effectivité) décrit Figure 2 s'enrichit alors d'une dimension supplémentaire, inhérente à l'éthique. Trois conditions de conformité, des notions respectives d'objectifs, moyens et résultats viennent compléter les conditions établies sur leurs associations, deux à deux. La figure 3 représente cette vision complétée de la performance dont la définition évoluerait dans ce sens et deviendrait :

L'efficacité dans l'atteinte d'objectifs éthiques, dans une utilisation efficiente et éthique des moyens disponibles, sous réserve de pertinence de ces objectifs et du système en matière de moyens utilisés et d'éthique des résultats obtenus.

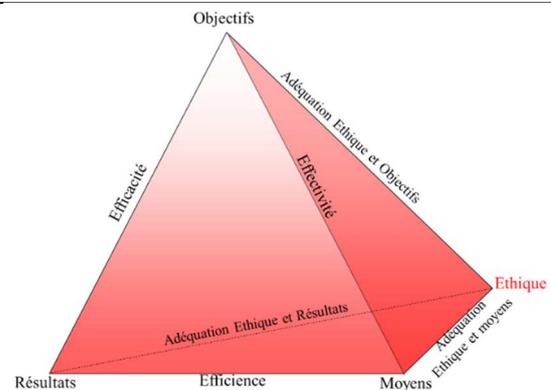


Figure 3. Le tétraèdre de la performance

Au-delà de porter sur les objectifs, les moyens et les résultats, l'éthique porte sur les relations entre eux, au travers des trois notions usuelles de la performance (efficacité, efficience et effectivité). Ainsi, les facettes décrivant les différentes relations entre l'éthique et ces notions sont représentées par les différentes surfaces du tétraèdre. Pour le triplet Objectifs-Moyens-Ethique par exemple, qui traite de la relation entre l'effectivité et l'éthique, l'idée est de fixer des objectifs atteignables sachant les moyens disponibles et aussi une utilisation éthique de ces moyens. Des conditions similaires se posent pour les trois autres surfaces associées aux trois autres triplets. Le tableau 2 contient des exemples de questionnements et illustrations selon la relation étudiée entre l'éthique et les différents concepts historiques de la performance.

Tableau 2. Tétraèdre de la performance : exemples de questionnement

Adéquation entre éthique et...	Exemples de questionnement	Illustrations
Objectifs (axe OE)	Les objectifs assignés sont-ils éthiques ?	Imposition d'un objectif ayant un impact social négatif
Résultats (axe RE)	Les résultats mesurés ou obtenus le sont-ils de manière éthique ?	Utilisation de données personnelles des opérateurs 4.0
Moyens (axe ME)	Les moyens sont-ils utilisés ou définis de manière éthique ?	Morceler les responsabilités pour limiter les salaires

Efficacité (plan ORE)	La mise en relation entre les objectifs et les résultats est-elle faite de manière éthique ?	Triche, cyber-espionnage ou mensonge sur l'atteinte des objectifs
Efficiéce (plan MRE)	La mise en relation entre les moyens et les résultats est-elle faite de manière éthique ?	Remplacement systématique de l'Homme par un robot autonome
Effectivité (plan MOE)	La mise en relation entre les moyens et les objectifs est-elle faite de manière éthique ?	Imposition volontaire d'un objectif inatteignable pour pousser à la faute

Si l'application de ce modèle se veut indépendante de tout contexte, son illustration se fera dans le contexte de l'Industrie 4.0. C'est dans ce sens que la section suivante décrit plus en détail les relations entre la notion d'éthique et celle de la performance dans le contexte de l'Industrie 4.0.

4 PERFORMANCE ETHIQUE ET DANS L'INDUSTRIE 4.0

4.1 Enjeux éthiques de l'Industrie 4.0

Portée par des disciplines telles que la philosophie, le droit, les sciences humaines et sociales de manière plus générale, l'éthique a suscité bon nombre de réflexions tant dans la recherche d'une définition la plus juste que dans les moyens de la mettre en œuvre. Globalement, le comportement d'une entité est éthique du point de vue d'un système, d'une société, s'il véhicule équité et morale tout en s'alignant sur les valeurs et la culture du système, de la société, considérés [Morahan, 2015]. Plusieurs paradigmes sont développés dans le cadre de la recherche d'un comportement éthique. Si l'on s'intéresse aux paradigmes facilement appropriables par le chercheur en ingénierie technique, il est alors possible d'identifier [Trentesaux et Caillaud, 2020] :

- la *déontologie* qui se conçoit au travers du respect d'un certain nombre de règles immuables, que l'on applique dans n'importe quelle situation ;
- le *conséquentialisme* qui se conçoit au travers de l'étude des conséquences d'actions possibles sur des valeurs éthiques.

Dans leur analyse concernant l'éthique dans le cadre de l'Industrie 4.0, [Pacaux-Lemoine et Trentesaux, 2019] mettent en avant l'importance de réfléchir à l'éthique rattachée au fonctionnement de l'ensemble des entités intelligentes qui y sont impliquées, qu'elles soient de nature humaine ou digitale (par exemple, Intelligence Artificielle), si elles présentent un certain degré d'autonomie décisionnelle. La spécificité du concept Industrie 4.0 par rapport à celui d'Industrie 3.0 est ainsi relative à la difficulté, *via* une agrégation de technologies digitales variées et à fort potentiel d'utilisation, de comprendre, prédire et garantir le comportement de ces entités en particulier sur le volet éthique. Pour l'équipe, il s'agit de considérer le comportement de ces entités à la fois lorsqu'elles fonctionnent en autonomie et lorsqu'elles fonctionnent en interaction avec les Hommes. Les auteurs décrivent les enjeux liés à ces technologies et prônent par exemple le recours au paradigme conséquentialiste dans le cas où une forte incertitude sur les comportements de ces dernières est présente, ou l'impact de leur fonctionnement sur l'environnement est mal connu. Typiquement, dans ce dernier cas, aucun objectif ne peut être explicitement associé aux entités considérées, ces dernières échapperont donc à toute mesure et toute action pour y pallier, si ce n'est éventuellement en termes de risque à minimiser. Dans de tels cas, les trois postulats inhérents à la performance trouvent leurs limites, comme illustré par le cas d'étude décrit dans la partie suivante.

4.2 Transition vers l'Industrie 4.0 et enjeux éthiques : étude de cas industriel

Filiale de la NTN Corporation, la société NTN-SNR Roulements est l'un des leaders mondiaux en tant que concepteur, développeur et fabricant de roulements. Réalisant un Chiffre d'Affaires de 6 Milliards d'Euros en 2019, l'entreprise est présente sur les marchés de l'automobile, du ferroviaire, des énergies renouvelables et de l'aéronautique. NTN-SNR Roulements emploie environ 6 000 personnes principalement en Europe et compte 13 sites de production, dont 7 en France. En particulier, le site spécialisé dans l'aéronautique bénéficie depuis 2015 d'investissements importants pour digitaliser la production, conformément aux principes du 4.0. Dans ce cadre, le projet « *Digital Manufacturing* » a été mis en œuvre. Il concerne le système de production dédié aux roulements du nouveau moteur LEAP destiné aux nouvelles générations de moyen-courriers tels que l'Airbus a320neo et le Boeing 737Max. Ce projet a donné la priorité à la mise en place d'une solution MES (*Manufacturing Execution System*) pour d'une part, répondre rapidement aux attentes des clients en matière de traçabilité produit-process, et d'autre part, mettre en place les tableaux de bord opérationnels permettant le pilotage des postes de production [Clivillé et al., 2019]. Il s'agit de mettre à disposition les données de production qui, combinées aux données commerciales, financières et humaines, permettront de proposer de nouvelles utilisations (Réalité augmentée, *Simulation*, *Cloud Computing*, Intelligence Artificielle) dans un cadre sécurisé (*Cyber Security*).

Les premiers résultats du projet ont été encourageants. L'entreprise a constaté que :

- la traçabilité s'était nettement améliorée et permettait de désormais répondre aux exigences des équipementiers ;
- les tableaux de bord étaient déployés sur l'ensemble des postes de production.

Le projet se poursuit actuellement par le déploiement successif des différents piliers de l'Industrie 4.0 (cf. Figure 1) ; ce qui vient transformer en profondeur le système et son usage. Par l'introduction des technologies proposées (*Big Data*, *IoT*, *System integration*, *Cyber-Physical Systems*), l'entreprise a pour but de rendre le système de production intelligent, afin de mieux répondre aux besoins du client, analyser et améliorer la performance, développer des approvisionnements parcimonieux et durables, rendre l'entreprise plus agile, etc. Cependant, au fur et à mesure de ces transformations, l'entreprise a pris conscience de l'importance de la dimension humaine, paradoxalement apparue comme essentielle pour maîtriser le déploiement et l'utilisation des technologies du 4.0. Ainsi, l'abandon de tâches à faible valeur ajoutée comme le *reporting* des données de production et la montée en expertise en termes de qualité et de gestion de flux au service d'un pilotage plus autonome bouleversent le rôle traditionnel des opérateurs de production. Il s'agit désormais de maîtriser les outils numériques et d'être partenaires de leur développement pour « customiser » le MES en fonction des besoins de chaque poste. La structure de pilotage est remise en cause par cette autonomie permise par la mise à disposition des Indicateurs de Performance directement actualisés par le système d'information sur les postes de travail, sans l'intervention d'échelons intermédiaires comme les chefs d'équipe.

Les questions posées par cette transition sont multiples et si les opportunités qu'elle offre sont de mieux en mieux connues, les risques associés se dévoilent progressivement. Dans ce sens, l'entreprise ressent le besoin de compléter sa vision traditionnelle de la performance (conforme à la Figure 2), où les

objectifs sont régulièrement reconsidérés à la hausse, les résultats sont analysés dans ce sens, et les moyens sont alloués en fonction des résultats et des nouvelles valeurs à atteindre. A ce titre, l'entreprise se définit une nouvelle nature d'objectifs qu'elle formule au travers d'un certain nombre d'actions :

- repenser les missions affectées aux opérateurs, et définir les nouvelles fiches de description du poste ;
- définir les nouvelles responsabilités au sein d'une structure de pilotage moins hiérarchisée et plus coordonnée ;
- définir un protocole d'accompagnement (*coaching*) des opérateurs en complément des écrans de supervision associés à leur travail ;
- reconsidérer son modèle de pilotage en adoptant une philosophie proche de celle du Tableau de Bord Prospectif (BSC) [Kaplan et Norton, 1992], où chaque centre de décision de la structure de pilotage dispose d'informations établies de façon cohérente sur tous les niveaux de décision ;
- établir sur cette base pour chaque poste, un modèle de simulation en temps réel des impacts locaux et globaux des décisions opérationnelles.

Pour mener à bien ce projet, l'entreprise connaît bien sûr les lois et règlements à respecter telles le Règlement Général sur la Protection des Données (RGPD) [RGDP 2020]. Mais plutôt que de les subir comme des contraintes, elle veut miser sur le développement humain en réinterprétant la définition de Ricoeur par « *une activité professionnelle heureuse, avec et pour autrui, dans une entreprise juste* ». Elle souhaite définir ses objectifs, évaluer ses résultats et allouer ses moyens dans cette vision éthique de la performance.

Tentons maintenant d'éclairer cette volonté de l'entreprise en questionnant le tétraèdre proposé en section 3.4 sur la considération de l'éthique dans la définition de la performance de l'entreprise au travers de deux exemples (Tableau 3). Considérons le triplet Objectifs-Moyens-Ethique. Pour atteindre l'objectif lié à la « Traçabilité Produits-Process », l'entreprise ajoute un nouveau composant de son système d'information dont elle sait qu'il autorisera également la traçabilité des opérateurs et plus généralement la disponibilité des personnels (logistique interne, qualité, maintenance...) impliqués. Pour éviter un détournement des données recueillies et stockées, l'entreprise doit encadrer les possibilités d'analyse et les restitutions qui en sont faites hors du pilotage de la production, en stipulant par exemple que la personne concernée devra être informée et consultée pour les utilisations potentielles en ce qui concerne la gestion des ressources humaines.

Considérons cette fois le triplet Objectifs-Résultats-Ethique en retenant l'objectif lié à la « Mise en place des tableaux de bord opérationnels ». Les changements induits par la pratique du MES dans la structure de pilotage, augmentation des responsabilités du niveau opérationnel, simplification de la hiérarchie, renforcement de l'autonomie mais aussi du risque d'isolement, sont importants. Sont-ils justes pour les opérateurs qui voient leurs compétences s'élargir en même temps qu'augmentent leurs responsabilités ? Sont-ils justes pour les chefs d'équipe qui voient une part importante de leurs missions devenir caduques et qui doivent se réinventer un nouveau rôle à la frontière entre hiérarchie et support ? Consciente de l'importance de la considération de ces aspects pour sa pérennité et des limites actuelles de ses compétences dans ce domaine, l'entreprise envisage la création d'un groupe de réflexion sur le sujet, afin de systématiser la prise en compte de considérations éthiques dans son modèle de performance.

Tableau 3. Quelques questionnements (étude de cas industriel)

Adéquation entre éthique et...	Questionnement de l'industriel
Objectifs (axe OE)	Fixer un objectif de réactivité difficilement atteignable par un opérateur est-il éthique ?
Résultats (axe RE)	Est-il bon d'atteindre une traçabilité de 100% sur les produits et de 90% sur les processus quand il existe un risque de voir cette traçabilité mesurée pour l'Homme ?
Efficience (plan MRE)	Comment empêcher, voire interdire, que les données permettant de tracer produits et processus ne servent à tracer les opérateurs et plus généralement mesurer la réactivité-disponibilité et la valeur ajoutée des personnels supports comme la logistique, la qualité, la maintenance ?
Effectivité (plan MOE)	Est-il juste de demander aux opérateurs <i>via</i> un MES de prendre des décisions en autonomie sans les former et en les mettant en situation d'isolement au risque de faire peser sur leurs épaules une pression inconnue auparavant ?

4.3 Discussion

La mise en œuvre d'un déploiement de la performance fondé sur la vision et la définition proposées précédemment ouvre le débat sur bon nombre de points. Du point de vue industriel, l'idée retenue est de parvenir à « borner » aussi bien les actions sur les moyens que les objectifs assignés de sorte à garantir l'éthique. L'éthique devient ainsi une condition à satisfaire au fur et à mesure du pilotage des systèmes de production. Et la satisfaction de cette condition serait à planifier tout comme les autres objectifs.

Le positionnement du concept d'éthique par rapport à la RSE reste de notre point de vue insuffisant. Au sens RSE, l'éthique relève essentiellement du volet financier (paradis fiscaux, emploi de main d'œuvre jeune et conditions d'emploi, etc.). L'éthique au sens de la RSE concerne également le comportement des employés mais suggère l'alignement de leur éthique avec la recherche de la pérennité de l'entreprise (j'aide mes collaborateurs à progresser, etc.). Cependant, même si l'on comprend que la finalité d'une entreprise reste et restera de générer du profit, on peut comprendre que l'éthique des employés vis-à-vis de l'entreprise peut ne pas être jugée suffisante si on tient compte des enjeux sociétaux et du monde qui entoure ces entreprises. En ce sens, l'éthique doit également, de notre point de vue, intégrer le Développement Durable, de par, pour le moins, son pilier relatif à la société. De ce fait, il est possible d'imaginer le dilemme d'un employé qui, afin d'être éthique vis-à-vis de la société, contrevient à la recherche de pérennité de son entreprise (par exemple un lanceur d'alerte dans les domaines environnementaux ou de la protection de données).

D'un point de vue méthodologique, le déploiement de la performance en appellera à considérer deux points de vue du système, le point de vue physique/opérant et le point de vue décisionnel. L'éthique serait à associer de ce fait au comportement de tous les éléments du système, en particulier ses parties autonomes et intelligentes. Cette association aura à être effectuée en cohérence avec le modèle décisionnel de l'entreprise. En particulier, il pourrait être imaginé que la question de l'adéquation efficacité-éthique soit posée dès le niveau stratégique, celle de l'adéquation efficacité-éthique dès le niveau tactique alors que celle de l'adéquation efficacité-éthique relève du niveau opérationnel. Des questions peuvent être alors posées sur les décideurs à même d'opérer cette intégration de l'éthique. De même, les aspects à prendre en compte du fait de la considération de l'éthique peuvent à leur tour soulever des questions en termes de caractérisation. Dans le contexte de l'Industrie 4.0, la compétence d'un opérateur vis-à-vis de son remplacement par un Système Intelligent ou

l'autonomie d'un opérateur par exemple auront besoin d'être caractérisées, aussi bien par des objectifs que par des mesures. Par ailleurs, en présence de comportements dont les conséquences sont mal connues, il sera nécessaire de recourir à des dispositifs de simulation permettant d'envisager d'une part les différents scénarii, d'en évaluer d'autre part les conséquences et de capitaliser les informations obtenues.

D'un point de vue technique, étant donné la complexité induite par le contexte de l'Industrie 4.0, le recours aux outils de la simulation permettrait de tester différentes alternatives (décisions) avant d'appliquer une décision pour tester l'impact sur la performance éthique (approche de type *what-if?*). Les avancées dans les développements relatifs au concept de jumeau numérique le permettent, tout comme la démocratisation des *serious games* permettant de simuler différentes organisations candidates pour une entreprise ou même des procès fictifs permettant d'assigner les responsabilités juridiques ou éthiques soit aux hommes, soit aux entités artificielles et intelligentes [Mucchieli, 2018]. Cependant, il restera à déterminer des Indicateurs de Performance quantifiables dans ce contexte et c'est certainement l'un des aspects les plus délicats à traiter. Comment mesurer et comparer deux alternatives ou deux situations selon un angle éthique ? Comment diagnostiquer une situation selon un angle d'analyse en lien avec l'éthique ? Les concepts de base de la performance disposent de mesures (valeur monétaire, temporelle, etc.) et d'outils de diagnostic mais comment construire une mesure de l'éthique ou établir un degré « d'éthicalité » et l'expliquer ? Bien qu'une mesure absolue nous semble difficile, il est raisonnable de penser qu'au cas par cas, selon la culture d'entreprise ou l'environnement social, certains indicateurs puissent contribuer à une mesure plus relative et plus ou moins consensuelle. Par exemple, une mesure du bien-être social par un questionnaire, une mesure de l'adéquation travail/salaire, une enquête sur l'impact social d'une entreprise, une capacité à recruter les jeunes générations attentives aux évolutions sociétales, etc. permettent une évaluation, certes indirecte, de ce degré « d'éthicalité ». Dans tous les cas, il ne sera pas possible d'atteindre « l'excellence éthique industrielle » du premier coup et une démarche itérative s'imposera d'elle-même [Berrah et Trentesaux, 2020].

Enfin, considérer l'éthique dans le modèle de performance est naturellement indépendant de son contexte d'application. L'approche proposée pourra dans ce sens bénéficier des développements proposés dans le cadre des philosophies voisines citées précédemment.

5 CONCLUSION

La performance industrielle a connu une évolution majeure dans son passage d'une recherche d'optimisation fondée sur une simple diminution des coûts de production à une recherche d'optimisation associant aux coûts des aspects techniques permettant à l'entreprise à la fois d'intégrer les notions relatives à la satisfaction des clients et d'occuper une position concurrentielle avantageuse. Le modèle d'efficacité s'était alors élargi à celui du couple (efficacité, efficacité) auquel s'était greffée la notion d'effectivité, pour veiller à l'équilibre entre la déclaration des objectifs et l'utilisation des moyens pour les atteindre. Par l'avènement des Systèmes Autonomes et Intelligents, la digitalisation des activités de production et la profusion des données, l'Industrie 4.0 vient marquer l'évolution de la performance par la nécessité de prendre en compte les nouvelles caractéristiques des systèmes de production. Le bien-fondé du modèle (efficacité, efficacité, effectivité) établi jusqu'alors peut en effet être questionné, dans sa capacité de prendre en compte l'incertitude du comportement des Systèmes

Intelligents ainsi que les conséquences d'une digitalisation qui permet certes d'augmenter l'efficacité et l'efficience des systèmes mais qui peut avoir des effets négatifs par ailleurs.

Cet article a traité de la question de l'intégration de l'éthique dans le modèle de performance. Une extension du triangle (efficacité, efficacité, effectivité) à un tétraèdre permettant de considérer l'adéquation de ces trois notions usuelles aux valeurs éthiques a été proposée. Une étude de cas (équipementier aéronautique) a été décrite pour étayer les propos. Une analyse des conséquences de la digitalisation de ses activités a permis d'enrichir les réflexions sur les bases d'un modèle de performance 4.0. Même si l'éthique peut sembler duale avec la notion de performance industrielle, dans sa part taylorienne et « courttermiste », elle lui promet un fondement pérenne.

6 REMERCIEMENTS

Certains éléments présentés ont été discutés à l'occasion d'un exposé au groupe de travail SIMPA de la SAGIP du 4 décembre 2020. Les auteurs remercient les animateurs et membres de ce groupe pour les échanges qui ont conduit à enrichir cet article.

7 REFERENCES

- Anthony, R. N. (1965). *Planning and Control Systems : A Framework for Analysis*. Graduate School of Business Administration, Harvard University Press.
- Berrah, L., & Trentesaux, D. (2020). *Decision-Making in Future Industrial Systems: Is Ethics a New Performance Indicator? 10th Workshop on Service Oriented, Holonic and Multi-Agent Manufacturing Systems, for Industry of the Future, Studies in Computational Intelligence*, Paris, France, 1-2 octobre 2020.
- Bescos P.L., Dobler P., Mendoza C., Naulleau G. (1995), *Contrôle de gestion et management*, Montchrestien.
- Boisvert, H. (1995). *Comprendre, mesurer et gérer la productivité, la comptabilité par activités*. Congrès international de Génie industriel de Montréal : la productivité dans un monde sans frontières, Montréal, Canada, 18-20 octobre, 835-844.
- Brimson, J. A. (1991). *Activity Accounting : An Activity-Based Costing Approach* (John Wiley&sons).
- Burritt, R., & Christ, K. (2016). *Industry 4.0 and environmental accounting : A new revolution? Asian Journal of Sustainability and Social Responsibility*, 1(1), 23-38.
- Clivillé, V., Berrah, L., Foulloy, L., & Chapel, C. (2019). *Mise en place du MES: Le témoignage d'un fournisseur aéronautique*. CIGI QUALITA 2019: 13ème Congrès International de Génie Industriel et Qualita, Montreal, Canada, 25-28 juin.
- Ducq, Y., Vallespir, B., & Doumeingts, G. (2001). *Coherence analysis methods for production systems by performance aggregation*. *International Journal of Production Economics*, 69(1), 23-37.
- Fettig, K., Gacic, T., Koskal, A., Kuhn, A., & Stuber, F. (2018). *Impact of Industry 4.0 on Organizational Structures*. *IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Stuttgart, Allemagne, 17-20 juin.
- Gasparatos, A., El-Haram, M., & Horner, M. (2009). *The argument against a reductionist approach for measuring sustainable development performance and the need for methodological pluralism*. *Accounting Forum*, 33(3), 245-256.

- Ghalayini, A. M., Noble, J. S., & Crowe, T. J. (1997). An integrated dynamic performance measurement system for improving manufacturing competitiveness. *International Journal of Production Economics*, 48(3), 207-225.
- Ghobakhloo, M. (2018). The future of manufacturing industry: A strategic roadmap toward Industry 4.0. *Journal of Manufacturing Technology Management*, 29(6), 910-936.
- Gilbert, G. R., Collins, R. W., & Brenner, R. (1980). Age and leadership effectiveness: From the perceptions of the follower. *Human Resource Management*, 29(2), 187-196.
- Globerson, S. (1985). Issues in developing a performance criteria system for an organization. *International Journal of Production Research*, 23(4), 639-646.
- GRI. (2020). Standards. <https://www.globalreporting.org/standards/>
- Imai, M. (1992). *Kaizen: La clé de la compétitivité japonaise*. Eyrolles.
- ISO 22400. (2014). Automation systems and integration- Key performance indicators (KPIs) for manufacturing operations management.
- ISO 26000. (2008). Lignes directrices relatives à la responsabilité sociétale. AFNOR.
- Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research*, 57(3), 829-846.
- Jacot, J. H., & Micaelli, J. P. (1996). La question de la performance globale, dans *La performance économique en entreprise*. J.H. Jacot et J.P. Micaelli (eds).
- Johnson, H. T. (1975). Management Accounting in an Early Integrated Industrial: E. I. duPont de Nemours Powder Company, 1903-1912*. *Business History Review*, 49(2), 184-204.
- Kanigel, R. (2005). *The One Best Way*. The MIT Press.
- Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2018). Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process Safety and Environmental Protection*, 117, 408-425.
- Kaplan, R., & Norton, D. (1992,1). *The Balanced Scorecard: Measures that Drive Performance*. Harvard Business Review.
- Le Moigne, J.-L. (1990). *La modélisation des systèmes complexes*. Dunod.
- Lorino, P. (1996). *Méthodes et pratiques de la performance*. Éditions d'Organisation.
- Mishra, D., Gunasekaran, A., Papadopoulos, T., & Dubey, R. (2018). Supply chain performance measures and metrics: A bibliometric study. *Benchmarking: An International Journal*, 25(3), 932-967.
- Morahan, M. (2015). Ethics in management. *IEEE Engineering Management Review*, 43(4), 23-25.
- Mucchieli, J. (2018). Procès du carambolage du siècle: « L'intelligence artificielle a-t-elle une éthique? » Dalloz actualités, accessible en ligne <https://www.dalloz-actualite.fr/flash/proces-du-carambolage-du-siecle-l-intelligence-artificielle-t-elle-une-ethique#.X85dARJCdEa>. https://www.dalloz-actualite.fr/flash/proces-du-carambolage-du-siecle-l-intelligence-artificielle-t-elle-une-ethique#.X8_k9BfZCb_
- Neely, A, Mills, J., Platts, K., Gregory, M., & Richards, H. (1996). Performance measurement system design: Should process based approaches be adopted? *International Journal of Production Economics*, 46-47, 423-431.
- Neely, Andy, Gregory, M., & Platts, K. (1995). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management*, 15(4), 80-116.
- Nudurupati, S. S., Bititci, U. S., Kumar, V., & Chan, F. T. S. (2011). State of the art literature review on performance measurement. *Computers & Industrial Engineering*, 60(2), 279-290.
- Ono, T. (1988). *Toyota Production System: Beyond Large-Scale Production* (productivity press). CRC Press.
- Pacaux-Lemoine, M. P., & Trentesaux, D. (2019). Ethical risks of human-machine symbiosis in Industry 4.0: Insights from the human-machine cooperation approach. *IFAC-PapersOnLine*, 52(19), 19-24.
- Popova, V., & Sharpanskykh, A. (2010). Modeling organizational performance indicators. *Information Systems*, 35(4), 505-527.
- RGPD. (2020). Le règlement général sur la protection des données (RGPD), mode d'emploi, accessible sur: <https://www.economie.gouv.fr/entreprises/reglement-general-sur-protection-des-donnees-rgpd>.
- Ricoeur, P. (1990). *Soi-même comme un autre*, Editions Seuil.
- Rosa, P., Sassanelli, C., & Terzi, S. (2019). Towards Circular Business Models: A systematic literature review on classification frameworks and archetypes. *Journal of Cleaner Production*, 236, 117696.
- Saturno, M., Pertel, V. M., Deschamps, F., & Loures, E. (2017). Proposal of an automation solutions architecture for Industry 4.0. *Transactions on Engineering and Technology Research*. 24th International conference on production research, ICPR 2017, Poznan, Poland, 30 July-3 August.
- Shah, L. A., Amjad, F., & Vernadat, F. (2020). Framework of performance measurement and management for Industry 4.0 era. 13th Conference on Modeling, Optimization and Simulation, MOSIM 2020, Agadir, Maroc, 12-14 novembre.
- Sikdar, S. K. (2003). Sustainable development and sustainability metrics. *AIChE Journal*, 49(8), 1928-1932.
- Taylor, F. W. (1911). *The principles of scientific management*. New York, London, Harper & Brothers.
- Tobon Valencia, E., Lamouri, S., Pellerin, R., & Moeuf, A. (2020). Modélisation du programme directeur de production en vue de la transition numérique des PME manufacturières. 13th Conference on Modeling, Optimization and Simulation, MOSIM 2020, Agadir, Maroc, 12-14 novembre.
- Trentesaux, D., & Caillaud, E. (2020). Ethical stakes of Industry 4.0. *IFAC World Congress*, 11-17 juillet.
- Wisner, J. D., & Fawcett, S. E. (1991). Link Firm Strategy to Operating Decisions Through Performance Measurement. *Production and Inventory Management Journal*. Third quarter, 5-11.
- Womack, J. P., Jones, D. T., & Roos, D. (1990). *Machine That Changed the World*. Free Press, New York.

The key catalyst events for the evolution of the sustainable development concept

BELMABROUK AMAL¹, ALAEDDINE ZOUARI¹, MICHEL TOLLENAERE²

¹UNIVERSITE DE SFAX, ISGIS , OLID , LR19ES21

3021 , Sfax , Tunisie

amalbelmabrouk1@gmail.com

ala.zouari@isgis.usf.tn

²UNIVERSITE GRENOBLE ALPES, GRENOBLE INP, G-SCOP

michel.tollenaere@grenoble-inp.fr

Résumé – La façon dont nous abordons le développement affecte tout le monde et presque tout. Les répercussions de nos décisions et de nos actes en tant que société ont des conséquences très graves sur la vie des générations actuelles et sur la vie des générations futures également. Le développement durable offre une approche pour prendre de meilleures décisions sur les questions qui affectent toutes nos vies. Dans cet article, nous cherchons à remonter loin dans l'histoire pour retracer les racines du concept de développement durable et illustrer comment l'idée de ce concept a évolué à travers les siècles en opposition aux notions de progrès. Le contexte historique du XIX^e siècle au siècle actuel est esquissé.

Abstract – The way we approach development affects everyone and almost everything. The impacts of our decisions and acts as a society have very critical consequences for the current generations' lives and for future generations' lives as well. Sustainable Development provides an approach for making better decisions on the issues that affect all of our lives. In this paper, we aim to go far back in history to trace the roots of the concept of Sustainable Development and illustrate how the idea of this concept evolved through the centuries as a counter to notions of progress. The historical context from the nineteenth century until the actual century is outlined.

Mots clés – Développement durable, Durabilité, Histoire, Événement clés

Keywords – Sustainable Development, Sustainability, History, Key events

1 INTRODUCTION

This paper sketches the evolution of the idea of Sustainable Development by studying the roots of this concept throughout history and highlighting the key events contributing to the emergence and progression of Sustainable Development from the 19th century until the actual century. In the remainder of the paper, we will provide some basic definitions and concepts (2). Then, an historical view will be presented with a structure inspired from [Mebratu, 1998], [Jordan and Voisey, 1994] as follows:

- Pre-Stockholm : covering the period until the Stockholm Conference on Environment and Development (19th century to the early 20th century) (section3);
- From Stockholm to World Commission on Environment and Development (WCED) : covering the period between the Stockholm Conference and the WCED (the late 20th century) (section4);
- From WCED to Johannesburg : covering the period between the WCED and the Johannesburg's Conference (late 20th century and early 21st century) (section5);
- Post-Johannesburg : covering the period after the Johannesburg's Conference (early 21st century) (section6)

2 BASIC DEFINITIONS AND CONCEPTS

2.1 Sustainable Development

The earliest meaning of sustain is to “support” uphold the course of or “keep into being”. Another meaning is to provide with food and drink, or the necessities of life [Pezzoli, 1997]. Yet, the very vagueness of the term Sustainable Development (SD) has generated different interests and views about its issues.

Thus, there is no single clear definition of SD [McCormick, 1986], [Slim, 1995]. Since the release of The Brundtland Commission Report, the definition given by WCED is the resolution, most influentially; it has been most commonly used at the global level to cover various perspectives of SD [Faucheux et al., 1998], [Gibson et al., 2013], [Sood and Sooden, 2016]. The Brundtland commission [WCED, 1987] has defined the SD as the development that meets the needs of the present without compromising the ability of future generations to meet their own needs. It contains within it two key concepts:

- the concept of 'needs', in particular the essential needs of the world's poor, to which overriding priority should be given; and
- the idea of limitations imposed by the state of technology

and social organization on the environment's ability to meet present and future needs.

2.2 Sustainable Development and Sustainability

Sustainability is an idea that has never been really defined [Worster, 1994]. The concept of Sustainability is both very ancient and quite recent. The term was definitely put on the political agenda in 1987, through the report of the World Commission on Environment and Development [Faucheux et al., 1998]. Our planet has an exhaustible store of resources and does not have an unlimited capacity to accept waste [Elms and Wilkinson, 1995]. In the same vein, the term Sustainability brings to mind the image of a society, including all human activities, in balance with natural surroundings, respecting ecological limits, and capable of persisting far into the future [Faucheux et al., 1998]. As for the difference between the SD and Sustainability, [Robinson, 2004] stated that the SD is the technical fix and the Sustainability is the change value. Generally, Sustainability is often thought of as a long-term target, while SD addresses processes and pathways to achieve this long-term target.

3 PRE-STOCKHOLM

3.1 19th Century

By 1798, many of the evil effects of the industrial revolution had surfaced. Unemployment, poverty, and disease were already problems calling for remedial treatment [Mebratu, 1998]. This makes it of a broad interest to propose prompt solutions and determine some limits to mitigate the undesirable impacts of the industrial revolution. In fact, more than a century before the term SD came into general use, a number of publications appeared which dealt with what we would today call SD [Malthus et al., 1992].

3.1.1 Theory of population

One of the most famous propositions in the 19th century was the "Theory of population". It belongs to Thomas Robert Malthus (1766-1834)¹. Malthus believed that though the produce of the earth would be increasing every year, population would have the power of increasing much faster, and this superior power must necessarily be checked by the periodical or constant action of moral restraint, vice, or misery [Wall and Ashford, 1994].

3.1.2 Emergence of Preservation and Conservation Movements

While the term SD was probably coined during the 1980s, its intellectual pedigree goes back much further. Concern with the environmental impacts of human activities in Western history goes back at least to the Greeks [Cohen et al., 1998]. In his retrospective assessment of the successes and failures of the 19th century, called "our wonderful century", published in 1898, Alfred Russell Wallace included a chapter on the plunder of the Earth. He discussed the damage done by the 'reckless destruction of the stored-up products of nature' and regarded the unlimited extraction of coal, oil, gas and minerals, and the exploitation of the rain forests as an "injury done to posterity" [Malthus et al., 1992]. At the end of the 19th century, two prominent schools have been developed of environmental thought in the US: the preservationists associated with Muir (1901 by way of Thoreau and Emerson), and the conservationists associated with Gifford

¹ An English cleric and scholar, influential in the fields of political economy and demography.

Pincho.² The preservationist tradition was transcendentalist in nature, arguing for the preservation of nature. The conservationists were explicitly utilitarian, arguing that conservation of nature was needed to ensure the continued availability of natural resources for human use in the future [Du Pisani, 2006]. From this overview it is clear that the roots of the concept of SD and the idea of Sustainability can be traced back to ancient times.

3.2 20th Century

3.2.1 Early 20th century

The concept of SD has greatly flourished in the 20th century. It emerged in the context of a growing awareness of an ecological crisis [De Souza Rolimi, 2008]. The era between the late 19th century and early 20th has witnessed a slightly rising interest with regard to the SD. According to [Pezzoli, 1997] concerns about Sustainability were expressed by the Regional Planning Association of America and the Tennessee Valley Authority. Noteworthy, the precise term of Sustainability was not yet used. Furthermore, in the 1937, the Institute of International Law had the foresight to adopt the Luxemburg Resolution³. In addition to the purpose of collaboration on the progressive codification of international law, this resolution served as legal background to the development of the new law of the sea and maritime and environmental law [Manulak, 2015].

However, the field of SD tended to be neglected during the Great Depression (lasting from 1929 to 1939), the Second World War (1939 to 1945), and the post-War boom [Du Pisani, 2006]⁴. It began to re-emerge in a new form in the 1960s, with the rise of protests over nuclear power and various forms of air and water pollution. Indeed, the urgent need for SD was discussed in depth at the 1949 United Nations Scientific Conference on the Conservation and Utilization of Resources held at Lake Success, New York [Slim, 1995]. In addition to that, [Manulak, 2015] reported that the evolution of environmental law in correlation with the law of the sea stands out as fast-paced from the 1970s. Still in the same context, the International Union for Conservation of Nature and Natural Resources (IUCN) -created in 1948- discussed the need to reconcile conservation and development.

All the actions mentioned above led to the emergence of the "environmental movement" of the 1970s and 1980s [Du Pisani, 2006].

3.2.2 Late 20th century

The late 20th century witnessed numerous actions and propositions. Starting with the increase of Canadian concerns over pollution in Arctic waters increased greatly after oil was discovered on Alaska's North Slope in 1968. The government introduced the Arctic Waters Pollution Prevention bill in April 1970 [De Rome and Meadows, 1974]. Later in the same year, the World Bank had established a new position of environmental advisor. Yet, the bank faced considerable resistance from Third World governments. They thought that would add 25 to 50 percent to total project costs [Slim, 1995]. Furthermore, in 1973, IUCN and the Conservation Foundation published a guide book entitled: Ecological Principles for Economic Development. It examined the basic ecological

² Back at the turn of the 20th Century Gifford Pinchot and John Muir had radically contrasting views of how to manage America's wild lands. More precisely, Muir promoted preservation and Pinchot advocated for conservation.

³ A referendum on the Order law was held in Luxembourg on 6 June 1937.

⁴ A broad period of worldwide economic expansion beginning after World War II and ending with the 1973-1975 recession.

concepts that needed to be considered in instances where economic development threatened major changes in natural systems [Slim, 1995]. At around the same time of the Stockholm Conference, the global environmental crisis was expanding at an alarming rate.

A group of environmental activists, later to be known as the Club of Rome⁵, produced a comprehensive report on the state of the natural environment entitled “Limits to growth”. It referred to a ‘condition of ecological and economic stability that is sustainable far into the future’ and capable of satisfying the basic material needs of all people [Mebratu, 1998], [Malthus et al., 1992], [Goldsmith, 1972]. Consecutively, a book entitled “A Blueprint for survival” was released later in the same year.

The author of this book: Goldsmith in [Goldsmith, 1972], believed that the principal defect of the industrial way of life with its ethos of expansion is that it is not sustainable [Kumar et al., 2016]. Even though the precise time and circumstances are in doubt, but still we can be sure that sooner or later it will end.

Figure 1 (elaborated by the authors) illustrates the pre-Stockholm phase relevant events of SD emergence.

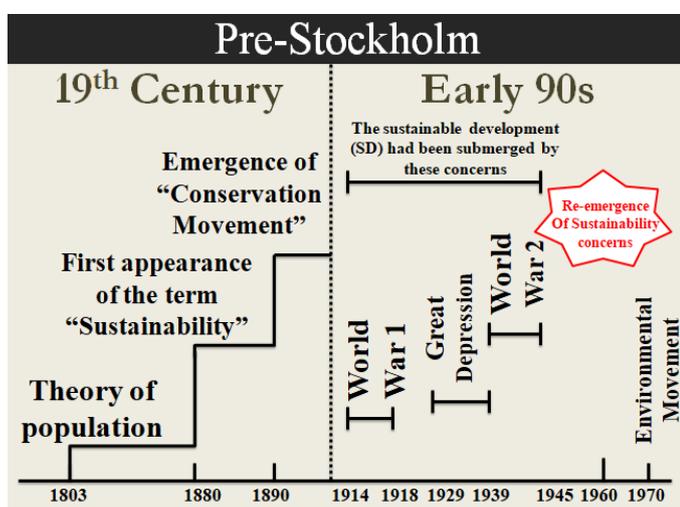


Figure 1. Pre-Stockholm phase

4 FROM STOCKHOLM TO WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT

4.1 Stockholm Conference

Aware of the predicament and profound challenge described by the Club of Rome, prominent scholars and leaders came together for a series of epoch-making meetings [Pezzoli, 1997]. The United Nations (UN) sponsored 1972 Stockholm Conference on the Human Environment. This led to the establishment of the United Nations Environment Program (UNEP) in 5 June 1972.

Furthermore, it adopted a non-binding Declaration of Principles and a Program of Action (Stockholm Declaration) in an attempt to find solutions to complex and interconnected environmental issues. Moreover, it placed the environment on the international agenda for the first time [Dodds et al., 2016]. The Stockholm Conference and the declaration that the participating States adopted at the end of the conference are generally considered as the starting point for what constitutes

⁵ An informal organization that has been aptly described as an “invisible college”.

today the modern international law and policy for the environment [WCED, 1987].

4.2 United Nations Environment Program and United Nations Conferences

Stockholm became the precursor to the international agenda of SD, which was first promoted at its ten-year review held in 1982 in Nairobi, the seat of UNEP [Forsch and Gallopoulos, 1989]. Thus, following Stockholm, UNEP and the UN Conference on Trade and Development proposed meetings. The first of these meetings, held in Geneva in April 1974, concluded that because of the complexity of international economic relations, difficulties associated with the distribution of natural resources would constrain development well before any problems of finite quantities arose [Slim, 1995].

4.3 Cocoyoc Declaration

The second meeting was a Symposium on Patterns of Resource Use, Environment and Development Strategies, held at Cocoyoc, Mexico, in October 1974 [Slim, 1995]. In 1974, a group of ten of the world’s development experts (all men) met at Cocoyoc in Mexico to try to set a new agenda of alternative development, to move forward from what they considered to be the failure of development in the 1950s and 1960s. They produced the Cocoyoc Declaration, in which they made a basic distinction between priorities relating to the “inner limits”⁶ and those relating to the “outer limits”⁷ of development [Slim, 1995]. As stated in [Pezzoli, 1997] the term SD was already in use at the time of the 1974 Cocoyoc seminar. These tendencies were reinforced by the energy crisis of 1973, its successor in 1979, which became a central focus of environmental concern [Du Pisani, 2006].

4.4 World Conservation Strategy

The World Conservation Strategy was launched internationally in 1980. It was the fruit of the hard work of the IUCN [Mebratu, 1998], with the advice, cooperation and financial assistance of the UNEP and the World Wildlife Fund and in collaboration with the Food and Agriculture Organization of the UN and the United Nations Educational, Scientific and Cultural Organization.

A decade after Stockholm, environmentalists were more conciliatory [Slim, 1995]. Nonetheless, a number of global environmental challenges had clearly not been adequately addressed [Wackernagel and Rees, 1998], [Raskin et al., 1998]. These challenges take credit for interlacing the goals of international security and SD [Nordquist et al., 2003]. In the 1980s the new paradigm of SD became more widely used. As incorporated for the first time into an international forum for the Conservation of Nature’s World Conservation Strategy [Faucheux et al., 1998], [Worster, 1994].

Figure 2 (elaborated by the authors) presents the main events of SD strategies in the late nineteen’s especially between Stockholm conference and the WCED.

⁶ The inner limits cover ‘fundamental human needs’ like food, shelter, health, and human rights.

⁷ The outer limits relate to aspects of ‘the planet’s physical integrity’ like the environment and population

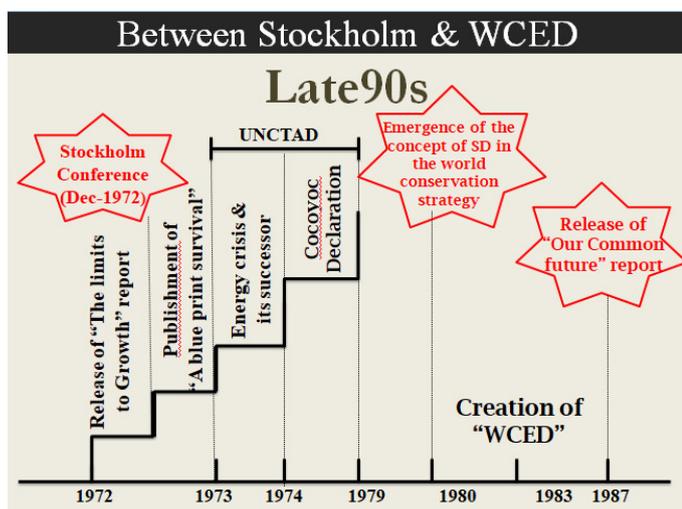


Figure 2. From the Stockholm conference to the WCED

5 BETWEEN WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT AND JOHANNESBURG

5.1 World Commission On Environment And Development

"A global agenda for change" - this was what the WCED was asked to formulate. It was an urgent call by the General Assembly of the United Nations [Sullivan, 1993]. Over a three-year period (1984–1987), the Commission conducted public hearings on five continents. This brought the discussion of SD to literally millions of people around the world [Pezzoli, 1997]. The WCED definition of SD was vague and ambiguous. Despite that, it has been highly instrumental in developing a "global view" with respect to our planet's future [Mebratu, 1998]. With regard to international issues, the creation of the WCED by the UN in 1985 represented a turning point in the sophistication of the debate and the incorporation of the concerns of developing countries [Gibson et al., 2013], [Du Pisani, 2006].

The report, *Our Common Future*⁸, published by WCED, is taken as a starting point for most current discussions on the concept of SD. It focused primarily on the needs and interests of humans, and was concerned with securing a global equity for future generations by redistributing resources towards poorer nations to encourage their economic growth in order to enable all human beings to achieve their basic needs [Malthus et al., 1992].

5.2 Earth Summit

The work of the WCED prefigured a series of international conferences concerning environmental development relations [Pezzoli, 1997]. And, as mentioned in [Death, 2011], Rio inaugurated the heyday of the era of mega-conferences. One of the most influential conferences after WCED is the UN Conference on Environment and Development (UNCED)⁹, which is also known as the "Rio Conference" or the "Earth Summit", which took place in Rio de Janeiro. The 12-day conference was the largest the UN ever held [Pezzoli, 1997].

In rhetoric, if not yet in agenda and action, the nations of the world pronounced their commitment to SD at the 1992 Earth

⁸ also known as the Brundtland Report, published in dozens of languages

⁹ It was attended by 116 heads of state or government, 172 states, 8000 delegates, 9000 members of the press and 3000 accredited representatives of non-governmental organizations

Summit [Nordquist et al., 2003]. UNCED led to production of major international documents such as the Rio Declaration, Agenda 21¹⁰. As a result, SD was taken to every corner of the world.

5.3 The World Conference on Human Rights

In June 1993, 171 states were gathered by the United Nations in Vienna to confer on human rights. It was the largest assembly ever on global human rights issues. The June 1993 World Conference on Human Rights witnessed the extraordinary success of efforts by women's rights activists worldwide to end the historic disregard of human rights violations against women [Boyle, K. 1995].

The outputs of this conference were the Vienna Declaration and Programme of Action. The World Conference on Human Rights in Vienna had been preceded by the Rio conference on Environment and Development, in 1992 and was followed in September 1993 by the UN conference on Population and Development held in Cairo. A conference on World Social Development will be held in March 1995 in Copenhagen, to be followed by the Fourth World Conference on Women in Beijing in September 1995 [Fomerand, 2009].

5.4 World Summit for Social Development

As pointed out by [Schechter, 2009], the year 1995 was worldwide, yearlong celebrations marking the fiftieth anniversary of the United Nations. Perhaps one of the most notable events was the World Summit for Social Development, lasted from 6 to 12 March 1995 in Copenhagen, Denmark, concerned with eradication of poverty, expansion of productive employment, reduction of unemployment, and social integration.

5.5 Habitat II Conference

Sustainability and SD were also the core theme at the 1996 Habitat II Conference held in Istanbul. Habitat II (the City Summit) was the last major UN conference to be held this century. Representatives of 171 countries and a total of 16400 people addressed two major themes: adequate shelter for all; and sustainable human settlements development in an urbanizing world. Habitat II concluded with publication of the Habitat Agenda [Pezzoli, 1997].

5.6 Rio Cluster

The Earth and City Summits together with several other UN conferences and conventions are collectively referred to as the "Rio Cluster". The Rio Cluster has filled interest in Sustainability in international as well as national, regional and local settings around the world. This rising interest has been accompanied by a rising tide of literature that mostly falls into the fields of managerialism, policy, and planning [Pezzoli, 1997].

5.7 Kyoto Protocol

Beyond promoting economic globalization and keeping the peace, global governance became increasingly irrelevant. The international negotiation continued on critical environmental and social problems. But they were either vastly insufficient like the Kyoto Protocol on greenhouse gasses [Nordquist et al., 2003]. On 1997, the nineteenth special session of the General Assembly was organized (23-27 June) to review progress made five years after the Earth Summit in Rio [Schechter, 2009]. Unfortunately, at the end of the twentieth century, the

¹⁰ A document that spells out some initiatives, promoting Education, Public Awareness..., to be put into action

international momentum for a sustainable future seems squandered.

5.8 Millenium Summit and Sustainable Development Conference

SD, the half-forgotten battle cry of the late 20th century and early 21st century, was resurrected [Nordquist et al., 2003]. Millenium Summit of the United Nations was held from 6 to 8 September 2000 at the United Nations head quarter, New York [Fisher, 2010]. The outcome of this Summit was the Millenium Summit Declaration [Burleson, 2016] and the Millennium Development Goals (MDGs). In this Summit several values were highlighted such as: Defense of civil rights, respect for human dignity, integrity, impartiality, tolerance for diversity, as well as protection of nature, the vulnerable segments of the world community and so on.

5.9 Johannesburg's Conference

Thirty years after the first United Nations conference on the environment in Stockholm, ten years after the Earth Summit in Rio de Janeiro, and five years after the adoption of Kyoto Protocol [Sachs, 2012]. The World Summit on SD was held in Johannesburg, South Africa, from 26 August to 4 September 2002. The Johannesburg's Conference was a hinge event [Nordquist et al., 2003].

The main objective was to look back on three decades of implementation since Stockholm and one decade since Rio, and to agree on a new and ambitious plan of action; the few advances made were, regrettably, quickly eroded [Forsch and Gallopoulos, 1989].

Figure 3 (elaborated by the authors) shows the temporal evolution of SD events in the late nineteen's after the WCED.

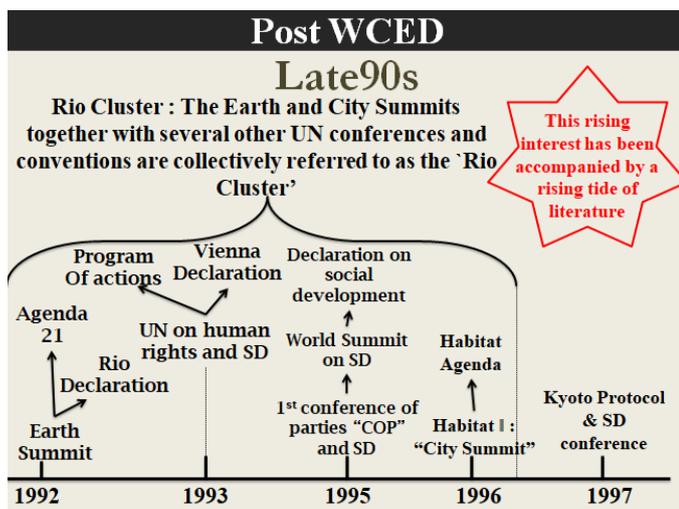


Figure 3. Pertinent events post WCED

6 THE EARLY 21ST CENTURY

By 2002, the 9/11 terrorist attack awakened a sleepwalking global elite to deep fissures cutting across the geo-political landscape. The nations of the world, mobilized in a cooperative effort to fight terrorism, they offered an unexpected opportunity to redirect development strategy and commit to a form of globalization that is more inclusive, democratic and sustainable [Nordquist et al., 2003].

6.1 Copenhagen's Conference

Since its inception at the United Nations Conference on Environment and Development in 1992, the climate regime has

been relatively inclusive of stakeholders from its many constituencies [Christoff, 2010]. The 15th Conference of the Parties (COP'15), held between 7 and 18 December 2009 in Copenhagen¹¹, is rightly seen as a critical moment in climate negotiations. Hence, it is considered as the largest international climate meeting to date. This conference needed to succeed in order to ensure orderly preparation for the expiry of the Kyoto Protocol's first commitment period at the end of 2012 [Bernstein, 2013].

6.2 The Global Sustainability Panel

On 9 August 2010, Mr. Ban Ki-moon, the UN Secretary-General, launched a High-level Panel on Global Sustainability (GSP). This is an initiative following the path of the Brundtland Commission on SD and its report Our Common Future, a key document for the milestone Rio conference on environment and development in 1992 [Burleson, 2016].

The role of the GSP was to bring together important world figures to formulate a new blueprint for a sustainable future on a planet under increasing stress resulting from human activities and to explore approaches for building a low carbon, green and resilient economy that can eradicate poverty and ensure a dignified life for all [Forsch and Gallopoulos, 1989].

6.3 Rio 20+

Global summits and major conferences – from Stockholm through Rio to Johannesburg and from Kyoto to Copenhagen – have become firmly established as landmark moments of environmental governance. The announcement of a “Rio + 20 Summit” in Brazil in 2012 confirms their continued prominence as benchmarks [Sachs, 2012]. The Rio 20+ meeting largely passed below the radar [Forsch and Gallopoulos, 1989]. Rio+20 is symptomatic of what we observe in a growing number of multilateral forums that require compromises to build coherence and momentum on global concerns [Sharma et al, 2018]. The legacy of Rio+20 consisted of the agreement to develop integrated goals; renew environmental and developmental institutions; develop green and inclusive economic policy frameworks [Forsch and Gallopoulos, 1989]. The final outcome of the conference, titled “The Future We Want”, contains 283 paragraphs and is 53 pages long.

6.4 COP'21 and the Paris Agreement

The Paris COP 21 Decision and Agreement map collective climate engagement on such thorny elements as [Asadnabizadeh, 2019]:

- Mitigation - reducing emissions fast enough to achieve the temperature goal
- A transparency system and global stock-take – accounting for climate action
- Adaptation - strengthening ability of countries to deal with climate impacts
- Loss and damage - strengthening ability to recover from climate impacts
- Support - including finance, for nations to build clean, resilient futures.

The Paris Agreement relies on the political will of respective parties to collectively offer ambitious climate action plans [Asadnabizadeh, 2019]. In this framework, UN executive Secretary Christiana Figueres, emphasized again and again at COP 21 “Paris was more of a beginning than an end”.

¹¹ It involved nearly 27,000 people, including 10,500 delegates representing 190 states, and over 120 heads of state and government.

6.5 From Millennium Development Goals to sustainable development goals

In 2015, 195 nations agreed on the fact that they can change the world for the better. This will be accomplished by bringing together their respective governments, businesses, media, institutions of higher education, local NGOs and so on, to improve the lives of the people in their country by the year 2030. Consequently, the idea of the SD goals (SDGs) has quickly gained ground because of the growing urgency of SD for the entire world [Burlison, 2016]. The SDGs can be broadly divided into three categories: First, an extension of MDGs that includes the first seven SDGs; second group is inclusiveness (jobs, infrastructure, industrialization, and distribution). It includes goals 8, 9, and 10; and the third group is on Sustainability and urbanization that covers the last seven goals: sustainable cities and communities, life below water “consumption and production; climate action; resources and environment; peace and justice; and the means of implementation and global partnership for it” [Jordan and Voisey, 1994].

6.6 COP'24 and COP'25

The 24th Conference of Parties (COP'24) to the UNFCCC was an important milestone for the international climate change regime, expected to further animate the Paris Agreement by adopting the detailed rules for its implementation [Burlison, 2016].

Although, [Asadnabizadeh, 2019] stated that the COP24 was particularly crucial because 2018 was the deadline that the signatories of the Paris Agreement agreed upon in order to adopt a work program for the implementation of the Paris Commitments.

On the last day of the Polish Presidency of COP'24, Michal Kurtyka, President of COP24 who is now the new and firstever Minister of Climate in Poland as well handed the Presidency over to Chile in Madrid and congratulated Ms. Carolina Schmidt on becoming President of COP'25. The Pre COP'25 was held at the Costa Rica Convention Center, from October 8th to 10th, 2019.

Figure 4 (elaborated by the authors) describes the chronology of most relevant events of SD in early 21st century.

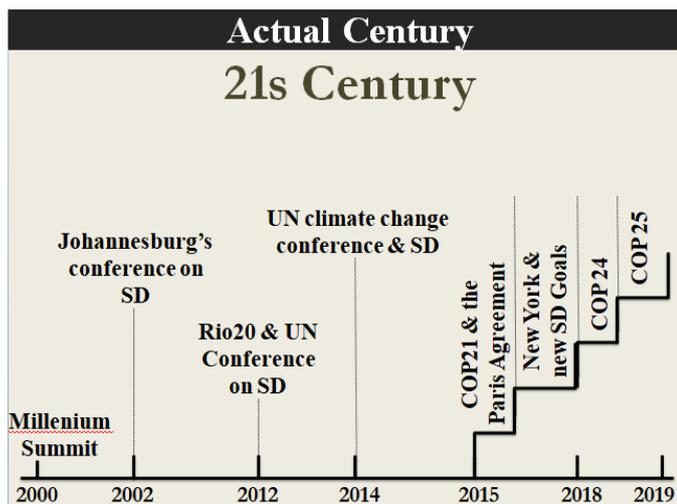


Figure 4. Relevant events of SD in 21st century

7 STATISTICS AND ANALYSIS

As illustrated in Figure 5 (elaborated by the authors), the major interests of almost half of the conferences and events

mentioned in this paper (42.8%) were dedicated to environmental issues. It is intuitive, because of the disasters caused due to human activities, industrial revolutions, wars, and so on. The SD as a whole (3 pillars: economic, social and environmental) was of great interest since the release of the Brundtland report (28.57%).

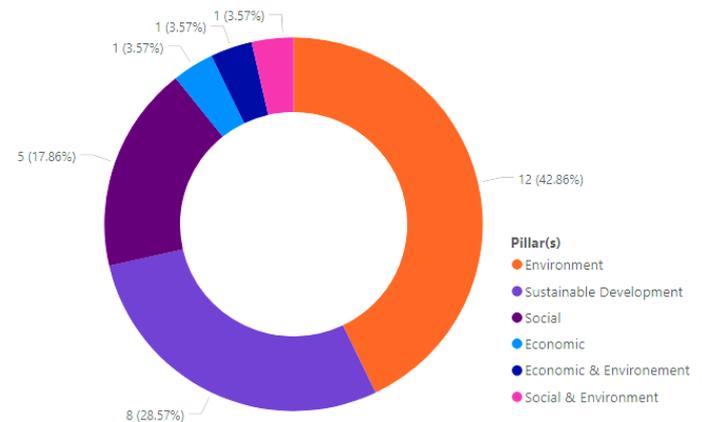


Figure 5. Percentage of key events per pillar

Concerning the social pillar, concerns have been continually expressed all along the 19th, 20th centuries and the actual century, to eradicate poverty and illiteracy, promote education, and defend women’s rights. Indeed, a very remarkable improvement and progress were noticed. Regarding the economic pillar, more attention needs to be consecrated to this pillar.

8 CONCLUSION AND DISCUSSION

In the 19th century, the environmental impacts of human activities were the center of interest and the necessity to control the exploitation of natural resources was expressed.

However, it was just rhetoric; there were no real actions or plans. But, the 20th century was the century of rising awareness, despite the fact that it was full of obstructive events: war, economic crisis, and so on. Several Summits and Conferences were made. Unlike the 19th century, social, environmental and economic pillars were almost equally important. Consequently, a slight progress has been noticed in many areas such as equity between men and women, a decrease in illiteracy rate, and so on. But still, no sufficient result was reached. In line with this evolution, a series of norms and standards (ISO 14040:1997...) have emerged to enhance industrials and manufacturers to follow the pathway of SD, and facilitate the procedure of implementation of SD principles. Taking into consideration the whole supply chain, standards, guidelines and norms insist on integrating, interactively, all the elements to achieve SD goals.

By the end of the 20th century, emergency calls were launched. Thus, in the 21st century, the peak of awareness has been reached. This century is witnessing a lot of events and actions enhancing the integration of SD in all activities: transport and logistics, manufacturing, agriculture, and so on. All the implemented actions are attempts to rescue the planet, mitigate the impact of human activities on earth, ensure the well-being of all the habitats of the earth, etc. Great results are generated such as: in 2019 the Ozone Hole was the Smallest on record since its discovery. But on the other hand, a lot of violations of SD principles still exist, and the results reached are not sufficient enough. As a result, so much more effort needs to be dedicated to spread the concept SD and integrate it in every area as one of the most important strategic axes.

9 REFERENCES

- Asadnabizadeh, M. (2019). Development of UN Framework Convention on Climate Change Negotiations under COP25: Article 6 of the Paris Agreement perspective. *Open Political Science*, 2(1), 113-119.
- Bernstein, S. (2013). Rio+ 20: Sustainable development in a time of multilateral decline. *Global Environmental Politics*, 13(4), 12-21.
- Boyle, K. (1995). Stock-taking on human rights: The world conference on human rights, Vienna 1993. *Political Studies*, 43(1), 79-95.
- Burleson, E. (2016). Paris agreement and consensus to address climate challenge. ASIL INSIGHT, Forthcoming.
- Christoff, P. (2010). Cold climate in Copenhagen: China and the United States at COP15. *Environmental Politics*, 19(4), 637-656.
- Cohen, S., Demeritt, D., Robinson, J. and Rothman, D. (1998). Climate change and sustainable development: towards dialogue. *Global Environmental Change*, 8(4), 341-371.
- De Rome, C. and Meadows, D. H. (1974). *The limits to Growth: a Report for the Club of Rome's Project on the Predicament of Mankind*. Universe Books.
- De Souza Rolim, M. H. F. (2008). *The international Law on Ballast Water: Preventing Biopollution*. Martinus Nijhoff.
- Death, C. (2011). Summit theatre: exemplary governmentality and environmental diplomacy in Johannesburg and Copenhagen. *Environmental Politics*, 20(1), 1-19.
- Dodds, F., Laguna-Celis, J. and Thompson, L. (2014). *From Rio+20 to a new development agenda: Building a bridge to a sustainable future*. Routledge, United Kingdom.
- Du Pisani, J. A. (2006). Sustainable development-historical roots of the concept. *Environmental sciences*, 3(2), 83-96.
- Elms, D. and Wilkinson D. (1995). *The Environmentally Educated Engineer: Focus on Fundamentals*. (eds). Centre for Advanced Engineering.
- Faucheux, S., O'Connor, M. and Van Der Straaten, J. (1998). *Sustainable development: Concepts, rationalities and strategies* (Vol 13). Springer, Dordrecht
- Fisher, D. R. (2010). COP-15 in Copenhagen: How the merging of movements left civil society out in the cold. *Global Environmental Politics*, 10(2), 11-17.
- Fomerand, J. (2009). *The A to Z of the United Nations*. Scarecrow Press.
- Forsch, R. A. and Gallopoulos, N. E. (1989). Strategies for manufacturing. *Scientific American*, 261(3), 144-153.
- Gibson, B., Hassan, S. and Tansey, J. (2013). *Sustainability assessment: criteria processes*. Routledge, United Kingdom.
- Goldsmith, E. (1972). *Blueprint for survival*. Penguin Books.
- Jordan, A. and Voisey, H. (1994). Institutions for global environmental change. *Global Environmental Change*, 1(8), 93-97.
- Kumar, S., Kumar, N. and Vivekadhish, S. (2016). Millennium development goals (MDGS) to sustainable development goals (SDGS): Addressing unfinished agenda and strengthening sustainable development and partnership. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, 41(1), 1-4.
- Malthus, T. R., Winch, D., and James, P. (1992). *Malthus: 'An Essay on the Principle of Population'*. Cambridge University Press.
- Manulak, M. W. (2015). Multilateral solutions to bilateral problems: The 1972 Stockholm conference and Canadian foreign environmental policy. *International Journal*, 70(1), 4-22.
- McCormick, J. (1986). The origins of the world conservation strategy. *Environmental Review: ER*, 10(3), 177-187.
- Mebratu, D. (1998). Sustainability and sustainable development: historical and conceptual review. *Environmental Impact Assessment Review*, 18(6), 493-520.
- Nordquist, M.H., Moore, J.N. and Mahmoudi, S. (eds). (2003) *The Stockholm declaration and law of the marine environment*. Martinus Nojhoff Publishers, 2003.
- Pezzoli, K. (1997). Sustainable Development: A Transdisciplinary Overview of the Literature. *Journal of Environmental Planning and Management*, 40(5), 549-574.
- Raskin, P., Banuri, T., Gallopin, G., Gutman, P., Hammond, A., Kates, R. and Swart, R. (2002). *Great transition: The promise and lure of the times ahead*. Stockholm Environmental Institute.
- Robinson, J. (2004). Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics*, 45(4), 369-384.
- Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The Lancet*, 379(9832), 2206-2211.
- Schechter, M. G. (2009). *United Nations global conferences*. Routledge, United Kingdom.
- Sharma, A., Schwarte, C., Bird, P., Michaelowa, A., and Müller, B., (2018). *COP24: key outcomes*. In: COP24, Katowice, 2-18 December 2018, Poland, p.1-19.
- Slim, H. (1995). What's development? *Development in Practice*, 5(2), 143-148.
- Sood, B. and Sooden, M. (2016). Sustainable Development: Evolution and Challenges Ahead. *Journal of Business Management and Information Systems*, 3(1) 86-96.
- Sullivan, D. J. (1993). Women's human rights and the 1993 World Conference on Human Rights. *American Journal of International Law*, 88(1), 152-167.
- Wackernagel, M. and Rees, W. (1998). *Our ecological footprint: reducing human impact on the earth*. New Society Publishers, Canada.
- Wall, D. and Ashford, P. (1994). Green History. *Journal of Environmental Psychology*, 14(3), 261
- WCED, R. (1987) Report of the World Commission on Environment and Development: *Our Common future*. <http://www.ask-force.org/web/Sustainability/Brundtland-Our-Common-Future-1987-2008.pdf>, Accessed Feb, 10, 2020.
- Worster, D. (1994). *The wealth of nature: environmental history and the ecological imagination*. Oxford University Press.

**Session 3.3 - Session spéciale 14 :
Supply chains of the future: Towards
greener and more local supply chains**

Chairman : Julien Boissière - Ramzi Hammami

Simulation of a full containerization approach in road transportation: preliminary results in the waste management industry

MILAD NOZARI¹, JEAN-FRANÇOIS AUDY², PASCAL FORGET¹ ET AMINA LAMGHARI²

¹ UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES, DÉPARTEMENT DE GÉNIE INDUSTRIEL
3351 BOUL. DES FORGES, TROIS-RIVIERES (QUEBEC), CANADA
milad.nozari@uqtr.ca; pascal.forget@uqtr.ca

² UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES, DÉPARTEMENT DE MANAGEMENT
3351 BOUL. DES FORGES, TROIS-RIVIERES (QUEBEC), CANADA
jean-francois.audy@uqtr.ca; amina.lamghari@uqtr.ca

Résumé – Cet article présente une analyse préliminaire de l'impact de la mise en œuvre de la conteneurisation intégrale dans le réseau logistique d'approvisionnement d'une entreprise en gestion des matières résiduelles du secteur CRD (construction, rénovation et démolition). En recourant à la simulation à événements discrets, divers indicateurs de performance sont évalués, dont la réduction du temps d'attente des camions lors de l'accès aux divers sites de l'entreprise. L'objectif de cette étude est d'améliorer l'efficacité du système d'approvisionnement et ainsi réduire les coûts de transport et de manutention. Effectués dans les centres de transbordement, à l'aide de deux types de camion opérant en amont et en aval du réseau d'approvisionnement, les processus de manutention et les contraintes d'espace sont identifiés comme les goulots générateurs de longues files d'attente. Les résultats préliminaires suggèrent que la mitigation de ces goulots par un transport conteneurisé de la matière, donc sans transbordement, mène à une réduction significative des temps d'attente total. De plus, les résultats montrent que le recours aux camions ayant un système de levage à crochet plutôt qu'un système de levage à treuil hydraulique facilite la conteneurisation intégrale et diminue la durée des opérations.

Abstract – This article presents a preliminary analysis of the impact of full containerization implementation in the supply logistics network of a waste material management company in the CRD (construction, renovation and demolition) sector. By using the discrete event simulation, various performance indicators such as waiting times of trucks at the entrances of various company sites are assessed. Ultimately, the objective of this study is to improve the efficiency of the supply system and thus reduce transport and handling costs. Done in collection centers between two types of trucks operating upstream and downstream of the supply network, handling processes and space constraints are identified as the bottlenecks generating long queues. The preliminary results suggest the mitigation of these bottlenecks by containerized transport of the material, therefore without transshipment, it leads to a significant reduction in the total waiting time. In addition, the results show that the use of trucks with a hook lifting system rather than a hydraulic winch lifting system facilitates full containerization and reduces the duration of operations.

Mots clés – Transport routier, simulation, conteneurisation, files d'attente, recyclage de matières résiduelles

Keywords – Road transportation, simulation, containerization, waiting queues, waste material recycling

1 INTRODUCTION

Waste generated by construction, renovation, and demolition activities, commonly known as CRD waste, is one of the most important solid waste streams, with a worldwide annual volume exceeding 10 billion tons (Wu et al., 2019). This waste stream typically contains wood, metal, glass, stone, rubble, etc. Although more than 50% of these materials can be recycled, landfilling remains the prevailing practice, mainly because recycling costs are higher than landfilling costs (Ulubeyli et al., 2017). For example, in the province of Quebec (Canada), more than 60% of the wood generated by CRD activities is used as landfilling (Recyc-Quebec, 2012). To reduce recycling costs, and thereby increase recycling rates, efficient logistics solutions are needed (Barton et al., 1996; Trochu et al., 2018).

A typical CRD industry's logistics chain is illustrated in Figure 1. It consists of a dispatch center, multiple collection centers (CC), and a sorting and processing plant (SPP). CRD waste is collected from customers' sites using a limited number of trucks owned by the recycling company, referred to as internal trucks. There are also external trucks, which do not belong to the recycling company, but which unload their waste materials at

the company collection centers. Truck drivers start their shifts from the dispatch center, where customer orders are received. Once the waste material is collected from a customer's site, it is transported by collection trucks either to the sorting and processing plant or to one of the collection centers. At these centers, material coming from different customers is piled up and then transferred to larger vehicles, namely trucks with moving floor trailers (TMFT). Figure 2 illustrates this process. The TMFTs then transport the collected material to the SPP. After unloading (whether at the SPP or a CC), a truck goes on to collect material from the next customer, identified by the dispatch center operator.

Due to the limited space and handling equipment capacity, long waiting queues are often formed at the entrance of the CCs or SPP, which results in longer delays in serving the next customers as well as higher costs for both the recycling company and external trucks. To reduce the negative impact of the long waiting queues, a new logistics approach is proposed by using a full containerization of the material transportation from customers to the SPP and thus exchanging full and empty containers upon arrival at the CCs or SPP.

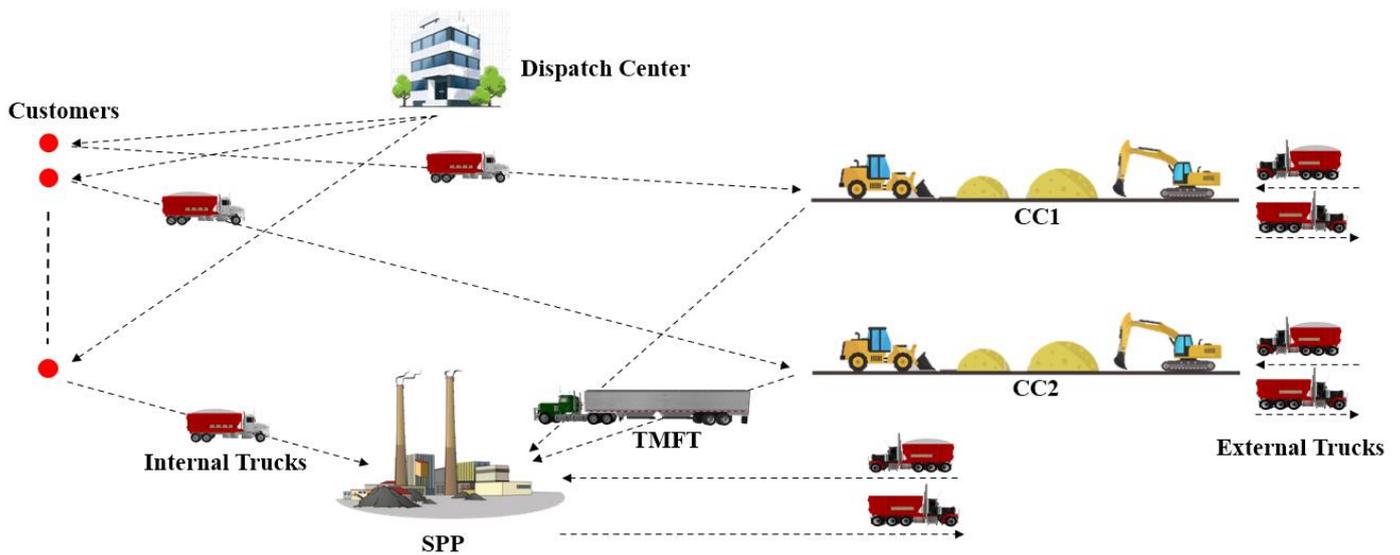


Figure 1. Typical logistics chain in the CRD industry

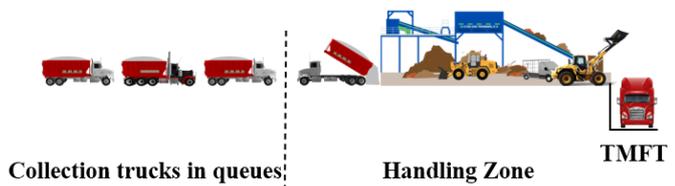


Figure 2. A typical collection center

Implementing this approach requires major modifications to the transportation fleet, more specifically:

- I. using a hook-lift technology instead of a hydraulic winch lifting system (cable roll-off) for containers handling;
- II. using road trains rather than TMFTs when transporting collected material from a CC to the SPP.

In this paper, a discrete-event simulation model is developed to assess the economic gains of the new logistics approach as well as its impact on reducing transit times.

The remainder of this paper is organized as follows. In Section 2, a literature review is provided, focusing on the recycling logistic and truck transportation optimization. The problem addressed in the paper is discussed in more details in Section 3. Section 4 describes data collection and Section 5 presents the simulation model. Simulation results follow in Section 6, and a conclusion is presented in Section 7.

2 LITERATURE REVIEW

In this study, the impact of containerization in reducing trucks waiting times in entrance queues of several sites of a waste recovering company is analyzed. The literature review of three fields of research is presented in this section. Firstly, an overview of studies conducted in CRD waste management will be investigated. Secondly, articles focusing on the reduction of waiting times in queues by improving the efficiency of transport or handling operations will be studied. At last, the concept of containerization will be investigated.

2.1 CRD waste management

So far, a significant body of research has been done in the area of CRD waste management. Asefi et al. (2020) have classified them into two main categories: system assessment models (SAM) and system engineering models (SEM). SAM mainly

analyzes waste management systems, material flow analysis, risk assessment techniques, and life cycle analysis. On the other hand, SEM involves the design of a waste management system and includes multi-criteria decision making (MCDM), system dynamics (SD), forecasting and statistic-based models, simulation-based approaches, cost-benefit analysis (CBA), and optimization models (Asefi et al., 2020). This study belongs to the second type.

Beliën et al. (2012) conducted a review of studies on the municipal solid waste collection and management problem. This study examined the classification of waste types (garbage, dumpsters/containers, hazardous materials, recyclables, and others), scope (case study, installation, truck allocation, and collection routing), resolution methods (programming mathematics, metaheuristics, other heuristics, and other methods), the stochastic or deterministic solution approach, the objective (number of trucks, costs, environmental impacts, number of employees, length of the route, collection time and others factors), etc. The focus of most of these papers is on presenting different mathematical computation and analysis methods, including vehicle routing problems in the CRD material supply chain. According to our research, the number of papers focusing on optimizing the waiting time at the entrances of collection or recycling sites is very low. Akpan & Akpofure (2013) are one of the few that examine and optimize the truck waiting times at CCs in the waste management industry.

2.2 Reduction of truck waiting times in queues

Studies that examine truck waiting times in queues are in the context of intermodal terminals such as seaports, cross-docking warehouses, or rail-road terminals. The study by Fleming et al. (2013) was carried out with an agent-based simulation to model the entrance gate of a seaport to reduce the waiting time for trucks in queues. The results indicate that using a pooled queue lowers significantly average wait times and variability in wait times. Briesemeister & Novaes (2017) aimed at minimizing the waiting times of trucks in a cross-dock terminal, the simulation of the proposed approach represented satisfying results. In the research of Lima et al. (2015), which focuses on the control of transport flows at the intermodal terminal in port logistics systems, a transport planning method carried out by an agent at the intermodal terminal was taken into account. Mokshin et al. (2019) worked on the optimization and simulation of queuing

systems from terminals to seaports to minimize total costs by reducing service time and waiting time in queues.

Wang et al. (2018) studied an entire logistics network (rail-road) by a bi-objective mixed-integer linear programming optimization via a genetic algorithm to minimize the maximum time between each origin-destination pair which is the sum of travel time required at intermodal hubs. Wang & Zhu (2019) studied the minimization of total handling time in a rail-road intermodal hub via genetic algorithm optimization. Yamada et al. (2003) proposed a mathematical model using queuing theory to minimize container handling costs by optimizing handling capacity in seaports.

Zhao & Goodchild (2010) studied the impact of truck arrival information on reducing handling work by improving terminal to seaport operations via simulation. Zhang et al. (2019) used a heuristic approach of the queuing model to optimize an appointment system to reduce waiting times for trucks outside the gates and wait times for internal trucks in the yard of a seaport. Dekker et al. (2013) proposed to exchange the chassis to avoid additional handling of containers in port terminals and studied by simulation its effects on waiting time and the environment. Zhao & Goodchild (2013) used a hybrid simulation and queue theory approach to reduce container handling and truck processing times.

2.3 Containerization

Kuznetsov et al. (2017) suggested containerization of bulk cargos in seaports, which is an effective method for the transportation and handling of general cargo, as the transportation of bulk cargo by rail and road modes causes fairly significant losses to the environment due to the production of dust. Matsuda et al. (2020) have proposed containerization of bulk cargoes, as it can address imbalances in the amount of cargo carried between the main and backhauls, thereby improving efficiency. They also identified the key factors in the choice of transport modes. Their study aims to clarify the cost competitiveness of container transport and to identify cost reductions likely to increase the use of containerization of bulk goods.

Clott et al. (2015) assessed the impact and challenges of increasing containerized movements of agricultural products from hinterland points to overseas markets, in terms of both operations and governance. They looked at several major operational issues, actions, rules, and policies affecting containerized flow and its total cost. Rodrigue & Notteboom (2014) studied the gain of the agricultural sector and its product chains in the speed, ubiquity, and flexibility of containerized freight distribution. They also discussed how containerization facilitates the management of individual freight shipments. Lirn & Wong (2013) studied the factors influencing the choice of mode of transport between bulk carriers and containers.

To the best of our knowledge, no research has been conducted on the implementation of full containerization in the field of the CRD waste management industry. This study aims to analyze containerization as a new logistic approach to reduce truck waiting times in queues of several sites of a waste recovering company and investigate its impact on logistics cost reduction and waiting reduction.

3 PROBLEM DESCRIPTION

3.1 Processes of the transport system

In this paper, the logistics chain of a company that recycles CRD materials is investigated. All activities are based on a task

list that contains customer orders and their priority. A fleet of internal trucks is responsible for transporting empty containers from the dispatch center or CCs to customer sites, as well as transporting full containers from customer sites to CCs or the SPP. In addition to internal trucks, there are also external trucks whose arrivals at CCs or the SPP are beyond the control of the company dispatchers. TMFTs are outsourced and in charge of transporting the waste materials from CCs to SPP, and their arrival times are based on a timetable that is predetermined by the company. Logistics is based on partial containerization for the transportation of materials between customer sites and CCs, and transportation and handling costs represent a significant portion of the total costs.

Once a customer container is ready to be collected, an internal truck ensures the delivery of an empty container to the customer and the transport of the full container to a CC, where each incoming truck unloads and leaves with an empty container to the next customer. Handling equipment is in charge of loading the material into a TMFT to be sent from the CC to the SPP. During the day, there are regular and repeated queues when the truck arrival rate exceeds the handling capacity of the CC. A similar observation is valid for the SPP. The main problem with this system is the long waiting queues at the entrance of CCs and SPP, where trucks must spend a significant part of their work time, which result in additional costs per year.

3.2 Customer orders

Based on customer calls, the company builds a daily list of tasks to be performed by a dedicated truck fleet according to customer priority as well as call time. A transport task consists of a sequence of three sites, involving the movement of an empty and/ or a full container. There are three types of tasks. Task type 1 is a task whereby the truck goes to the customer's site carrying an empty container, then after dropping off the empty container, it comes back to the CC to take a new empty container and start a new task. This type of task mostly occurs when it is the first visit of the customer's site and there are no full containers at the site. Task type 2 is a task by which the truck goes to the customer's site carrying an empty container, then after delivery of it, the truck will lift a full container and move towards CC or SPP. Most tasks are of this type where the container at the site is full and the customer needs an empty container. Task type 3 is a task by which the truck goes to the customer's site without carrying a container. After having lifted a full container, it heads for CC or SPP. Most of this type of task occurs when a customer no longer needs an empty container.

3.3 Full containerization as a new logistics approach

A promising logistic approach to improve the actual system is the full containerization from all customers' sites to the SPP. More explicitly, once the collection trucks arrive at the CCs, they move their full containers in the exchange area, lift an empty container, and head to a new customer. Then, another truck type called a road train (RT), capable of transporting two containers and belonging to the company, moves the containers to the SPP. This way, there would be no need for TMFTs and their required handling operations. Moreover, this study examines the use of hook-lift trucks instead of current container trucks which operate based on a hydraulic winch lifting system (roll-off cable) and are the common system in North America. Figure 3 represents the material and vehicle flow in current and proposed approaches.

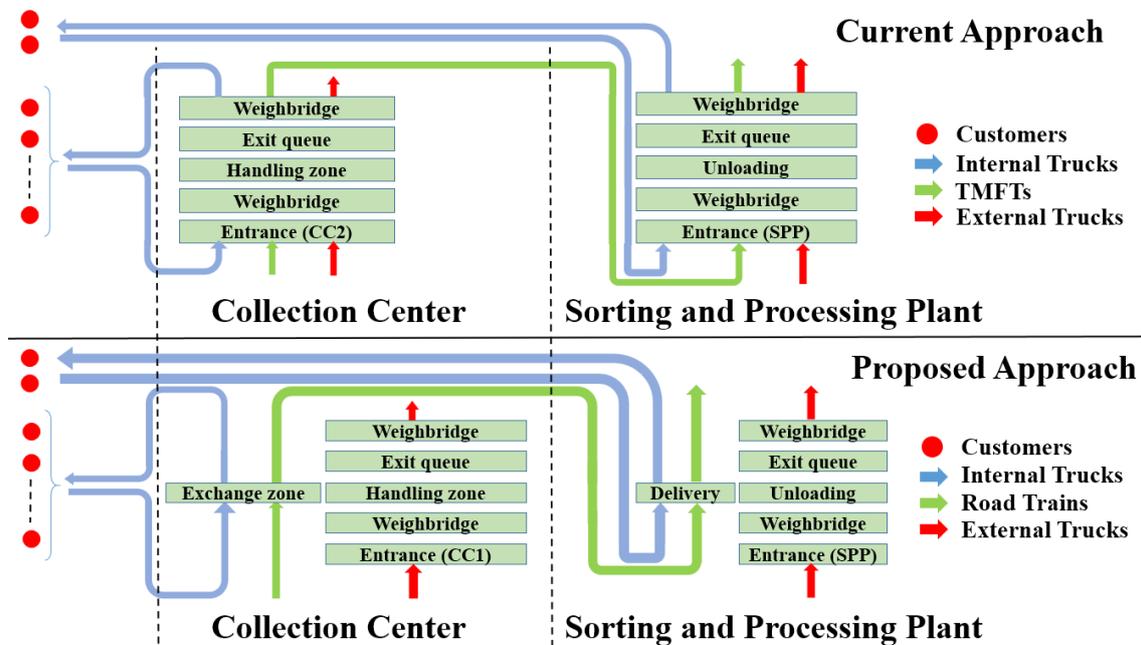


Figure 3. Material and vehicle flow in current and proposed approaches

4. DATA COLLECTION AND ANALYSIS

To generate a simulation model that adequately represents reality, various input parameters are adjusted based on historical data in model development. This historical data is collected using the information provided by dispatchers. Working day starts at 5:30 a.m. and at this time, collection trucks leave the dispatch center and go to the customers' sites. At the end of the shift, the collection trucks come back to the dispatch center before 3:30 p.m. to avoid traffic jam. To determine the truck arrival rate, the queues length, and the waiting time, historical data of two working days from 5 a.m. to 4 p.m. was collected by field observation. Subsequently, the data collected from customer sites, CC1, CC2, and SPP were verified. Using the Kolmogorov-Smirnov method, the samples were fitted to appropriate distributions of each parameter, and the characteristics of these distributions were determined.

5 SIMULATION

The simulation model for this study was developed by using SIMIO simulation software. Google Earth was used to plot landmarks and travel paths between customers and the company sites. Arrival rate, processing time, and load and unload distributions adjusted to historical data were used in the development of the model. 86 customers in the Greater Montreal region as well as the locations of the dispatch center, collection center 1 (CC1), collection center 2 (CC2), and SPP are considered as they are shown in Figure 4.

The model follows the order of the tasks on the list when allocating the next free truck to a task. The simulation begins when the site opens by allocating the first truck to the first task on the list, then the second truck to the second task and so on, until the task list is complete.

After developing the simulation model, the input data are entered. Statistical distributions adjusted to historical data are considered for the times of operations in the reference scenario.

Regarding the input parameters of the two scenarios, as presented in Table 1, company experts were consulted. In order to preserve the confidentiality of the company's financial information, the related data is given in the form of a coefficient of value X.

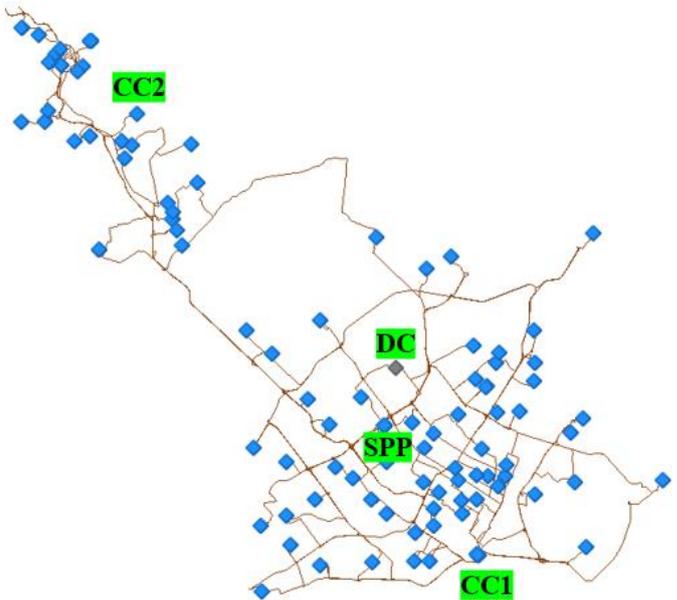


Figure 4. Distribution of customers' sites

- The hypotheses that were considered in the simulation are:
- the quantity of empty containers at the sites is infinite;
 - the capacity of the sites for the storage of materials is infinite;
 - the capacity of all containers is the same;
 - all containers are fully loaded;
 - there is no failure of resources;
 - there is no resource break.

Table 1. Input data of two scenarios

	Item	Reference scenario	Alternative scenario
1	Truck operating time at customers' sites	Beta (average: 7.92 min)	Triangular (3, 4, 5) min
2	Passage time on weighbridge	Weibull (average: 3.6 min)	Weibull (average: 3.6 min)
3	Unloading time by internal and external trucks	Lognormal (average: 7.3 min)	Triangular (1.5, 2, 2.5 min)
4	Loading time by TMFTs/ RTs	Gamma (average: 45.6 min)	Triangular (7, 8, 9 min)
5	Unloading time by TMFTs/ RTs	Gamma (average: 20.82 min)	Triangular (7, 8, 9 min)
6	Equivalent capacity of the handling zone	5	5
7	Transport costs of Internal Trucks	Roll-off : X	Hook-lift : X
8	Transport costs of TMFTs/ RTs	2.1 X	1.25 X
9	Handling costs at each site	2.5 X	2.5 X

After implementing the reference scenario in the simulation software and entered the various input parameters, a discrete-event simulation was run for a period of 11 hours (5 a.m. to 4 p.m.). Since in this study, the average waiting time and average service time are calculated from the start to the end of the day, no warm-up period is taken into account. Also, 50 replications were considered to obtain results.

To verify the accuracy of the model, a validation and verification phase was carried out. The first validation was done by making sure that the right number of entities were created and that these entities took the correct path in the model. The arrival times of the different types of trucks at the site entrances were validated by recording the number of trucks arrivals at the different time slots with the interval of one hour. Besides, the average service times of different trucks at the sites were compared with those collected in the field and the results were shared with the company operation managers.

In addition to the reference and alternative scenarios, four transition scenarios are defined to monitor the evolution of waiting time and cost reduction during the transition period of

implementing 20%, 40%, 60%, and 80% of the proposed approach.

6 PRELIMINARY RESULTS

This section presents the simulation results of the reference scenario, the alternative scenario, and four transition scenarios to compare their performance. The simulation results of the alternative scenario show that by implementing the proposed approach, the average service time at customers' sites has been reduced up to 58%.

Regarding queues at entrances of company sites, as shown in Figure 5, in the peak hour of the reference scenario, the number of trucks in these queues even reaches 18. However, the results of the alternative scenario simulation indicate that the average waiting times of trucks at these queues have a reduction of 92 to 95%. In other words, the queues at the entrances of the CCs and the SPP have almost disappeared for internal trucks so that the maximum number of trucks in the queue is not greater than 4. This reduction in waiting time reduced also the average service time by approximately 90% at the company's three sites.

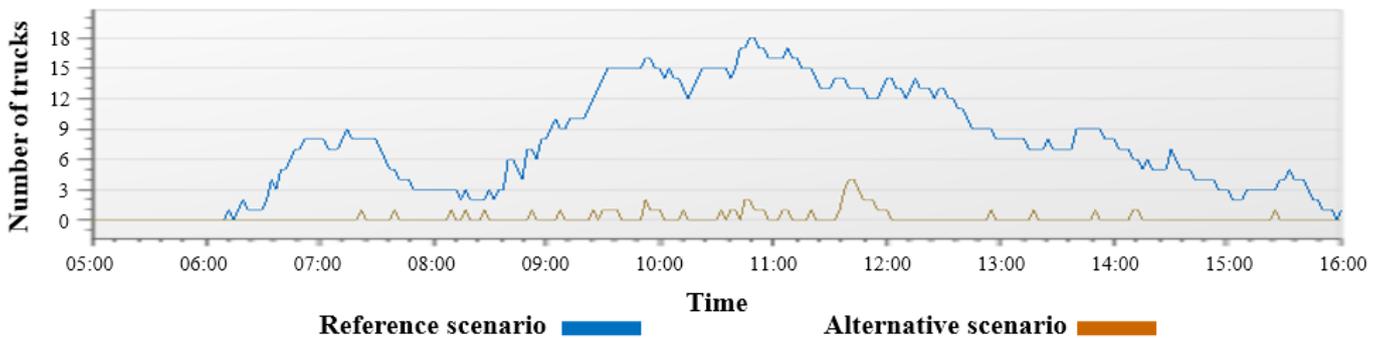


Figure 5. Number of trucks in queues over the working day

Since the transition from the current approach to the proposed approach is not possible at once and it takes a transition period of several years, an evaluation of the reduction in waiting time and the associated costs is achieved by defining four extra scenarios in which, different levels of the proposed system are implemented. These scenarios respectively represent the implementation of 20, 40, 60, and 80% of the new approach.

As Figure 6 represents the simulation results of scenarios in terms of time, the reduction in waiting time in queues compared to the percentage of implementation of the new approach is not linear. More precisely according to the simulation results of scenario 1, by replacing 20% of the current fleet with the proposed one, an average reduction of 34% in waiting time will be achieved. This reduction level is respectively 67%, 77%, 85% and 91% for scenarios 2, 3, 4 and alternative scenario.

These results indicate especially the importance of implementing the initial 40% of the new system.

Since in the current system, the trucks spend a lot of time in queues and this leads to important costs, the use of the alternative scenario will significantly reduce transportation costs as well as the handling costs because of the need to transfer materials from internal trucks to the TMFTs by handling equipment is eliminated. Figure 7 shows the impact of the proposed approach on cost reduction of the operations at customers' sites, waiting times in the queues, transports from CCs to SPP, and handling operations. In the alternative scenario, since each external truck unload their container in the exchange zone and this volume have to be reloaded in containers to be delivered by a RT, there is still the need for the presence of handling equipment.

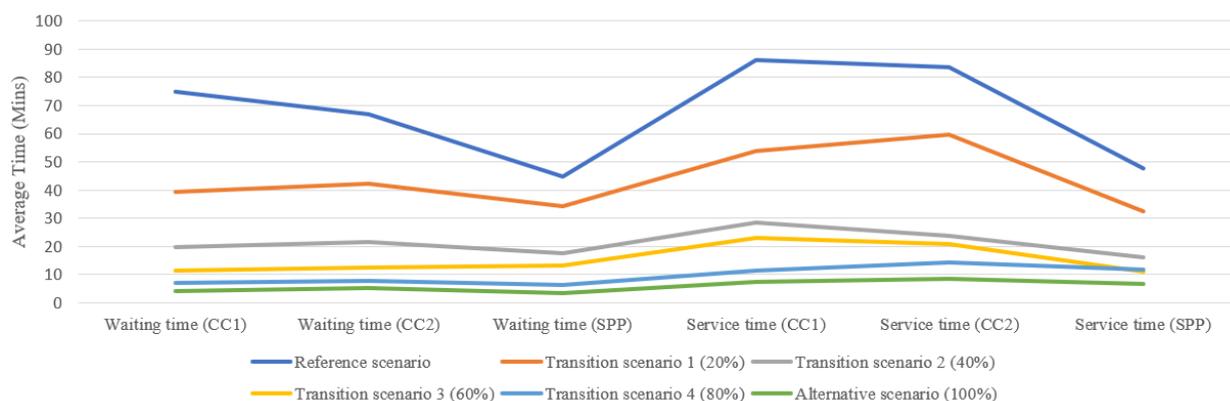


Figure 6. Average truck waiting time and average service times per site

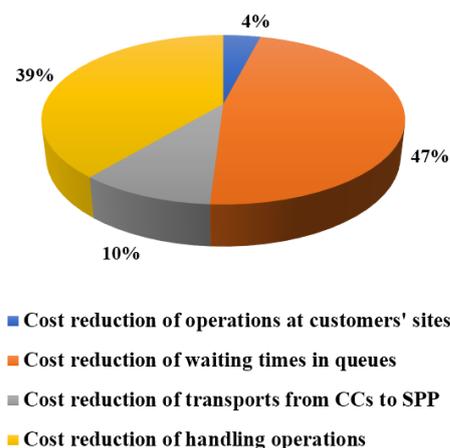


Figure 7. Share of cost reduction in different fields

There are still handling costs, but they are much lower. Taking into account the costs associated with the transport and handling of the waste materials, the cost is calculated based on the reduction of waiting times and handling operations achieved by

simulating the defined scenarios. Table 2 presents the procedure of this calculation. The calculations are based on the simulation results in terms of time reduction and the input data given in Table 1. The results show that by benefiting from the new fleet of trucks, we will have a reduction of 29.35% in transport and handling costs.

Assuming that the company internal truck fleet has a useful life of around 10 years and each year the company renews 10% of its fleet, our simulation model shows that if the company replaces 10% of its old fleet by the proposed one each year, at the end of the fourth year there will be a reduction of 17.32% in costs, in other words, 59% of the total cost reduction. Figure 8 shows the reduction in average waiting time for internal trucks as well as the reduction in costs over a 10-year transition period, following the replacement of 10% of the fleet per year. However, a universal container's attachment point is required to enable both lifting systems. This will necessarily incur additional costs during the transition period, e.g., at the purchasing of a new container once a unit reaches the end of its normal lifetime or, for the more recent units, to include in their maintenance plan a task to modify their attachment point.

Table 2. Simulation results compared to the reference scenario

	Indicators	Transition scenario 1 (20%)	Transition scenario 2 (40%)	Transition scenario 3 (60%)	Transition scenario 4 (80%)	Alternative scenario
1	Reduction in total trucks operation time at customers' sites (hrs)	1.105	2.18	3.715	5.4	7.3
2	Percent of reduction in trucks operations cost at customers' sites	11.85%	23.38%	39.84%	57.91%	78.29%
3	Reduction in average waiting time for trucks at CT1 (hrs)	0.59	0.92	1.06	1.13	1.18
4	Reduction in average waiting time for trucks at CT2 (hrs)	0.41	0.76	0.90	0.98	1.03
5	Reduction in average waiting time for trucks at SPP (hrs)	0.17	0.45	0.52	0.64	0.68
6	Average reduction in total waiting time (hrs)	1.17	2.13	2.49	2.75	2.90
7	Average waiting costs reduction per day (CAD)	1757	3197.75	3729.50	4130	4346
8	Number of TMFTs-RTs for the transport of the same quantity	13-6	10-12	7-18	3-26	0-32
9	Cost reduction by using RTs	75	150	225	325	400
10	Usage time of handling equipment at each site	7	5,5	4	3	2
11	Reduction in handling costs at all sites (CAD)	900	1800	2700	3000	3600
12	Percentage of expected cost reduction	9.17%	17.32%	22.67%	25.90%	29.35%

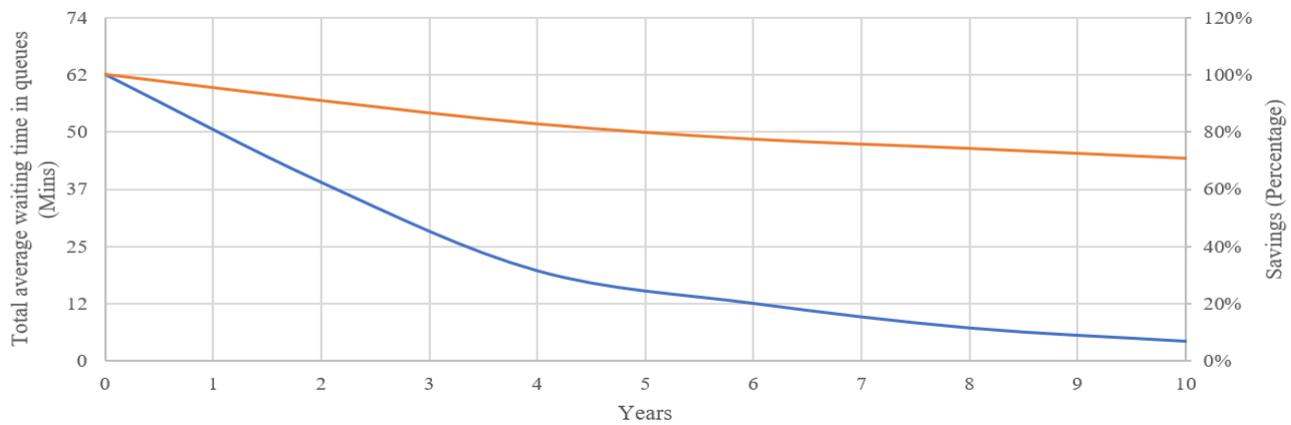


Figure 8. Evolution of total waiting time and costs reduction in a 10-year transition plan

7 CONCLUSION

In this paper, full containerization has been proposed as a different logistic approach in the CRD material recovery industry and has been compared to the conventional transportation approach in this industry. Different aspects of using full containerization have been investigated using discrete-event simulation, including waiting time of trucks in queues of entrances and transportation and handling costs. Also, a new type of truck that facilitates full containerization has been simulated to analyze its impact on cost reduction. The results of the simulation show a significant difference in terms of waiting time and operation costs. At last, an implementation strategy taking into account the useful life of the truck fleet is proposed.

The preliminary results presented in this article are based on the historical data of one single day. Since the number of customers' orders and thus the amount of material transported may vary depending on the season, climate, and regional market conditions, more numerical testing will be needed to support a ten-year earnings prediction. The presented studies, though, deserve to explore part of the theoretical path to follow. Complementary studies are going on to ensure the reliability of the proposed approach, especially in the handling operations. Also, the role of full containerization in the reduction of greenhouse gas emissions should be analyzed to assess the environmental advantage of the logistic approach proposed.

9 REFERENCES

- Akpan, N., & Akpofure. (2013). Queuing Theory and its Application in Waste Management Authority (A Focus on Lawma Igando Dump Site, Lagos State). *4*, 4-438
- Asefi, Y., Gohari Mahmoudabad, A., Habibian Sezavar, A., Mirshahvaladi, S., Abyadeh, M., & Abyareh, M. (2020). Association between maternal cadmium exposure and preterm birth: a meta-analysis. *International Journal of Environmental Health Research*, 1-10.
- Beliën, J., De Boeck, L., & Van Ackere, J. (2012). Municipal Solid Waste Collection and Management Problems: A Literature Review. *Transportation Science*, *48*
- Briesemeister, R., & Novaes, A. G. N. (2017). Comparing an Approximate Queuing Approach with Simulation for the Solution of a Cross-Docking Problem. *Journal of Applied Mathematics*, *2017*, 4987127
- Clott, C., Hartman, B. C., Ogard, E., & Gatto, A. (2015). Container repositioning and agricultural commodities: Shipping soybeans by container from US hinterland to overseas markets. *Research in Transportation Business & Management*, *14*, 56-65
- Dekker, R., Heide, S. v. d., Asperen, E. V., & Ypsilantis, P. (2013). A chassis exchange terminal to reduce truck congestion at container terminals. *Flexible Services and Manufacturing Journal*, *25*, 528-542
- Fleming, M., Huynh, N., & Xie, Y. (2013). Agent-Based Simulation Tool for Evaluating Pooled Queue Performance at Marine Container Terminals. *Transportation Research Record: Journal of the Transportation Research Board*
- Kuznetsov, A., Kirichenko, A., & Slitsan, A. (2017). Simulation for assessment of bulk cargo berths number. *IOP Conference Series: Earth and Environmental Science*, *87*
- Lima, A. D. P., Werner de Mascarenhas, F., & Frazzon, E. M. (2015). Simulation-Based Planning and Control of Transport Flows in Port Logistic Systems. *Mathematical Problems in Engineering*, *2015*, 862635
- Lirn, T.-C., & Wong, R.-D. (2013). Determinants of grain shippers' and importers' freight transport choice behaviour. *Production Planning & Control*, *24*(7), 575-588
- Matsuda, T., Hanaoka, S., & Kawasaki, T. (2020). Cost analysis of bulk cargo containerization. *Maritime Policy & Management*, *47*(6), 736-755
- Mokshin, V., Kirpichnikov, A., & Soiko, A. (2019). Simulation and optimization of the cargo terminal in the AnyLogic environment. *Journal of Physics: Conference Series*, *1368*, 042082
- Rodrigue, J.-P., & Notteboom, T. (2014). Looking Inside the Box: Evidence from the Containerization of Commodities and the Cold Chain. *Maritime Policy & Management*, *42*
- Wang, L., & Zhu, X. (2019). Container Loading Optimization in Rail-Truck Intermodal Terminals Considering Energy Consumption. *Sustainability*, *11*, 2383
- Wang, R., Yang, K., Yang, L., & Gao, Z. (2018). Modeling and optimization of a road-rail intermodal transport system under uncertain information. *Engineering Applications of Artificial Intelligence*, *72*, 423-436
- Yamada, T., Professor, A., Yoshizawa, G., Bona, R., & Mori, K. (2003). Optimizing the handling capacity in a container terminal for investigating efficient handling systems.
- Zhang, X., Zeng, Q., & Yang, Z. (2019). Optimization of truck appointments in container terminals. *Maritime Economics & Logistics*, *21*(1), 125-145
- Zhao, W., & Goodchild, A. V. (2010). The impact of truck arrival information on container terminal rehandling. *Transportation Research Part E: Logistics and Transportation Review*, *46*(3), 327-343
- Zhao, W., & Goodchild, A. V. (2013). Using the truck appointment system to improve yard efficiency in container terminals. *Maritime Economics & Logistics*, *15*(1), 101-119

(This paper is not included in this version, please contact authors for more information)

CIGI 2021

Analysis of Price and Production Configuration in the Olive Oil Industry with organic and conventional products: The French Case

IMEN NOUIRA¹, ALINA FERNANDEZ-ARIAS¹, NATACHA GONDRAN², RAMZI HAMMAMI¹

¹ RENNES SCHOOL OF BUSINESS

2 rue Robert d'Arbrissel, CS76522, 35065 Rennes CEDEX, France
{imen.nouira,alina.fernandez-arias,ramzi.hammami}@rennes-sb.com

² ECOLE NATIONALE SUPÉRIEURE DES MINES DE SAINT ÉTIENNE

158 Cours Fauriel, 42023 Saint-Étienne, France
natacha.gondran@mines-stetienne.fr

Résumé – Dans cet article, nous développons un modèle de conception de chaîne logistique adapté au secteur d'huile d'olive, dans lequel nous intégrons les décisions de localisation et de prix. Nous considérons deux variétés du produit : une variété biologique et une variété conventionnelle. Nous proposons un modèle mathématique dans lequel nous intégrons les principales décisions et contraintes du problème. Le modèle proposé étant non linéaire, nous développons une version linéarisée que nous résolvons à l'aide de CPLEX. Nous appliquons le modèle au cas de la France et utilisons des données réelles pour le marché français d'huile d'olive. Les résultats obtenus montrent que notre approche améliore le profit de 42.96% en moyenne par rapport au cas où les prix de vente sont fixés.

Abstract – In this paper, we develop a supply chain design model that is adapted to the olive oil sector and includes facility location and pricing decisions. We consider the production of organic and conventional products. We propose a mathematical model in which we consider the main decisions and constraints that are relevant to the olive oil production. Since the proposed model is nonlinear, we develop a linearized version that we solve using CPLEX. We analyse the French case and use real data from the French olive oil market. The obtained results show that our approach improves the profit by 42.96% on average with comparison to the case in which the selling prices are fixed.

Mots clés – Configuration de la chaîne logistique, Industrie d'huile d'olive, Demande endogène, Huile d'olive bio.

Keywords – Supply chain configuration, Olive oil industry, Endogenous demand, Organic olive oil.

1 INTRODUCTION

The consumption of the olive oil has historically been confined to the Mediterranean region. But recent generic and commercial promotion efforts have stimulated interest in the consumption of olive oil and created new opportunities for Mediterranean producers to export to non-traditional markets. The average world consumption has almost doubled over the past 25 years. Consequently, consumption that was traditionally restricted to the Mediterranean area (mainly in Spain, Greece, and Italy) is currently increasing in non-producing countries or emerging markets (especially USA, Canada, Australia, China, Japan, Argentina, Brazil, Chile, and Mexico) (Karanikolas et al., 2018; Roselli et al., 2016; Sayadi et al., 2017). Historically destined for local market with a short distribution channel, new market opportunities have generated new challenges in term of logistics for the olive oil industry. This increasing demand of olive oil in new European markets and the increasing demand for organic products, emphasize the need to rethink the olive oil Supply Chain (SC) design.

Olive oil production and in general, agri-food industry are characterized by a specific market structure since the market is currently segmented into ecofriendly customers that prefer

to buy organic products despite the higher price, and traditional (non-ecofriendly) customers that are more sensitive to price and, thus, prefer to buy the non-organic products as they are often cheaper. Many authors have been reported the impact of the organic olive oil labelling in the customers preferences. For instance, Vlontzos and Duquenne (2014) showed that Greek consumers accept payment premiums for organic oils and Palacios-Argüello et al., (2019) report an increment on the number of customer that are switching to organic products in the French agrifood sector. The attributes that most influence olive oil demand are the final price, the oil origin, and the organic certification (Yangui et al. 2014, Romo-Muñoz et al. 2015, Cacchiarelli et al. 2016, Boncinelli et al. 2016).

In this paper, we analyse price and production configuration for olive oil production. We consider a single manufacturer that must decide to produce conventional and/or organic olive oil and, consequently, select the location and the processing capacity of the facility. The selling price is different for organic and conventional olive oil and demand in each segment is sensitive to the selling price.

Our study case is centered in the French market. In 2019, the olive oil consumption in France was approximate 108,000

tons and only 5,500 tons were produced in France (Afidol, 2019). This higher demand is attracting the Mediterranean olive oil producers to the French market.

This paper is organized as follows: in section 2 are summarized some related works. In section 3, we formulate the problem studied in this paper. The mathematical model and the solution strategy are discussed in section 4. The French case is introduced in section 5. We finally conclude and discuss future research directions.

2 LITERATURE REVIEW

The supply chain in olive oil industry consists of many phases that include olive fruits extraction, processing, packaging, and distribution to consumers. Each of these phases is a complex process in which several decisions must be made. For example, there are two sources for the fruits: it can be purchased in the spot market after the recollecting season or the trees can be leased during the growing season (Kazaz and Webster, 2011). Another example is the packaging process, the olive oil can be selling in bulk or in small containers. At the same time, the location of the supplier, the facility, and the customer determine the effective cost incurred in purchasing, processing and distribution process (IOC, 2015).

This problem has already received attention in the literature. Kazaz (2004) focused on olive oil production and investigated production planning decisions under yield and demand uncertainty. Kazaz (2008) addressed the problem of determining the sale price and the production quantity under supply uncertainty. The problem is commonly observed in the agricultural industry. Kazaz and Webster (2011) proposed a single-period production planning model that maximizes the expected return by determining the optimal amount of space to be leased for production and the quantity of olives to be provided from external sources under yield and yield-dependent cost uncertainties in a Turkish olive oil supply chain. None of these papers deal with the design of SC (in particular, production facility location) or consider the production selection of organic and/or conventional varieties.

As mentioned, price and organic certification are two of the most influence parameters in the olive oil demand. According to Yangui et al. (2014), the most important attributes affecting consumers' preferences towards olive oil are the price and the product's origin. Balcco and Gracia (2020) indicated that the most important olive oil attributes for consumers, are price, origin of production, local production, and territory, and the Protected Designation of Origin (PDO) quality certification. Scalco et al. (2020) studied the organic labeling products in the olive oil sector. Sazvar et al. (2018) have observed a switching from conventional products towards organic products thanks to better communication about organic product benefits for the environment and the consumers' health.

Several authors have been suggested that the organic certification adds value to the olive oil (Roselli et al. 2016, 2018, Liberatore et al. 2017, Boncinelli et al. 2016, Cacchiarelli et al. 2016, Cavallo and Piqueras-Fiszman 2017, Del Giudice et al. 2015, Vlontzos and Duquenne 2014). In particular, Bonti-Ankomah and Yiridoe (2006) reports that most consumers are not willing to pay a price premium higher than 10-20%. All these studies justify the interest to include the organic certification as a decision to the supply chain. And the importance to consider an endogenous demand that depends on the price and organic certification.

In this paper we present a first attempt to present a model with demand sensitivity to the type of olive oil (organic or conventional).

3 PROBLEM DEFINITION

In this paper, we study the price and production decisions in the olive oil industry. We consider a single manufacturer that must decide to produce conventional and/or organic olive oil and, consequently, select the location and the processing capacity of the facility.

The olive oil market is segmented into organic customers and conventional customers. For each market segment, the effective demand depends on the price. There is a pair of base demand and base price associated with each variety of olive oil. The base price represents the standard market price. It may represent the competitors' price. We let p_b^o and D_b^o (respectively, p_b^n and D_b^n) denote the pair of base price and base demand of organic variety (respectively, conventional variety).

We consider the effective price of each variety of olive oil as a decision variable. We let p^o and p^n represent the effective price offered by the producer for organic product and conventional product, respectively. As usual in the operations management literature, we assume that demand linearly decreases in price. We let α^o and α^n respectively denote the price-sensitivity of organic and conventional customers. Typically, we have $\alpha^o \leq \alpha^n$, which means that organic customers are less sensitive to price compared to conventional customers. Thus, the effective demand of each olive oil variety is obtained as follows.

$$d_i^o = D_b^o - \alpha^o (p^o - p_b^o) \quad i = \{1, 2, \dots, 5\} \quad (1)$$

$$d_i^n = D_b^n - \alpha^n (p^n - p_b^n) \quad i = \{1, 2, \dots, 5\} \quad (2)$$

Another important decision is the location of the facility. There is a set of potential locations J indexed by j . Each location j is characterized by a production (processing) cost per Kg of final product (m_j), a distribution cost per Kg of final product from facility j to the distribution center (h_j), the purchasing cost per Kg of organic (v_j^o) and conventional olives (v_j^n), and an annual amortized facility opening cost (F_j). The aim of this paper is not to go in the technical details related to the production process. At our low technical details level, it can be assumed that the processing cost m_j does not depend on the variety of olive oil (organic or traditional).

The producer must also decide the total extraction capacity to implement in the selected facility. We let T denote the set of available olive oil extraction machines, indexed by t . Each machine is characterized by its capacity Q^t , and the annual amortized acquisition cost τ^t .

All the elements described above are common to both olive oil varieties. The production of organic olive oil must satisfy the following conditions:

The olive oil is considered as organic only when the olives used are organic. Notice that some countries accept up to a certain percentage of conventional olives in the composition of organic varieties (for instance, up to 5%). This rate may differ across countries. As legislations are becoming increasingly tighter, we consider in this study that 100% of the quantity of olives used to extract the organic olive oil variety must be certified as organic. Our model can be easily adjusted to consider other composition rates.

The transportation and conditioning of organic olives must satisfy specific requirements (strict cleaning conditions, small quantity of olives per container, etc.) which increases

the cost per Kg of olives. This extra cost can be imputed to the purchasing cost.

The same type of extraction machine can be used for both varieties as explained earlier. However, the organic olive oil requires specific parametrization of the extraction machine (temperature, pressure, etc.) which induces less productivity. According to our industrial partners, we approximately need $R^o = 5$ Kg of organic olives to obtain 1 Kg of organic olive oil, whereas only $R^n = 2.5$ Kg of olives are required to obtain 1 Kg of conventional olive oil.

Each facility is linked with a supplier zone. And, as we mention before, the purchasing price depends on the facility location and the variety (organic or conventional) of the olives.

4 MATHEMATICAL MODEL

This section presents the model formulation. We first introduce the notation used for the parameters and decision variables (some notations have already been given but will be recalled here).

Demand and prices:

- D_b^o : base demand of organic olive oil variety,
- D_b^n : base demand of conventional (non-organic) olive oil variety,
- p_b^o : base price of the organic product in the market,
- p_b^n : base price of the conventional product in the market.

Facilities:

- J : set of all potential facility locations,
- m_j : production cost at facility $j \in J$ per Kg of final product,
- h_j : distribution cost from facility $j \in J$ to the demand zone per Kg of final product,
- v_j^o : purchasing cost per Kg for the organic olives by facility $j \in J$,
- v_j^n : purchasing cost per Kg for the conventional olives by facility $j \in J$,
- F_j : annual amortized opening cost of facility $j \in J$.

Olive oil extraction machines

- T : set of potential extraction (mill) machines,
- Q^t : input capacity of machine $t \in T$,
- τ^t : annual amortized acquisition cost of one unit of machine $t \in T$,
- R^o : quantity of organic olives required to obtain one unit (Kg) of organic olive oil,
- R^n : quantity of conventional olives required to obtain one unit (Kg) of organic olive oil.

Decision variables:

- $y_j \in \{0,1\}$: equals 1 if facility $j \in J$ is selected; 0 otherwise,
- $p^o \in \mathbb{R}^+$: price of organic olive oil (per Kg),
- $p^n \in \mathbb{R}^+$: price of conventional olive oil (per Kg),
- $x_j^o \in \mathbb{N}$: quantity of organic olive oil produced in facility $j \in J$,
- $x_j^n \in \mathbb{N}$: quantity of conventional olive oil produced in facility $j \in J$,
- $\mu_j^t \in \mathbb{N}$: quantity of machines of type $t \in T$ installed in facility $j \in J$.

Using the above notation, we formulate the problem as follows.

$$\begin{aligned} \text{Max } \Pi = & \sum_{j \in J} \left((p^o - m_j - h_j - R^o v_j^o) x_j^o + (p^n - m_j - h_j - R^n v_j^n) x_j^n \right) \\ & - \sum_{j \in J} F_j y_j - \sum_{j \in J} \sum_{t \in T} \tau^t \mu_j^t \end{aligned} \quad (3)$$

Subject to,

$$\sum_{j \in J} y_j \leq 1 \quad (4)$$

$$x_j^o \leq \Psi y_j \quad \forall j \in J \quad (5)$$

$$x_j^n \leq \Psi y_j \quad \forall j \in J \quad (6)$$

$$\sum_{j \in J} x_j^o \leq D_b^o - \alpha^o (p^o - p_b^o) \quad (7)$$

$$\sum_{j \in J} x_j^n \leq D_b^n - \alpha^n (p^n - p_b^n) \quad (8)$$

$$R^o x_j^o + R^n x_j^n \leq \sum_{t \in T} Q^t \mu_j^t \quad \forall j \in J \quad (9)$$

$$p^o \geq p^n \quad (10)$$

The objective function (3) maximizes the total profit which is equal to the revenues generated by the sales of organic and conventional products minus the different costs of production, distribution, purchasing, extraction plant opening, and mill machines acquisition. Constraint (4) imposes that at most only one facility can be selected. Using a sufficiently big constant, denoted by Ψ , constraints (5) and (6) guarantee for organic and conventional olive oil, respectively, that the production cannot occur in a facility that is not open. Constraint (7) imposes that the produced quantity of organic olive oil cannot exceed the incurred demand for this variety. The demand function of the organic variety was given in equation (1). In the model, price is a decision variable (p^o for organic olive oil). With the same approach, we impose in constraint (8) that the produced quantity of conventional olive oil cannot exceed the incurred demand for this variety. Constraint (9) ensures that the total processed quantity of olives in facility j (i.e., $R^o x_j^o + R^n x_j^n$) does not exceed the total installed capacity (i.e., $\sum_{t \in T} Q^t \mu_j^t$). Finally, constraint (10) guarantees that the organic price (p^o) is higher than the conventional price (p^n).

The proposed model is non-linear due to the consideration of pricing decisions in the objective function as this leads to the multiplication of an integer variable (quantity) by a continuous variable (price). In its present form, our model cannot be solved with commercial optimization software like CPLEX due to the complex non-linearity of the objective function. Moreover, the use of non-linear solvers is not an interesting option since these solvers have often a weak performance when there are many constraints in the model. If we express the prices and costs in cents, we can consider the prices p^o and p^n as integer variables ($p^o, p^n \in \mathbb{N}$). In this case, the non-linearity of the objective function is given by the product of two integer variables (price times quantity of olive oil). Since our problem is based in a real case, we can consider, without loss of generality, that there is an upper bound for the price and for the quantity of olive oil produced. The upper bound of the price is fixed based on the prices observed in the market and the upper bound for the quantity of olive oil produced depends on the demand.

The linearization of the product of two integer variables requires the representation of the original integer variables

(price and quantity of olive oil) in a binary format. Consequently, for each integer variable (p^o, p^n, x_j^o, x_j^n) is required a set of binary variables. The number of binary variables depends on the upper bound of the integer variable. For example, if the upper bound for the organic price (p^o) is fixed at 20, five binary variables are needed to represent all the possible prices, lets denote them as $\rho_1^o, \rho_2^o, \dots, \rho_5^o$. In particular, $p^o = 19 = \sum_{i=1}^5 (2^{i-1})\rho_i^o$ where $\rho_1^o = 1, \rho_2^o = 1, \rho_3^o = 0, \rho_4^o = 0, \rho_5^o = 1$. Based on the binary representation of the integer variables, the product of two integer variables is transformed into the product of two binary variables. Therefore, an extra set of binary variables is necessary to represent this product. For each pair of binary variables used on the representation of the organic price (ρ_i^o) and the quantity of organic olive oil (o_j^k), is declared a binary variable z_j^{ik} . For the conventional price (ρ_i^n) and quantity of conventional product (η_j^s), it proceeds in a similar way (the product is represented by the binary variable w_j^{is}).

Hereafter are detailed the extra data and variables used on the linearization of the model.

- p_u^o : upper bound for the organic price,
- p_u^n : upper bound for the conventional price,
- I : maximum number of binary variables required for the price (organic and conventional) representation,
- ρ_i^o : binary variables used for the organic price representation,
- ρ_i^n : binary variables used for the conventional price representation,
- K : maximum number of binary variables required for the representation of the quantity of organic olive oil produced,
- o_j^k : binary variables used for the representation of the quantity of organic olive oil produced at facility j ,
- S : maximum number of binary variables required for the representation of the quantity of conventional olive oil produced,
- η_j^s : binary variables used for the representation of the quantity of conventional olive oil produced at facility j ,
- z_j^{ik} : binary variables used on the linearization of the product of the organic price times the quantity of organic olive oil produced at facility j ,
- w_j^{is} : binary variables used on the linearization of the product of the conventional price times the quantity of conventional olive oil produced at facility j .

We recall that the objective function (expression 3) maximizes the total profit obtained for the sales of all the olive oil produced. This expression can be described as: revenue obtained from the sales of all the olive oil, minus the total cost incurred on the production of the olive oil (it includes: the olives purchasing cost, the production/pressing cost, and the distribution cost), minus the opening cost (annual amortized opening cost of the selected facility), minus the total acquisition cost for the required machines. The nonlinearity is observed in the revenue expression, in which are multiplied the sales price times the total quantity of olive oil produced. As we mention before, it is possible to linearize this expression using the auxiliar binary variables

z_j^{ik} and w_j^{is} . The linearized objective function is expression (11).

$$\begin{aligned} \text{Max } \Pi = & \sum_{j \in J} \sum_{i=1}^I \left(\sum_{k=1}^K (2^{i-1})(2^{k-1})z_j^{ik} + \sum_{s=1}^S (2^{i-1})(2^{s-1})w_j^{is} \right) \\ & - \sum_{j \in J} \left((m_j + h_j + R^o v_j^o)x_j^o + (m_j + h_j + R^n v_j^n)x_j^n \right) \\ & - \sum_{j \in J} F_j y_j - \sum_{j \in J} \sum_{t \in T} \tau^t \mu_j^t \end{aligned} \quad (11)$$

Constraints (4) to (10) remains without modifications. And the following constraints are added to the model to describe the relation between the auxiliary binary variables ($\rho_i^o, \rho_i^n, o_j^k, \eta_j^s, z_j^{ik}, w_j^{is}$) and the original variables (p^o, p^n, x_j^o, x_j^n).

$$x_j^o = \sum_{k=1}^K (2^{k-1})o_j^k \quad \forall j \in J \quad (12)$$

$$x_j^n = \sum_{s=1}^S (2^{s-1})\eta_j^s \quad \forall j \in J \quad (13)$$

$$p^o = \sum_{i=1}^I (2^{i-1})\rho_i^o \quad (14)$$

$$p^n = \sum_{i=1}^I (2^{i-1})\rho_i^n \quad (15)$$

$$p^o \geq p_b^o \quad (16)$$

$$p^o \leq p_u^o \quad (17)$$

$$p^n \geq p_b^n \quad (18)$$

$$p^n \leq p_u^n \quad (19)$$

$$z_j^{ik} \leq o_j^k \quad \forall i \in [1, \dots, I], k \in [1, \dots, K], j \in J \quad (20)$$

$$z_j^{ik} \leq \rho_i^o \quad \forall i \in [1, \dots, I], k \in [1, \dots, K], j \in J \quad (21)$$

$$z_j^{ik} \geq o_j^k + \rho_i^o - 1 \quad \forall i \in [1, \dots, I], k \in [1, \dots, K], j \in J \quad (22)$$

$$w_j^{is} \leq \eta_j^s \quad \forall i \in [1, \dots, I], s \in [1, \dots, S], j \in J \quad (23)$$

$$w_j^{is} \leq \rho_i^n \quad \forall i \in [1, \dots, I], s \in [1, \dots, S], j \in J \quad (24)$$

$$w_j^{is} \geq \eta_j^s + \rho_i^n - 1 \quad \forall i \in [1, \dots, I], s \in [1, \dots, S], j \in J \quad (25)$$

Constraints (12) and (13) establish a relation between the integer variables x_j^o and x_j^n respectively and their representation in binary variables. Likewise, constraints (14) and (15) describes the relation between the integer price variables and their representation in binary variables. Constraints (16) to (19) guarantee that the price variables belong to their definition interval determined by the corresponding lower and upper bound. Constraints (20) to (22) are related with the linearization of the organic price times the quantity of organic olive oil. And in the same way, constraints (23) to (25) correspond with the linearization of the conventional price times the quantity of conventional product.

Finally, the mathematical model is given by the objective function (11) and the constraints (4) to (10) and (12) to (25).

5 THE FRANCE CASE

In this study we are focusing on the French scenario. The average sales volume of olive oil in large distribution in France is about 65, 000 tons per year, with 57400 tons for conventional olive oil and 7600 tons of organic olive oil (Afidol, 2017).

The market share of south east of France represents approximately 13.6% (Afidol, 2017), which means around 7806 tons and 1034 tons for conventional and organic olive oil, respectively.

In our practical case study, we consider that the producer can capture a base demand $D_b^n = 1561 \text{ tons}$ and $D_b^o = 207 \text{ tons}$, which represent approximately 20% of average yearly sales in large distribution in south east France for the conventional and organic variety, respectively. We consider the base prices $p_b^n = 5.8 \text{ €/Kg} = 580 \text{ c/Kg}$ and $p_b^o = 8.7 \text{ €/Kg} = 870 \text{ c/Kg}$ for conventional and organic olive oil, respectively. To obtain these prices, we first considered the final prices offered in the large distribution sector in France (we used the prices announced on the online purchase platforms of the 3 largest hypermarkets). Second, we calculated the average final price per Kg of each variety. Finally, we assumed that the hypermarkets have a net margin of 20% and deduced the base price of olive oil for the producer.

For the price-sensitivity of conventional demand, we assume that the base demand drops by 75% when the conventional product is offered at the base price of the organic product (i.e., $8.7 \text{ €/Kg} = 870 \text{ c/Kg}$), which makes sense because most of the customers will switch to buy the organic product when both varieties have the same price. Therefore, we obtain $\alpha^n = 404 \text{ c}^{-1}$. Since organic customers are typically much less sensitive to price than conventional customers, we consider that $\alpha^o = 101 \text{ c}^{-1}$. We are aware that the values of α^o and α^n are difficult to estimate in practice even with empirical studies.

As we mention before, the production, distribution cost, purchasing cost and annual amortized open cost depends on the facility location. In our study we consider the following potential locations: France, Spain, Tunisia, Italy, Turkey, Greece, Morocco, Portugal, and Algeria. All these countries belong to the International Olive Oil Council and all are located in the Mediterranean zone. We assume that each potential location of the extraction plant gives access to a set of potential suppliers (farmers). In some countries like Tunisia and Spain that are among the largest producers of olives in the world, almost all olive oil producers rely on local supply. Purchasing olives from another country is not justified in these cases. In the case of France, the production volumes of olives are very low, which imposes some procurement constraints. Many of the French producers rely on the olives imported from Spain. Since the aim of this paper is to present a general configuration scheme, we assume that the transportation cost is including in the purchasing cost. The procurement costs of conventional olive are taken from (IOC, 2015). For the organic olive, our industrial partner considers that it is 25% more expensive. Table 1 summarize the practical data (IOC, 2015):

- m_j : Processing cost (cents per Kg of final product),
- h_j : Distribution cost (cents per Kg of final product),
- F_j : Annual amortized opening cos (in cents),
- v_j^n : Purchasing cost of Conventional Olives (cents per Kg of olive),
- v_j^o : Purchasing cost of Organic Olives (cents per Kg of olive).

We recall that the producer must also decide the total extraction capacity to implement in the selected plant. In this

study, we consider two types of extraction machines, with different capacities and investment costs. Table 2 shows the characteristics of each machine (Addinson, 2019). Notice that the producer can acquire one or several units of each type of machine.

Table 1: Cost relative to facility.

	m_j	h_j	F_j	v_j^n	v_j^o
France	30	2	6000000	100	80
Spain	3	5	4000000	85	68
Tunisia	4	10	1000000	68	54
Italy	10	5	4500000	80	64
Turkey	6	30	900000	23	19
Greece	7	8	3800000	75	60
Morocco	4	13	1100000	40	32
Portugal	4	6	4000000	55	35
Algeria	3	10	1050000	70	57

Table 2: Characteristics of olive oil extraction machines

	Small size extraction machine	Medium size extraction machine
Capacity of 1 machine per season, Q^t (Kg of processed olives)	486000	1965600
Annual amortized acquisition cost of 1 machine, τ^t (in cents)	1080000	3920000

5.1 Numerical example

In this section, we evaluate the efficiency of the proposed model by comparing it to the scenario of fixed prices. In this latter scenario, the selling prices are equal to the base prices p_b^o and p_b^n and, consequently, the effective demand is equal to the base demand.

To make the experiments, we generate test cases with a different number of potential facilities: from 3 until 9. For a fixed number of potential facilities, we vary the values of α^o from 100 to 500 with a step of 50. For each α^o , we consider 5 different values of α^n as follows: $\alpha^n = 2\alpha^o$, $\alpha^n = 3\alpha^o$, ... $\alpha^n = 6\alpha^o$. Hence, for a fixed number of potential facilities, it was tested 45 instances. Each instance is solved first with our approach and then, under fixed price conditions. We finally calculate, for each instance, the gain in profit that results from using our approach (O^*) instead of the fixed price approach (O_f) (expression 50). We also report the execution time in CPLEX (average time per test case).

$$Gain = \frac{O^* - O_f}{O_f} \times 100 \quad (50)$$

Table 3 summarizes the average execution time, the average gain, and the standard deviation of the gain for a given number of facilities. In table 4 are detailed the potential facilities considered in each test case group.

The obtained results show that our approach improves the profit by 42.96% on average in comparison to the case in which the selling prices are fixed and equal to the base prices. This shows the importance of considering the price as a decision variable as we do in our model. Another important remark is the viability of the procedure in terms of the computation time. Note that in all the cases it is possible to solve the linearized model using CPLEX in approximately 30 seconds.

Table 3: Numerical Results

Potential Facilities	Execution Time (s)	Gain	STD
3	6.68	51.80	44.49
4	10.27	51.79	44.48
5	12.01	39.43	36.58
6	16.04	39.43	36.58
7	25.86	39.43	36.58
8	31.92	39.43	36.58
9	34.42	39.43	36.58
Average	19.60	42.96	38.84

Table 4: Facility by test group

Potential Facilities	3	4	5	6	7	8	9
France	X	X	X	X	X	X	X
Spain	X	X	X	X	X	X	X
Tunisia	X	X	X	X	X	X	X
Italy		X	X	X	X	X	X
Turkey			X	X	X	X	X
Greece				X	X	X	X
Morocco					X	X	X
Portugal						X	X
Algeria							X

5.2 Effect on the price and the profit of the price sensitivity parameters

We also observe that, for a fixed number of potential facilities, there are a significant variance in the gain. We recall that the difference between these test cases is the value of the price sensitivity parameters (α^o, α^n). It means, for a fixed number of potential facilities, there are generate different scenarios, each one represents different types of customer. This result suggests the importance to select an adequate value of the price sensitivity parameters. As we mention before, the values of α^o and α^n are difficult to estimate in practice even with empirical studies.

In this subsection we will illustrate the effect of the price sensitivity parameter in the price and the profit. Let us consider the simplest test case group: only three potential facilities (France, Spain, and Tunisia).

Figures 1, 2, and 3 represent organic price (p^o) and conventional price (p^n) as function of the price sensitivity parameters (α^o, α^n). In all figures $\alpha^o \in [100, 400]$ and α^n is calculated using a positive factor: in Figure 1 $\alpha^n = 5\alpha^o$, in Figure 2 $\alpha^n = 6\alpha^o$ and in Figure 3 $\alpha^n = 7\alpha^o$. It is also represented the organic base price (blue line) and conventional base price (orange line).

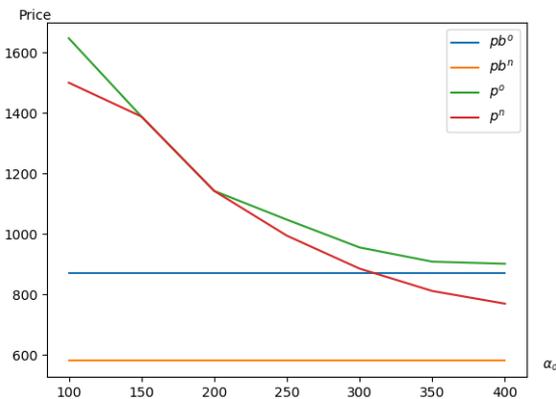


Figure 1: Price vs Price Sensitivity $\alpha^n = 5\alpha^o$

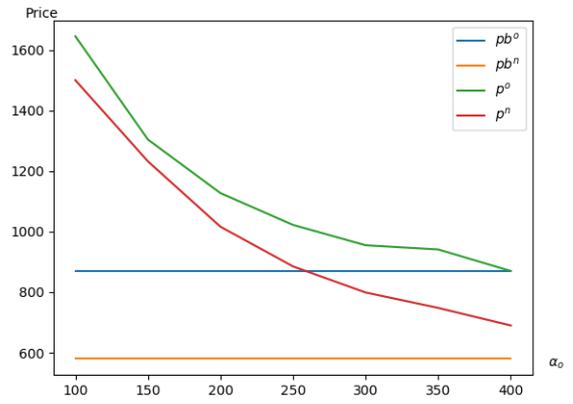


Figure 2: Price vs Price Sensitivity $\alpha^n = 6\alpha^o$

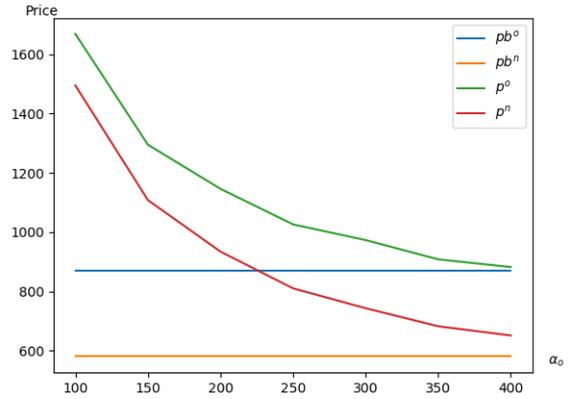


Figure 3: Price vs Price Sensitivity $\alpha^n = 7\alpha^o$

In these figures is observed that the price is decreasing in the price sensitivity parameter. And the difference between the organic and conventional price increase when the difference between the organic and conventional price sensitivity parameters increased.

Figure 4 illustrates the relation between the profit and the price sensitivity parameters. The organic price sensitivity parameter $\alpha^o \in [100, 400]$ and there are consider three relations for the conventional price sensitivity parameter: $\alpha^n = 5\alpha^o$, $\alpha^n = 6\alpha^o$ and $\alpha^n = 7\alpha^o$. It is also represented the base profit (blue line), it means, the profit obtained when the sales price is equal to the base price.

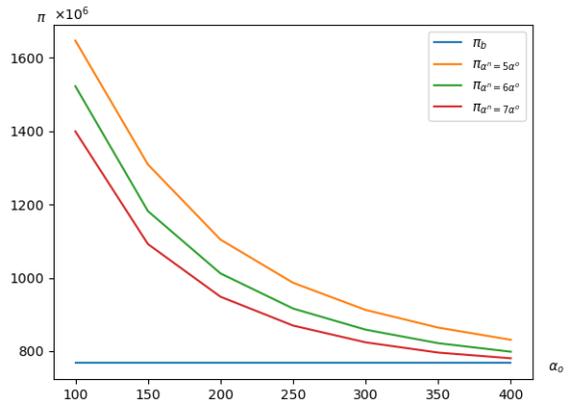


Figure 4 : Profit vs Price Sensitivity

As it is observed in Figure 4, the profit decrease in the price sensitivity parameter. For the same value of α^n , the higher profit is obtained for the small α^n . We recall that it is

assumed an endogenous demand functions given by the expressions $D^o = Db^o - \alpha^o(p^o - pb^o)$ for the organic product and $D^n = Db^n - \alpha^n(p^n - pb^n)$ for the conventional product. Hence, when the price sensitivity parameter increases, the effective demand decreases and consequently the profit decrease.

The obtained results can be very useful to the managers since can suggest a price policy adequate the characteristics of their customers. If the customers are extreme sensitivity to the price variations then, the selling price should be closer to the base price. Another important point is the relation between the organic and conventional customers. The greater difference between the parameters α^o and α^n , the greater differences between the organic and the conventional price.

6 CONCLUSIONS

In the last years, olive Oil consumption has been significantly increased in non-traditional markets. These new market opportunities have generated new challenges in term of logistics of the olive oil.

In this paper, we developed a model to analyze the price and production configuration of olive oil with organic and conventional products. We formulate a mathematical model that aims to reflect the main strategic-tactical decisions in the olive oil sector: facility location, capacity decisions, pricing and endogenous demand function that depends on the quality of the product (organic or conventional). We describe and solve a linearized version of the model.

Our study focused on the French market. The obtained results show that our approach improves the profit by 42.96% on average compared to the case in which the selling prices are fixed and equal to the base prices. Our results show the importance of including pricing decisions with supply chain design decisions.

At the moment of writing this document we are performing additional experiments to derive managerial insights from our model. We are interested in the following topics:

- The influence of the price sensitivity parameter in the facility location, it means, which customers prefers the French product?
- The impact of organic and conventional market size, i.e., how is the influence of the organic and conventional demand in the olive oil supply chain configuration? What is more profitable, offer a single variety or both?

7 ACKNOWLEDGMENT

This research has been supported by the French National Research Agency (ANR) through Project CONCLuDE and by the Area of Excellence G3D at Rennes School of Business.

8 REFERENCES

Addison, S., Addison, A., Dautricourt, M., Gloria, B., & Martin, B. (n.d.). *Sample Costs: The Olive Oil Source*. The Olive Oil Source. Retrieved November 28, 2020, from <https://www.oliveoilsource.com/page/sample-costs>

Association Française Interprofessionnelle de l'Olive (AFIDOL) 2017. Memo 2017 Huile d'olive en France. Available at: https://afidol.org/wp-content/uploads/2018/06/Memo2017_Huile-dolive-en-France.pdf

Association Française Interprofessionnelle de l'Olive (AFIDOL) 2019. Rapport d'activités 2019. Available at <https://afidol.org/wp-content/uploads/2020/06/AG2020-Pr%C3%A9sentation-Rapport-dActivit%C3%A9s-2019-vdef-PF.pdf>

Ballco, P., & Gracia, A. (2020). Do market prices correspond with consumer demands? Combining market valuation and consumer utility for extra virgin olive oil quality attributes in a traditional producing country. *Journal of Retailing and Consumer Services*, 53, 101999.

Boncinelli, F., Contino, C., Romano, C., Scozzafava, G., & Casini, L. (2016). Territory, environment, and healthiness in traditional food choices: Insights into consumer heterogeneity. *International Food and Agribusiness Management Review*, 20(1), 143–157.

Bonti-Ankomah, S., & Yiridoe, E. K. (2006). *A literature Review of the Economic of Consumer Perceptions and Preferences* (p. 63). Organic Agriculture Centre of Canada.

Cacchiarelli, L., Carbone, A., Laureti, T., & Sorrentino, A. (2016). The value of the certifications of origin: A comparison between the Italian olive oil and wine markets. *British Food Journal*, 118(4), 824–839.

Cavallo, C., & Piqueras-Fizman, B. (2017). Visual elements of packaging shaping healthiness evaluations of consumers: The case of olive oil. *Journal of Sensory Studies*, 32(1), e12246.

Del Giudice, T., Cavallo, C., Caracciolo, F., & Cicia, G. (2015). What attributes of extra virgin olive oil are really important for consumers: A meta-analysis of consumers' stated preferences. *Agricultural and Food Economics*, 3(1), 20.

International Olive Council (IOC) 2015. International Olive Oil Production Cost Study. Available at <https://www.internationaloliveoil.org/wp-content/uploads/2019/11/INTERNATIONAL-OLIVE-OIL-PRODUCTION-COSTS-STUDY-.pdf>

Karanikolas, P., Correia, T. P., Martinez-Gomez, V., Galli, F., Hernandez, P. A., Fastelli, L., Arnalte-Mur, L., Mendez, M. R., Prosperi, P., & Goussios, G. (2018). Food system integration of olive oil producing small farms: A comparative study of four Mediterranean regions. *13th European International Farming Systems Association (IFSA) Symposium, Farming Systems: Facing Uncertainties and Enhancing Opportunities, 1-5 July 2018, Chania, Crete, Greece*, 1–20.

Kazaz, B. (2004). Production Planning Under Yield and Demand Uncertainty with Yield-Dependent Cost and Price. *Manufacturing & Service Operations Management*, 6(3), 209–224.

Kazaz, B. (2008). *Pricing and production planning under supply uncertainty*.

Kazaz, B., & Webster, S. (2011). The Impact of Yield-Dependent Trading Costs on Pricing and Production Planning Under Supply Uncertainty. *Manufacturing & Service Operations Management*, 13(3), 404–417.

Liberatore, L., Casolani, N., & Murmura, F. (2017). Perception of Extra-Virgin Olive Oil Certifications: A Commodity Market Perspective. *Izvestiya. Journal of Varna University of Economics*, 61(1), 77–82.

Palacios-Argüello, L., Gondran, N., Nouira, I., Girard, M.-A., & Gonzalez-Feliu, J. (2020). Which is the relationship between the product's environmental criteria

- and the product demand? Evidence from the French food sector. *Journal of Cleaner Production*, 244, 118588.
- Romo Muñoz, R., Lagos Moya, M., & Gil, J. M. (2015). Market values for olive oil attributes in Chile: A hedonic price function. *British Food Journal*, 117(1), 358–370.
- Roselli, L., Carlucci, D., & Gennaro, B. C. D. (2016). What Is the Value of Extrinsic Olive Oil Cues in Emerging Markets? Empirical Evidence from the U.S. E-Commerce Retail Market. *Agribusiness*, 32(3), 329–342.
- Roselli, L., Carlucci, D., Rover, O. J., & Gennaro, B. D. (2018). The Effects of Extrinsic Cues on Olive Oil Price in Brazil. *Journal of International Food & Agribusiness Marketing*, 30(1), 70–87.
- Sayadi, S., Erraach, Y., & Parra-López, C. (2017). Translating consumer's olive-oil quality-attribute requirements into optimal olive-growing practices: A quality function deployment (QFD) approach. *British Food Journal*, 119(1), 190–214.
- Sazvar, Z., Rahmani, M., & Govindan, K. (2018). A sustainable supply chain for organic, conventional agro-food products: The role of demand substitution, climate change and public health. *Journal of Cleaner Production*, 194, 564–583.
- Scalco, A. R., Ganga, G. M. D., De Oliveira, S. C., & Baker, G. (2020). Development and validation of a scale for identification of quality attributes of agri-food products in short chains. *Geoforum*, 111, 165–175.
- Vlontzos, G., & Duquenne, M. N. (2014). Assess the impact of subjective norms of consumers' behaviour in the Greek olive oil market. *Journal of Retailing and Consumer Services*, 21(2), 148–157.
- Yanguí, A., Font, M. C., & Roig, J. M. G. (2014). Revealing additional preference heterogeneity with an extended random parameter logit model: The case of extra virgin olive oil. *Spanish Journal of Agricultural Research*, 3, 553–567.

On Retailers' Profit in A Price- and Greenness-Sensitive Market with Different Demand Functions

ERFAN ASGARI¹, YANNICK FREIN¹, RAMZI HAMMAMI²

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP*, G-SCOP, 38000 Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes

erfan.asgari@grenoble-inp.fr

yannick.frein@grenoble-inp.fr

² Rennes School of Business, 35065 Rennes, France

hammami.ramzi@gmail.com

Résumé – Un détaillant propose un produit à ses clients sur un marché sensible aux prix et à l'écologie. La demande baisse le prix et augmente l'amélioration des émissions de carbone (CEI). Nous étendons les travaux existants en considérant des fonctions de demande plus réalistes au lieu de la fonction de demande linéaire communément utilisée. Tout d'abord, nous considérons la nième racine de la fonction de demande où la demande augmente de manière non linéaire en termes d'amélioration des émissions de carbone. Ensuite, nous considérons une limite supérieure de la demande pour modéliser le fait qu'après un certain seuil limite, l'amélioration de la performance environnementale n'a pas d'effet sur la demande. Nous appliquons cette contrainte de limite supérieure pour les fonctions de demande linéaires et non linéaires. Les nouveaux problèmes (non-linéaire, non linéaire et linéaire bornée) sont formulés dans un environnement stochastique et résolu par une approche analytique. Le modèle avec racine carrée bornée est considéré comme une référence. Nous réalisons des expériences numériques approfondies et calculons l'écart de profit entre les différents modèles et le modèle de référence. La demande non linéaire bornée est considérée comme une référence et est comparée aux autres, les résultats révèlent que lorsque la borne est relativement faible (moins de 20% pouvant être obtenue par l'amélioration des émissions), le modèle linéaire borné fonctionne mieux que les autres mais au delà, le modèle non linéaire fonctionne mieux.

Abstract – A retailer offers a product to customers in a price- and greenness- sensitive market. The demand decreases in price and increases in Carbon Emissions Improvement (CEI). We extend the extant works by considering more realistic demand functions instead of the commonly-used linear demand function. First, we consider the n th root demand function where demand non-linearly increases in CEI. Second, we consider an upper-bound on demand to model the fact that after a certain threshold limit, improving the environmental performance does not have any effect on demand.. We apply this upper-bound constraint for both linear and non-linear demand functions. The new problems (non-linear, non-linear with cap, and linear with cap) are formulated in a stochastic environment and solved by an analytical approach. The model with square root demand and cap is considered as a reference. We perform extensive numerical experiments and calculate the profit gap between the different models and the reference model. We find that the linear-cap demand is the best approximation when η (the parameter that determines demand's cap) is low (less than 19%). For a higher η , the square root demand provides better approximation. The classical linear demand (without cap), which is often used in the literature, is never the best one. However, the linear-cap demand is a good approximation when the cap is small. The linear-cap demand is interesting because we can easily adapt it to existing models in the literature.

Mots clés - gestion de la chaîne d'approvisionnement, détaillant, qualité environnementale, demande complexe, prix.

Keywords – supply chain management, retailing, greenness, complex demand function, pricing.

1 INTRODUCTION

The product's environmental level is often assessed by the carbon emissions emitted in production and transportation. The emission intensity of the product (production carbon emissions + transportation carbon emissions) is becoming an essential purchase criterion for many customers (Borin et al., 2013; Hammami et al., 2018; Palacios-Argüello et al., 2020). In 2019, Accenture executed a survey of 6,000 consumers. The survey is executed in 11 countries across North America, Europe, and Asia. The results show that 72% of consumers are currently buying more environmentally-friendly products than they were buying five years ago. Also, 81% of them expect to buy more over the next five years. Moreover, more than half of

consumers would pay more for environmentally-friendly products (Accenture, 2019). Carbon Trust surveys indicate that approximately 20% of customers prefer to buy green products even if they are more expensive than regular products (Hong et al., 2019).

Empirical studies and surveys reveal the connection between the product's demand and its emission intensity. The classical models in Operation Research (OR) and Supply Chain Management (SCM) areas have been revised to embedding a greenness-sensitive demand (see e.g., Hovelaque et Bironneau, 2015; Liu et al., 2012; Zhang et al., 2015). Customers' Environmental Awareness (CEA) will directly influence the

enterprise's strategy to produce green products and to invest and adopt cleaner technology.

Customers' desire to purchase greener products and willingness to pay the premium price lead many retailers to have a better and clear understanding of customers' preferences. Thus, as highlighted by Ramanathan et al., 2012, these retailers asked their suppliers to reduce carbon emissions in production and transportation phases. It is common in practice that retailers decide the product's greenness level (see Bestseller, 2018; Walmart, 2017 for more details).

In the literature, the researchers considered a linear demand in function of product's emission intensity to present the link between demand and emission intensity. However, Palacios-Argüello et al., 2020 found out that the linear sensitivity of demand to carbon emissions is insufficient. Therefore, we develop and analyze a more general demand function with respect to the product's emissions intensity.

We consider a retailer that offers one product to customers to maximize its profit with different demand functions. The retailer's profit maximization problem consists in deciding the product's price, greenness, and order size.

In the following section, we review and present related papers. In section 3, we describe our problem and develop mathematical models. All models are solved by an analytical approach (all the proofs are presented in the appendices). Then in section 4, we compare the results and derive insights. Finally, the conclusion is presented in section 5.

2 LITERATURE REVIEW

In this section, we review the recent studies that considered a linear endogenous demand in function of greenness, sensitive to the product's price and carbon intensity.

Ghosh et Shah, 2012, considered a SC that is composed of one manufacturer and one retailer. The manufacturer produces the product and sells it through the retailer. The product's demand is a linear function of its greenness and price. In centralized SC, the manufacturer decides product's greening level and the retailer decides product's retail price. Different scenarios are considered in decentralized SC. They considered a quadratic cost function for producing greener product.

Ma et al., 2013 investigated the different contracts' impact, in a two-stage SC that is composed of one manufacturer and one retailer, on SC's profit and product's greenness. The product's demand is a linear function of the retailer's price, product's quality level, and retailer's marketing efforts. They formulated a deterministic model to analyze and compare the profit of each contract.

Ghosh and Shah, 2015 introduced "Cost Sharing Contract" and studied a similar problem like Ghosh and Shah, 2012. One product is offered in a SC composed of one manufacturer and one retailer. The product's demand is linearly decreasing in the product's retail price and increasing in greening improvement level. According to this new contract, the retailer takes a share of green improvement costs.

Xu et al., 2016 investigated the CEA's impact, in a closed-loop SC composed of a manufacturer and a retailer, on the optimal strategies. The used products are not abandoned and returned to the life cycle by re-manufacturing. The product's demand is linearly decreasing in the product's price and increasing in the product's eco-friendly level. Different scenarios are used to find out more about the SC system in economic, environmental, and social aspects.

Qin et al., 2017 studied a Make-To-Order (MTO) SC composed of one manufacturer and one retailer. The problem is formulated with a nonlinear stochastic model. The demand

is a linear function and sensitive to price and carbon emission reduction. They combined various demand forecasting and SC strategies. Three scenarios are considered: Without Information Sharing, Full Information Sharing, and Retailer-only Forecasting. In all scenarios, the manufacturer is the leader and the retailer is the follower.

Hong and Guo, 2019, considered a MTO supply chain that is composed of one manufacturer and one retailer under different cooperation contracts: Price-only contract, Green-marketing cost-sharing contract, Full channel coordination. The product's demand is linearly decreasing in the product's price and increasing in the product's greenness level. The profit maximization problem (manufacturer and retailer) is formulated in a stochastic environment. The manufacturer is the leader in Stackelberg game, and the retailer is the follower. Liu et al., 2020 considered a SC that is composed of one manufacturer and one retailer. The manufacturer invests in green innovation, produces a product, and sells it to the retailer. The retailer offers the product to customers. Product's demand is a linear function of its price and greenness. They attempt to investigate the effects of cooperative mechanisms (i.e., revenue sharing and cost-sharing contracts) on innovation investment and pricing strategies under different scenarios.

To the best of our knowledge, most relevant papers in the literature consider linear demand in function of product's greenness. In the following section, we develop different demands functions and solve the models by an analytical approach.

3 PROBLEM DESCRIPTION

As shown in the literature review, to the best of our knowledge, most studies in the literature used the same linear demand. We are interested in considering general (especially non-linear) demand function with respect to the greenness. The objective of this paper is to investigate the effects of other demands functions that may be more realistic under some circumstances. The analysis that have been done by Palacios-Argüello et al., 2020 reveal that the demand (in function of environmental level) is not linear. In addition the increase of this demand thanks to the improvement of environmental quality, and more precisely thanks to the decreasing of emissions, is bounded. Hence, it is not useful to decrease these emissions beyond a given value. For that, we are going to study the retailer's profit optimization problem under different demands functions.

We study the optimal strategies of a retailer who offers a product to customers at a price- and carbon emission- sensitive market. The retailer keeps a stock, S , to serve customers. The inventory level decreases one unit as soon as a customer arrives. The retailer sells products at the retail price, p . We consider x_0 as the product's carbon emission reference. Therefore, x_0 is the maximum Carbon Emission Improvement (CEI) possible. CEI's cost is given by bx^2 , where b is the CEI's cost factor, and x is the CEI. The retailer will maximize its profit by deciding the optimal price, carbon emission improvement, and order size. Figure 1 shows the described problem.

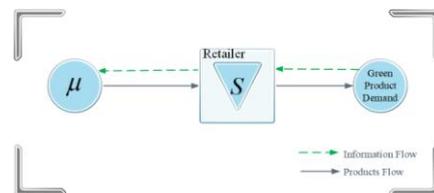


Figure 1. Supply chain structure

The retailer makes an order to fill inventory from a supplier that adopts a MTS manufacturing policy. The inventory is managed according to the standard (q, S) policy where q is the reorder point, and S is the replenishment lot size. We assume that the retailer makes an order to refill the stock when there is no product left in its inventory. During the replenishment time, demand is satisfied with safety stock. We assume that the safety stock is large enough to satisfy all demands. The sizing and management of the safety stock are out of the scope of this study.

Since customers' arrival and inventory refilling service time follow the Exponential distribution with mean rates μ and λ , respectively, the number of part in main stock is a Markovian process. In this case, we calculate the probability of stock-out, in which there is no item in the inventory, using continuous Markov chain process properties. Regarding Figure 2, let us define ψ_i , which stands for the probability of system in state i (i.e., having i item in stock). The probability of having zero stock is $\psi_0 = \frac{\lambda}{\lambda + \mu S}$ (See Asgari et al., 2020 for more details).

Hence, service level constraint for retailer is given by $\frac{\lambda}{\lambda + \mu S} \leq r$. Thus, the expected value of the average inventory level for the retailer can be given by $\frac{S}{2}$. We let $2h$ denote the unit inventory cost. Thus, the inventory cost for the retailer is hS .

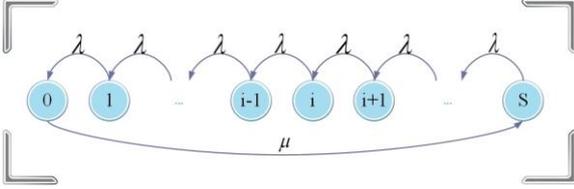


Figure 2. Inventory refilling policy

As we already said above, the linear demand function is well known in the literature. However, we mentioned before that our partners in project ANR CONCLuDE at École des Mines de Saint-Étienne found out that the linear sensitivity of demand to carbon emissions is insufficient. Therefore, we are interested in considering a new demand function (precisely, non-linear) to see how results differ from traditional linear demand functions. We consider that we know the product's demand when the retailer offers the standard product (i.e., without improvement). We also know the demand's shape of the curve (e.g., square root). Therefore, we can deduce the last point (when we improve 100%). Hence, we can deduce the linear demand from the two extreme points and calculate the corresponding greenness sensitivity parameter, $\beta_{e(L)}$.

In the literature, the researchers consider the linear demand without any constraint, but we know that in reality, the demand does not increase beyond a maximum value (Palacios-Argüello et al., 2020). This value is obtained for a given x that we call x_l . In other words, the demand increases in CEI until x_l , and after that, the demand is constant even when CEI increases. Thus, we consider a constraint on the demand that we call cap. We consider that the maximum increasing obtained thanks to carbon emissions improvement is ηA , for a given price p . Figure 3 shows four demands functions for a given price that we consider: non-linear, non-linear cap, linear, and linear cap.

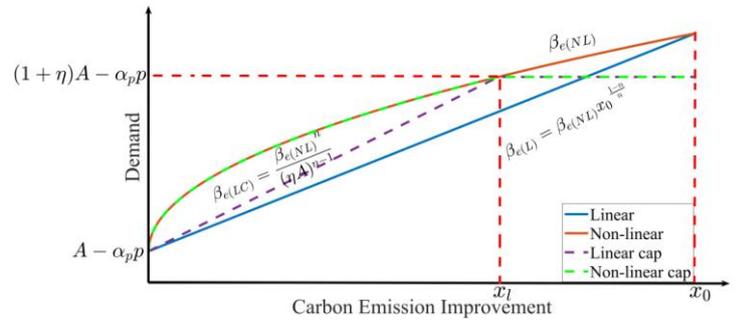


Figure 3. Demands functions for a given price

The mathematical problems and the optimal solutions are provided in the following section.

4 MATHEMATICAL MODEL

4.1 Linear demand function (M_0)

In this benchmark model, the retailer is alone in the market. The problem consists in deciding price p , carbon emission level x , and order size S to maximize the retailer's expected profit. A single product is offered to customers in this case. The demand function is presented in the following.

$$\lambda = A - \alpha_p p + \beta_{e(L)} x$$

Where A is market potential, α_p and $\beta_{e(L)}$ are price's- and CEI's- sensitivity parameters, respectively. p and x represent the product's retail price and carbon emission improvement, respectively, which are decision variables. Note that x_0 indicates the carbon emission reference (thus, maximum carbon emission improvement is equal to x_0 ; $x \leq x_0$). The last decision variable of our model is order size, S , as we mentioned earlier.

The mathematical model of the stochastic problem is presented in the following.

$$\max_{x,p,S} \pi = (p - (c + bx^2))\lambda - hS \quad (1)$$

subject to

$$\psi_0 = \frac{\lambda}{\lambda + \mu S} \leq r \quad (2)$$

$$\lambda = A - \alpha_p p + \beta_{e(L)} x \quad (3)$$

$$x, p, S \geq 0, \quad x \leq x_0, \quad \lambda \geq 0.$$

The objective is to maximize the total expected profit given in equation 1. This profit is equal to the revenue (i.e., $p\lambda$) – the procurement cost (i.e. $(c + bx^2)\lambda$ – the inventory cost (i.e., hS). Equation 2 is the service level constraint. It ensures that stock-out probability does not exceed a predetermined level of r (i.e., $1 - r$ is the minimum service level). The mean demand rate is given in equation 3. To solve this model, we first transform it into a single-variable model based on the results of Lemma 1 and 2 given below. All proofs are presented in the Appendices.

Lemma 1. For any given values of p and x , service level constraint (Equation 2) is binding and the optimal amount of order size is $S^* = \frac{1-r}{r\mu} (A - \alpha_p p + \beta_{e(L)} x)$.

Since we assume that $\lambda \geq 0$, then, S^* is positive and can be substituted by its expression. The equivalent model with two variables, p and x , is presented in the following.

$$\max_{x,p} \pi = \left(p - \left(c + bx^2 + \frac{h(1-r)}{r\mu} \right) \right) \lambda \quad (4)$$

subject to

$$\lambda = A - \alpha_p p + \beta_{e(L)} x$$

$$x, p, S, \lambda \geq 0, \quad x \leq x_0.$$

Lemma 2. For any given value of x , the optimal price is

$$p^* = \frac{\alpha_p bx^2 + \beta_{e(L)} x + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p}.$$

The proof is presented in Appendix B. We present a constraint on A that ensures demand is positive. By substituting p^* in Equation 4, the equivalent model with only one variable, x is presented in the following:

$$\max_{0 \leq x \leq x_0} \pi = \frac{\left(-\alpha_p b x^2 + \beta_{e(L)} x + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)^2}{4\alpha_p} \quad (5)$$

$$\text{subject to} \quad \lambda = \frac{-\alpha_p b x^2 + \beta_{e(L)} x + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)}{2} \quad (6)$$

Proposition 1. The optimal solutions for linear demand model (M_0) are properly presented in the following.

$$x^* = \min \left\{ \frac{\beta_{e(L)}}{2\alpha_p b}, x_0 \right\}$$

$$p^* = \begin{cases} \frac{3\beta_{e(L)}^2}{4\alpha_p b} + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right) & \text{for } x^* \neq x_0 \\ \frac{\alpha_p b x_0^2 + \beta_{e(L)} x_0 + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)}{2\alpha_p} & \text{for } x^* = x_0 \end{cases}$$

$$S^* = \begin{cases} \frac{(1-r) \left(\frac{\beta_{e(L)}^2}{4\alpha_p b} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)}{2r\mu} & \text{for } x^* \neq x_0 \\ \frac{(1-r) \left(-\alpha_p b x_0^2 + \beta_{e(L)} x_0 + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)}{2r\mu} & \text{for } x^* = x_0 \end{cases}$$

4.2 n th root demand (M_1)

In this problem, we consider a general demand function. When $n = 2$, the demand function is the square root. The demand function is linearly decreasing in retail price and non-linearly (for $n \geq 2$), increasing in carbon emission improvement. The demand function is presented in the following.

$$\lambda = A - \alpha_p p + \beta_{e(NL)} \sqrt[n]{x}$$

The mathematical model of the stochastic problem is presented in the following.

$$\max_{x,p,S} \pi = (p - (c + bx^2))\lambda - hS \quad (7)$$

$$\text{subject to} \quad \psi_0 = \frac{\lambda}{\lambda + \mu S} \leq r \quad (8)$$

$$\lambda = A - \alpha_p p + \beta_{e(NL)} \sqrt[n]{x} \quad (9)$$

$$x, p, S \geq 0, \quad x \leq x_0, \quad \lambda \geq 0.$$

Lemma 3. For any given values of p and x , service level constraint (Equation 2) is binding and the optimal amount of order size is $S^* = \frac{1-r}{r\mu} (A - \alpha_p p + \beta_{e(NL)} \sqrt[n]{x})$.

The proof is similar to Lemma 1. Since we assume that $\lambda \geq 0$, then, S^* is positive and can be substituted by its expression. The equivalent model with two variables, p and x , is presented in the following.

$$\max_{x,p} \pi = \left(p - \left(c + bx^2 + \frac{h(1-r)}{r\mu}\right)\right)\lambda \quad (10)$$

$$\text{subject to} \quad \lambda = A - \alpha_p p + \beta_{e(NL)} \sqrt[n]{x}$$

$$x, p, S, \lambda \geq 0, \quad x \leq x_0.$$

Lemma 4. For any given value of x , the optimal price is

$$p^* = \frac{\alpha_p b x^2 + \beta_{e(NL)} \sqrt[n]{x} + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)}{2\alpha_p}.$$

The proof is similar to Lemma 2. By substituting p^* in Equation 4, the equivalent model with only one variable, x is presented in the following:

$$\max_{0 \leq x \leq x_0} \pi = \frac{\left(-\alpha_p b x^2 + \beta_{e(NL)} \sqrt[n]{x} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)^2}{4\alpha_p} \quad (11)$$

subject to

$$\lambda = \frac{-\alpha_p b x^2 + \beta_{e(NL)} \sqrt[n]{x} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)}{2}$$

Proposition 2. The optimal solutions for n th root demand model (M_1) are properly presented in the following.

$$x^* = \min \left\{ \left(\frac{\beta_{e(NL)}}{2\alpha_p b n}\right)^{\frac{n}{2n-1}}, x_0 \right\}$$

$$p^* = \begin{cases} \frac{\left(\frac{(2n+1)\beta_{e(NL)}}{2n} \left(\frac{\beta_{e(NL)}}{2\alpha_p b n}\right)^{\frac{1}{2n-1}} + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)}{2\alpha_p} & \text{for } x^* \neq x_0 \\ \frac{\alpha_p b x_0^2 + \beta_{e(NL)} \sqrt[n]{x_0} + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)}{2\alpha_p} & \text{for } x^* = x_0 \end{cases}$$

$$S^* = \begin{cases} \frac{(1-r) \left(\frac{(2n-1)\beta_{e(NL)}}{2n} \left(\frac{\beta_{e(NL)}}{2\alpha_p b n}\right)^{\frac{1}{2n-1}} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)}{2r\mu} & \text{for } x^* \neq x_0 \\ \frac{(1-r) \left(-\alpha_p b x_0^2 + \beta_{e(NL)} \sqrt[n]{x_0} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu}\right)\right)}{2r\mu} & \text{for } x^* = x_0 \end{cases}$$

The proof is similar to Proposition 1.

4.3 Capacitated n th root demand (M_2)

In previous section, we have studied the problem with a non-linear demand function (precisely, n th root). In this section, we consider the same problem with considering one more constraint on demand function. As we mentioned in the beginning of this chapter, our colleagues found out that demand can not go beyond a certain value. Thus, we are introducing a new constraint that ensures demand function cannot go beyond a predetermined level. Therefore as already said previously, we define a new parameter, η such that $0 \leq \eta \leq 1$. We consider that carbon emission improvement can at maximum attract ηA more customers. Thus for a given p , the demand non-linearly (for $n \geq 2$) increases in CEI up to $(1 + \eta)A - \alpha_p p$, and after that, it is constant. Since improving carbon emission level after that point just increases the cost, it is logical that the optimal solution can not be after that. Therefore, we can introduce the constraint such that carbon emission improvement has a new upper bound, x_l , obtained from this inequality: $\beta_{e(NL)} \sqrt[n]{x} \leq \eta A$. Figure 4 shows the capacitated demand function under given price.

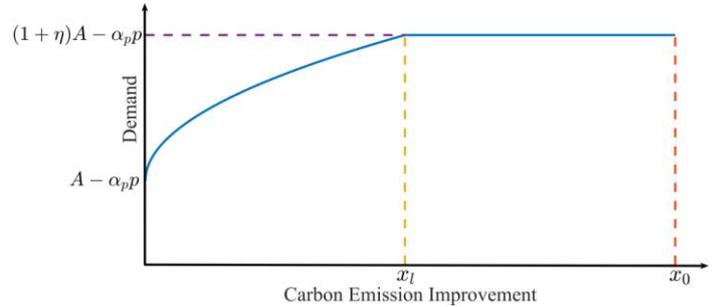


Figure 4. Capacitated n th root demand for a given price

The mathematical model of described problem is presented in the following:

$$\max_{x,p,S} \pi = (p - (c + bx^2))\lambda - hS \quad (12)$$

$$\text{subject to} \quad \psi_0 = \frac{\lambda}{\lambda + \mu S} \leq r \quad (13)$$

$$\beta_{e(NL)}\sqrt[n]{x} \leq \eta A \quad (14)$$

$$\lambda = A - \alpha_p p + \beta_{e(NL)}\sqrt[n]{x} \quad (15)$$

$$x, p, S, \lambda \geq 0, \quad x \leq x_0.$$

Lemma 5. For any given value of x , the optimal order size and price are $S^* = \frac{1-r}{r\mu} (A - \alpha_p p + \beta_{e(NL)}\sqrt[n]{x})$ and $p^* = \frac{\alpha_p b x^2 + \beta_{e(NL)}\sqrt[n]{x} + A + \alpha_p (c + \frac{h(1-r)}{r\mu})}{2\alpha_p}$, respectively.

The proof is similar to Lemma 1 and 2. By substituting optimal order size and price by their expressions, the equivalent model with only one variable, x is presented in the following:

$$\max_{0 \leq x \leq x_0} \pi = \frac{(-\alpha_p b x^2 + \beta_{e(NL)}\sqrt[n]{x} + A - \alpha_p (c + \frac{h(1-r)}{r\mu}))^2}{4\alpha_p} \quad (16)$$

subject to

$$x \leq \left(\frac{\eta A}{\beta_{e(NL)}} \right)^n$$

$$\lambda = \frac{-\alpha_p b x^2 + \beta_{e(NL)}\sqrt[n]{x} + A - \alpha_p (c + \frac{h(1-r)}{r\mu})}{2}$$

Proposition 3. The optimal carbon emission improvement for capacitated n th root demand model (M_2) is $x^* = \min \left\{ \left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}, \left(\frac{\eta A}{\beta_{e(NL)}} \right)^n, x_0 \right\}$.

The proof is similar to Proposition 2. $\left(\frac{\eta A}{\beta_{e(NL)}} \right)^n$ presents the carbon emissions improvement's upper bound (Equation 14).

4.4 Capacitated linear demand (M_3)

In this section, we study the benchmark problem with capacitated traditional linear demand function. Like previous section, we consider that the retailer cannot keep getting more costumers than ηA . So, after some point (x_l), the demand is constant as carbon emission improvement is increasing. The figure 5 shows the demand function.

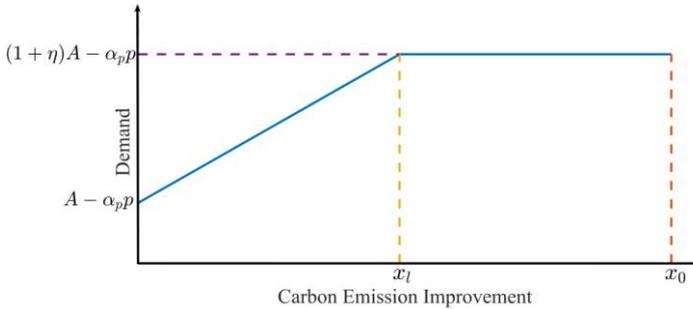


Figure 5. Capacitated linear demand for a given price

The demand function is linearly increasing in carbon emission improvement until x_l . It is constant after that. The mathematical model of the stochastic problem is presented in the following:

$$\max_{x,p,S} \pi = (p - (c + b x^2))\lambda - h S \quad (17)$$

subject to

$$\psi_0 = \frac{\lambda}{\lambda + \mu S} \leq r \quad (18)$$

$$\beta_{e(LC)} x \leq \eta A \quad (19)$$

$$\lambda = A - \alpha_p p + \beta_{e(LC)} x \quad (20)$$

$$x, p, S, \lambda \geq 0, \quad x \leq x_0.$$

We skip model explanation since it is similar to benchmark model. The optimal solution is presented in the following.

Lemma 6. For any given value of x , the optimal order size and

price are $S^* = \frac{1-r}{r\mu} (A - \alpha_p p + \beta_{e(NL)}\sqrt[n]{x})$ and $p^* = \frac{\alpha_p b x^2 + \beta_{e(NL)}\sqrt[n]{x} + A + \alpha_p (c + \frac{h(1-r)}{r\mu})}{2\alpha_p}$, respectively.

Proposition 4. The optimal carbon emission improvement for capacitated linear demand model (M_3) is $x^* = \min \left\{ \frac{\beta_{e(LC)}}{2\alpha_p b}, \frac{\eta A}{\beta_{e(LC)}}, x_0 \right\}$.

The proof is similar to Proposition 1. $\frac{\eta A}{\beta_{e(LC)}}$ presents the carbon emissions improvement's upper bound (Equation 19).

5 ANALYSIS AND INSIGHTS

In this section, we provide some important insights. First, we start to provide insights concerning to the optimal retailer's profit under different demands functions. As we mentioned before, the main idea of this paper is that the linear demand is not sufficient and realistic (Palacios-Argüello et al., 2020) and we consider that a general n th root demand function with a cap is more appropriate. For our numerical experiments, we consider the square root cap model to be our reference, and compare with other models. Then, we compare the optimal solutions. Finally, we provide the effect of n on the optimal solutions.

Table 1 presents the parameters that are used to compare the gap between square root cap demand ($n = 2$), as the reference, and other demands' (square root, linear, and linear cap) profits. To present a comprehensive result, we vary the parameters to consider a large number of combinations.

Table 1. Parameters

Parameter	Value	Parameter	Value
A	1200:100:2000	c	10:5:20
α_p	4:1:8	h	1:1:3
β_e	3:1:6	b	0.005:0.005:0.015
μ	30:10:50	r	0.05
x_0	100	η	0.1:0.3

We consider a variation of parameters, which are presented in Table 1, that leads to 7.2K different combinations. The gap between the optimal profit of the square root cap model (π_{M_2}) and other models ($\pi_{M_{0,1,3}}$) is calculated as $\frac{(\pi_{M_2} - \pi_{M_{0,1,3}}) * 100}{\pi_{M_2}}$. The results are presented in Table 2.

Table 2. Gap between profits

Profit	Linear (M_0)	Linear Cap (M_3)	Square Root (M_1)
$\eta = 0.02$	-11.87	0.00	-25.36
$\eta = 0.04$	-7.41	0.00	-20.35
$\eta = 0.06$	-3.39	0.00	-15.84
$\eta = 0.08$	0.003	0.003	-12.00
$\eta = 0.10$	2.8	0.21	-8.88
$\eta = 0.12$	4.97	0.61	-6.44
$\eta = 0.14$	6.63	1.24	-4.56
$\eta = 0.16$	7.93	1.71	-3.35
$\eta = 0.18$	8.96	2.28	-2.41
$\eta = 0.19$	9.39	2.60	-2.03
$\eta = 0.20$	9.73	3.15	-1.64
$\eta = 0.30$	12.18	6.49	-0.28

As Table 2 shows, with respect to profits, the capacitated linear demand is the closest approximation to the capacitated square root demand when η is less than 0.18, while, the square root demand provides better approximation when $\eta \geq 0.19$.

Now we are going to compare the decision variables (x^* , p^* and S^*) obtained from models M_0 , M_3 and M_1 to those obtained with model M_2 . The gap between the decision variables (DV)

of the square root cap model (DV_{M_2}) and other models ($DV_{M_{0,1,3}}$) is calculated as $\frac{(DV_{M_2} - \pi_{DV}) * 100}{DV_{M_2}}$.

Let us first consider x^* (Table 3), then p^* (Table 4) and finally S^* (Table 5).

Table 3. Gap between x^*

x^*	Linear (M_0)	Linear Cap (M_3)	Square Root (M_1)
$\eta = 0.02$	-117.11	0.00	-91.20
$\eta = 0.04$	-28.52	0.00	-22.05
$\eta = 0.06$	-12.12	0.00	-9.24
$\eta = 0.08$	-6.39	-0.01	-4.77
$\eta = 0.10$	-3.76	-0.04	-2.72
$\eta = 0.12$	-2.36	-0.07	-1.65
$\eta = 0.14$	-1.55	-0.10	-1.03
$\eta = 0.16$	-1.11	-0.12	-0.69
$\eta = 0.18$	-0.83	-0.14	-0.46
$\eta = 0.19$	-0.72	-0.15	-0.38
$\eta = 0.20$	-0.61	-0.16	-0.30
$\eta = 0.30$	-0.31	-0.25	-0.05

Table 3 shows that the product obtained with a capacitated linear demand has the closest optimal carbon emission improvement approximation with respect to the capacitated square root demand when η is less than 0.20. While, the product obtained with a square root demand provides better approximation (closer carbon emission improvement) to the product obtained with the capacitated square root demand when $\eta \geq 0.20$.

Table 4. Gap between optimal prices

p^*	Linear (M_0)	Linear Cap (M_3)	Square Root (M_1)
$\eta = 0.02$	-17.87	0.00	-18.65
$\eta = 0.04$	-15.70	0.00	-16.47
$\eta = 0.06$	-13.52	0.00	-14.27
$\eta = 0.08$	-11.31	-0.08	-12.04
$\eta = 0.10$	-9.19	-0.27	-9.91
$\eta = 0.12$	-7.28	-0.47	-7.99
$\eta = 0.14$	-5.59	-0.55	-6.30
$\eta = 0.16$	-4.50	-0.71	-5.13
$\eta = 0.18$	-3.56	-0.92	-4.08
$\eta = 0.19$	-3.15	-1.02	-3.61
$\eta = 0.20$	-2.60	-1.16	-3.08
$\eta = 0.30$	-1.25	-2.40	-0.83

As Table 4 shows, the product obtained with a capacitated linear demand has the closest optimal price approximation with respect to the product obtained with a capacitated square root demand when η is less than 0.20, while, the product obtained with a square root demand provides better approximation (closer price) to the product obtained with a capacitated square root demand when $\eta \geq 0.20$.

Table 5. Gap between optimal order sizes

S^*	Linear (M_0)	Linear Cap (M_3)	Square Root (M_1)
$\eta = 0.02$	-5.61	0.00	-11.82
$\eta = 0.04$	-3.48	0.00	-9.56
$\eta = 0.06$	-1.53	0.00	-7.49
$\eta = 0.08$	0.15	0.01	-5.71
$\eta = 0.10$	1.53	0.11	-4.25
$\eta = 0.12$	2.61	0.31	-3.10
$\eta = 0.14$	3.45	0.63	-2.20
$\eta = 0.16$	4.12	0.87	-1.62
$\eta = 0.18$	4.65	1.16	-1.17
$\eta = 0.19$	4.87	1.33	-0.99
$\eta = 0.20$	5.04	1.61	-0.80
$\eta = 0.30$	6.33	3.34	-0.14

Table 5 shows that the product obtained with a capacitated linear demand has the closest optimal order size approximation with respect to the product obtained with a capacitated square root demand when η is less than 0.19. While, the product obtained with a square root demand provides better

approximation (closer order size) to the product obtained with a capacitated square root demand when $\eta \geq 0.19$.

The results show that the linear demand that is well known in the literature is never the best one. However, the capacitated linear demand is a good approximation to the capacitated square root demand when η is small. The bright thing about capacitated linear demand is that we can easily adapt it to the previous more complex existing models in the literature.

We have considered for our numerical experiments a square root demand. Let us recall that models M_1 and M_2 have been developed for a general n th root demand. An interesting future work would be to compare results for different values of n . Intuitively we can claim that when n increases the influence of decreasing the emissions will decrease.

Proposition 5. In non-linear models (M_1 and M_2), increasing n leads to zero carbon emission improvements and in this case

$$p^* = \frac{A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p}$$

The n is a introduction of customers' desire for carbon emission improvement. Increasing n means the product's carbon emission improvement becomes less and less important to them. In other words, improving the carbon emission level attracts fewer and fewer new customers. So, improving one unit of carbon emission becomes less interesting from economical point of view. As a result, retailer's (or manufacturer's) motivation to offer greener products becomes fewer and fewer. Therefore for a given η , increasing n leads to lower $\left(\frac{\beta e(NL)}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}$, which is optimal solution without considering improvement constraint, and after some point, the constraint becomes non-binding at optimality.

6 CONCLUSION

In this paper, we consider general demands functions than the classical linear model. The demand functions are linearly decreasing in price, while, non-linearly increasing in carbon emission improvement. We consider the retailer's profit maximization problem with linear demand as the benchmark model and we reformulate it under new demand function (non-linear). The new problem is solved by an analytical approach. Then, we consider a constraint that allows demand to increase only to a certain level and we call it cap. The problems with new demands functions (i.e. linear cap and non-linear cap) are formulated and solved.

The numerical examples show that when demand's cap is small, the linear cap model provides the better approximation to square root demand than others, whereas, the square root model provides the better approximation to square root demand for the bigger cap.

The work that has been done in this paper can be extended in several ways. First, the new demand function can be used in product differentiation and competition models (e.g., reformulate the model that is presented by Asgari et al., 2020). It should be interesting to see the effect of cap on optimal strategies of retailers, especially on greenness, along with retailers' power (i.e. market share). Another extension can be the government regulations. The government regulations such as carbon tax, carbon cap, and etc. have their own effects on greenness decisions. However, it is interesting to see the effects of demand's cap alongside the government regulations' effect.

7 ACKNOWLEDGMENTS

This research has been supported by the French National Research Agency (ANR) through Project CONCLuDE.

8 REFERENCES

- Accenture. (2019). More than Half of Consumers Would Pay More for Sustainable Products Designed to Be Reused or Recycled, Accenture Survey Finds. Retrieved from <https://newsroom.accenture.com/news/more-than-half-of-consumers-would-pay-more-for-sustainable-products-designed-to-be-reused-or-recycled-accenture-survey-finds.htm>
- Asgari, E., Hammami, R., & Frein, Y. (2020). Retailing Competition for Substitutable Products in Greenness-and Price-Dependent Market. *International Conference in Information, Logistics and Supply Chain*.
- Bestseller. (2018). Taking action on climate change. Retrieved from <https://about.bestseller.com/news/taking-action-on-climate-change>
- Borin, N., Lindsey-Mullikin, J., & Krishnan, R. (2013). An analysis of consumer reactions to green strategies. *Journal of Product & Brand Management*.
- Ghosh, D., & Shah, J. (2012). A comparative analysis of greening policies across supply chain structures. *International Journal of Production Economics*, 135(2), 568–583.
- Ghosh, D., & Shah, J. (2015). Supply chain analysis under green sensitive consumer demand and cost sharing contract. *International Journal of Production Economics*, 164, 319–329.
- Hammami, R., Noura, I., & Frein, Y. (2018). Effects of Customers' Environmental Awareness and Environmental Regulations on the Emission Intensity and Price of a Product. *Decision Sciences*, 49(6), 1116–1155.
- Hong, Z., & Guo, X. (2019). Green product supply chain contracts considering environmental responsibilities. *Omega*, 83, 155–166.
- Hong, Z., Wang, H., & Gong, Y. (2019). Green product design considering functional-product reference. *International Journal of Production Economics*, 210, 155–168.
- Hovelque, V., & Bironneau, L. (2015). The carbon-constrained EOQ model with carbon emission dependent demand. *International Journal of Production Economics*, 164, 285–291.
- Liu, G., Yang, H., & Dai, R. (2020). Which contract is more effective in improving product greenness under different power structures: Revenue sharing or cost sharing? *Computers & Industrial Engineering*, 148, 106701.
- Liu, Z. L., Anderson, T. D., & Cruz, J. M. (2012). Consumer environmental awareness and competition in two-stage supply chains. *European Journal of Operational Research*, 218(3), 602–613.
- Ma, P., Wang, H., & Shang, J. (2013). Contract design for two-stage supply chain coordination: Integrating manufacturer-quality and retailer-marketing efforts. *International Journal of Production Economics*, 146(2), 745–755.
- Palacios-Argüello, L., Gondran, N., Noura, I., Girard, M.-A., & Gonzalez-Feliu, J. (2020). Which is the relationship between the product's environmental criteria and the product demand? Evidence from the French food sector. *Journal of Cleaner Production*, 244, 118588.
- Qin, J., Ren, L., & Xia, L. (2017). Carbon emission reduction

and pricing strategies of supply chain under various demand forecasting scenarios. *Asia-Pacific Journal of Operational Research*, 34(01), 1740005.

- Ramanathan, R., Ramanathan, U., & Hsiao, H.-L. (2012). The impact of e-commerce on Taiwanese SMEs: Marketing and operations effects. *International Journal of Production Economics*, 140(2), 934–943.
- Walmart. (2017). *Walmart Launches Project Gigaton to Reduce Emissions in Company's Supply Chain*. Retrieved from <https://corporate.walmart.com/newsroom/2017/04/19/walmart-launches-projectgigaton-to-reduce-emissions-in-companys-supply-chain>
- Xu, Y., Xie, H., & others. (2016). Consumer environmental awareness and coordination in closed-loop supply chain. *Open Journal of Business and Management*, 4(03), 427.
- Zhang, L., Wang, J., & You, J. (2015). Consumer environmental awareness and channel coordination with two substitutable products. *European Journal of Operational Research*, 241(1), 63–73.

9 APPENDICES

Appendix A. Proof of Lemma 1.

Since the objective function is linearly decreasing in S ($\frac{\partial^2 \pi}{\partial S^2} = 0$ and $\frac{\partial \pi}{\partial S} < 0$), the smallest possible S is the optimal stock. According to the constraint (Equation 2), $S \leq \frac{(1-r)\lambda}{r\mu}$. Therefore, the optimal value is $S^* = \frac{(1-r)}{r\mu} (A - \alpha_p p + \beta_{e(NL)} \sqrt[n]{x})$, which implies that service level constraint (Equation 2) is binding.

Appendix B. Proof of Lemma 2.

The first and second derivatives of π with respect to p are presented in the following. The second derivative is negative that means objective function is concave in p . Therefore, the root of first derivative (called p^{max}) maximizes the objective function.

$$\frac{\partial \pi}{\partial p} = -2\alpha_p p + \alpha_p b x^2 + \beta_{e(L)} x + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)$$

$$\frac{\partial^2 \pi}{\partial p^2} = -2\alpha_p < 0$$

$$\frac{\partial \pi}{\partial p} = 0 \Leftrightarrow p^{max} = \frac{\alpha_p b x^2 + \beta_{e(L)} x + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p}$$

To consider p^{max} as an optimal solution, we need to verify $p^{max} \geq 0$ and $\lambda(p^{max}) \geq 0$. Since $x > 0$, this is obvious that $p^{max} \geq 0$. Then we need to verify that $\lambda(p^{max})$ is greater than zero. To show demand positivity, we substitute p^{max} in the demand function, which is presented in the following:

$$\lambda(p^{max}) \geq 0 \Leftrightarrow \frac{-\alpha_p b x^2 + \beta_{e(L)} x + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p} \geq 0$$

$$\Leftrightarrow -\alpha_p b x^2 + \beta_{e(L)} x + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right) \geq 0$$

The first and second derivatives of $\lambda(p^{max})$ with respect to x are presented in the following.

$$\frac{\partial \lambda}{\partial x} = \frac{-2\alpha_p b x + \beta_{e(L)}}{2\alpha_p}$$

$$\frac{\partial^2 \lambda}{\partial x^2} = \frac{-b}{2} < 0$$

The second derivative is negative, which means the demand is concave in x . The demand function is increasing from zero ($\lim_{x \rightarrow +\varepsilon} \frac{\partial \lambda}{\partial x} > 0$) to the root of the first derivative. If the first root is higher than x_0 , it means that the demand function is positive in the feasible region if $A \geq \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)$. In the second possibility, the root of first derivative is in the feasible region, and the demand is decreasing after this root. Thus, also, we would need to make sure that demand is positive when $x = x_0$, which the condition is $A \geq \alpha_p b x_0^2 - \beta_{e(L)} x_0 + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)$.

As summary, while $A \geq \max \left\{ \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right), \alpha_p b x_0^2 - \beta_{e(L)} x_0 + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right) \right\}$ the demand function is positive in the feasible region. Therefore we can consider p^{max} as the optimal solution (i.e., $p^* = p^{max}$).

Appendix C. Proof of Proposition 1.

From Equation 6, the objective function (Equation 5) is equivalent to $\frac{\lambda^2}{\alpha_p}$. Since $\lambda \geq 0$, then maximizing λ is equivalent to maximizing $\frac{\lambda^2}{\alpha_p}$. We showed earlier that the demand function is concave in x . Therefore, the root of the first derivative (called x^{max}) maximizes the demand (equivalently, objective function).

If x^{max} is in the feasible region, then, it is the optimal solution.

We know that $0 \leq x \leq x_0$. Therefore, $x^* = \min \left\{ \frac{\beta_{e(L)}}{2\alpha_p b}, x_0 \right\}$.

By substituting $x^* \leq x_0$ into p^* , the optimal price is $p^* = \frac{\frac{3\beta_{e(L)}^2}{4\alpha_p b} + A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p}$.

In the same way, $S^* = \frac{(1-r) \left(\frac{\beta_{e(L)}^2}{4\alpha_p b} + A - \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right) \right)}{2r\mu}$.

Appendix D. Proof of Proposition 5.

Thanks to Propositions 2 and 3, we know that the x^* of non-

linear and non-linear cap models are $\min \left\{ \left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}, x_0 \right\}$

and $\min \left\{ \left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}, \left(\frac{\eta A}{\beta_{e(NL)}} \right)^n, x_0 \right\}$, respectively. It is clear

that n has no effect on x_0 . However, the other

elements, $\left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}$ and $\left(\frac{\eta A}{\beta_{e(NL)}} \right)^n$, depend on n . In the

following, we show values of these elements when n increases and goes to infinity.

$$\lim_{n \rightarrow \infty} \left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}} = 0$$

$$\lim_{n \rightarrow \infty} \left(\frac{\eta A}{\beta_{e(NL)}} \right)^n = \begin{cases} 0 & \text{for } \frac{\eta A}{\beta_{e(NL)}} < 1 \\ 1 & \text{for } \frac{\eta A}{\beta_{e(NL)}} = 1 \\ \infty & \text{for } \frac{\eta A}{\beta_{e(NL)}} > 1 \end{cases}$$

The upper bound's value in non-linear cap, $\left(\frac{\eta A}{\beta_{e(NL)}} \right)^n$, when n goes to infinity has different values that depends on different

situations. However, the other element's value, $\left(\frac{\beta_{e(NL)}}{2\alpha_p b n} \right)^{\frac{n}{2n-1}}$, goes to zero. Therefore, increasing n in non-linear model leads to decrease

x^* to zero. It is true for non-linear cap also (there is at least one element that goes to zero when n increases to infinity). The optimal price when optimal CEI is equal to zero is $p^* =$

$$\frac{A + \alpha_p \left(c + \frac{h(1-r)}{r\mu} \right)}{2\alpha_p}$$

Session 3.4 - Session spéciale 2 et 9 :
Ingénierie de l'obsolescence et
Ingénierie de la soutenabilité

Chairman : Marc Zolghadri - Daniel Brissaud

Sustainable urban factory: proposal of a viable scenario applied to Grenoble

SARAH PEIGNÉ¹, DAMIEN EVRARD²

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP
46 avenue Félix Viallet 38031 GRENOBLE CEDEX 01, FRANCE
sarah.peigne@grenoble-inp.org

² Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP
46 avenue Félix Viallet 38031 GRENOBLE CEDEX 01, FRANCE
damien.evrard@univ-grenoble-alpes.fr

Cet article présente une recherche d'un mode de production proche des lieux de consommation et soutenable. L'approche proposée consiste à remplacer une partie de la production massive et centralisée par une production centrée sur la satisfaction de besoins et localisée à proximité des lieux de consommations. Seules les zones urbaines sont considérées ici. Pour cette étude, le point de départ a été de se baser sur les applications possibles du plastique recyclé disponible sur place, dans les mines urbaines. À partir de là, la manière de fabriquer un ensemble de produits utilisables en milieu urbain et les volumes de production nécessaires pour atteindre l'équilibre économique tout en restant dans les limites du gisement de matière ont été considérées. Enfin, la pertinence environnementale du scénario proposé a été estimée grâce à une analyse du cycle de vie comparative de celui-ci avec un scénario industriel mondialisé. Il ressort de cette étude que le procédé d'impression 3D, malgré sa forte consommation électrique, est apparu comme ayant des impacts moindres sur l'environnement par rapport à un procédé de moulage par injection. Une étude économique succincte montre que pour être rentable, il est nécessaire de produire des objets ayant une valeur ajoutée suffisante pour compenser une vitesse de production plus faible.

This work wonders if a new means of production which meets the sustainable development goals is possible. The proposed idea is to switch a part of the massive and centralized production by a production centred on the satisfaction of needs and located close to the places of consumption. Only urban areas are included here. The starting point for this study was based on the possible applications of locally available recycled plastic in urban mines. From there, the way to manufacture a set of products usable in an urban environment and the production volumes required to achieve economic equilibrium while remaining within the limits of the material deposit were considered. Finally, the environmental relevance of the proposed scenario was assessed through a life cycle analysis comparing it with a globalized industrial scenario. The study shows the 3D printing process, despite its high electric consumption, looks like to have a lower environmental impact than the injection moulding process according to the assumptions made. The brief economic study shows that in order to be profitable, it is necessary to produce objects with sufficient added value to compensate for lower production speed.

Mots clés - Soutenabilité, Usine urbaine, Analyse de Cycle de Vie, Impression 3D.

Keywords – Sustainability, Urban's factory, Life Cycle Assessment, 3D printing.

1 INTRODUCTION

À partir des années 1960, des délocalisations d'usines vers des pays aux salaires faibles et aux réglementations environnementales moins contraignantes ont commencé. Ce mouvement, qui se poursuit dans cette décennie, permet de produire à moindre coût, mais au détriment de l'emploi local. Cela a également mené à une hyperspécialisation des usines, avec la nécessité de larges productions d'une même catégorie de produits, à un endroit donné. Aujourd'hui, au XXI^e siècle, des moyens existent pour rendre les industries moins polluantes et bruyantes. Par conséquent, il semblerait envisageable de ne plus autant séparer zone de production et zone de consommation. C'est ainsi que se pose la question

de réimplanter des usines urbaines à l'inverse de la tendance de la dernière décennie.

2 ÉTAT DES LIEUX DES USINES URBAINES DANS LE MONDE

En 2020, la chaîne de production est majoritairement globalisée c'est-à-dire qu'un produit avant d'arriver aux consommateurs parcourt des milliers de kilomètres. Ce transport a des impacts négatifs sur l'environnement (Juraschek, Becker, Thiede, Kara, & Herrmann, 2019). Aussi, de plus en plus de gens habitant en milieu urbain, en 2050, il est prévu que 80 % de la population européenne soit urbaine. Pour des raisons sociales, économiques et environnementales, les villes ne doivent plus être de simples places de résidences,

mais un lieu mixte entre travail, divertissement et lieu de vie (Burggräf, Dannapfel, Uelpenich, & Kasalo, 2019). L'usine urbaine vise à répondre à ces problématiques. Un système de production urbain ou « usine urbaine » est un site de production situé dans un environnement urbain. Le terme « usine » est défini comme un « regroupement local des facteurs de production pour la réalisation de l'ensemble ou d'une partie de la chaîne de valeur des biens ». La définition du terme « urbain » est plus controversée selon les disciplines scientifiques ou les pays d'application. Il n'existe pas de limite de nombre d'habitants, ou de surface de ville pour délimiter la notion d'urbain (Juraschek et al., 2019). De plus, ces usines « urbaines » pourraient être intéressantes dans des zones rurales éloignées des centres de productions et d'emploi ce qui permettrait une autonomie dans leur production et une relance économique durable. Par exemple Aquapôle, l'usine de traitement des eaux de l'agglomération Grenobloise peut être considéré comme une ébauche d'usine urbaine car une partie des boues produites lors des traitements est méthanisée à des fins énergétiques. Le biométhane produit est injecté dans le réseau de gaz de la commune du Fontanil-Cornillon. L'objectif est de produire 90% du gaz consommé par la commune (pour environ 2500 familles).

2.1 Les avantages et les défis des usines urbaines

L'un des principaux avantages des usines urbaines serait de produire près des consommateurs ce qui permettrait aux fabricants une totale transparence vis-à-vis de ceux-ci. Cependant, cet atout est à nuancer. Il est possible de produire loin des consommateurs et d'avoir une bonne traçabilité de la production et inversement. Cette proximité a également un impact positif sur la réduction des GES liés aux transports des lieux de productions vers les lieux de consommations (Juraschek et al., 2019). Aussi, des études montrent une plus grande facilité aux recrutements de cadres. Puis, l'utilisation des services urbains et ces infrastructures permettent de réduire certains coûts. Une étude montre que 2/3 des compagnies interrogées voient un potentiel dans la mise en réseau des sites urbains le long de la chaîne de valeur grâce à la microfabrication et aux prototypages rapides. Ils soulignent aussi que ce type de production encourage l'innovation (Burggräf et al., 2019).

Les défis liés à ces usines sont le manque d'espace urbain pour se développer, les émissions de bruit, le trafic engendré et les apparences physiques de l'usine (Juraschek et al., 2019). Il faudrait donc trouver des solutions pour mitiger ces nuisances. Certaines technologies de production trop sonore même avec une bonne isolation acoustique des bâtiments ou qui demande une grande surface de bâtiments devront être compensées par une autre manière de répondre au besoin qui soit plus adapté à l'environnement urbain. Par exemple, les papeteries seraient difficilement implantables en milieu urbain, car il faut une surface suffisamment grande pour accueillir une machine à papier, pour stocker la production, de l'eau à proximité et il y a un désagrément sonore important. Le besoin en papier/cartons peut être compensé partiellement par l'utilisation du numérique ou en limitant l'emballage des produits fabriqués. Il faudra importer le reste dans la zone d'habitation en essayant de limiter les nuisances sonores et environnementales du transport par exemple avec des camions au gaz naturel, ou en utilisant le fret ferroviaire ou encore fluvial.

2.2 Usines urbaines liées avec les mines urbaines

La production de déchets solides est un problème majeur en milieu urbain. L'élimination incontrôlée des déchets, l'absence de la collecte des déchets solides municipaux dans les quartiers informels, les bords de route, les rivières et les égouts contaminés ou bouchés par des déchets, ou les déchets toxiques éliminés de manière inappropriée sont des situations courantes dans de nombreuses villes particulièrement dans les pays en développement (Gutberlet, 2015). Cependant, de plus en plus de déchets sont reconnus comme une ressource susceptible de maintenir les moyens de subsistance et de générer des profits (Gutberlet, 2015 ; Savini, 2019). Les solutions actuelles (mise en décharge, incinération, recyclage) sont coûteuses et ne sont pas exemptes de risques pour l'environnement. Les solutions de réduction et évitement de la production de déchets ainsi que la récupération des ressources pour la réutilisation et le recyclage sont encore embryonnaires et très insuffisantes (Gutberlet, 2015).

L'exploitation minière urbaine s'étend de l'exploitation des décharges au processus de récupération des composés et des éléments de tout type et de déchets, y compris ceux issus du bâtiment, des infrastructures, des industries, des milieux environnementaux recevant des émissions anthropiques et les produits (en et hors d'usage). Les matériaux stockés peuvent représenter une source importante de ressources avec des concentrations souvent comparables ou excédant les sources naturelles. La figure 1 montre les différents flux de matériaux possibles. Pour obtenir un bénéfice économique, les mines urbaines se concentrent majoritairement sur les déchets électroniques avec une concentration élevée en métal et éléments rares. Mais les modèles économiques nécessitent d'être clarifié pour quantifier le potentiel bénéfice économique de chaque ressource (Brunner, 2011). Un cadre législatif européen qui faciliterait la récupération est pour l'instant inexistant (Cossu & Williams, 2015) et (Gutberlet, 2015).

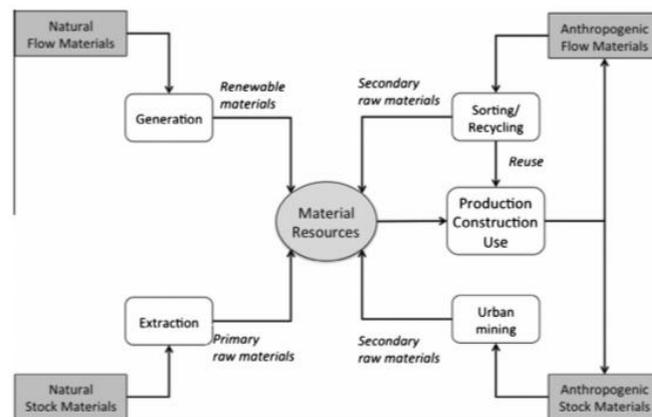


Figure 1 Description d'une mine urbaine versus une mines naturelle pour créer une nouvelle matière première (d'après Cossu & Williams, 2015).

Les mines urbaines sont également un moyen de mobiliser les habitants des villes. La notion de « prosommateur » rend compte de ce changement. Les individus sont « plus que des utilisateurs ou des consommateurs » : ils deviennent les ressources nécessaires à leurs propres pratiques de consommation. Le « prosommateur » appartient à un système économique dans lequel les deux processus de production et de

consommation, autrefois distincts, se chevauchent dans le temps et dans l'espace (Savini, 2019). Les mines urbaines ont une contribution à jouer pour réindustrialiser la ville (Brunner, 2011).

3 MODELISATION D'UN SCENARIO D'USINE URBAINE A GRENOBLE QUI EXPLOITE UNE MINE URBAINE

3.1 Le choix d'un gisement de plastique urbain

Le but de cette étude est de proposer un scénario d'usine urbaine pour en connaître les impacts environnementaux par rapport à une usine actuelle pour une gamme d'objet. Il a été décidé de partir d'un gisement de matière plastique disponible en milieu urbain dans l'objectif d'exploitation d'une mine urbaine et de dérouler la chaîne de production jusqu'aux fonctions des objets réalisables avec cette matière première. Cette usine a été imaginée afin d'estimer l'intérêt de l'usine urbaine pour satisfaire les besoins spécifiques sur un territoire fixé. Son objectif n'est pas de produire pour créer un besoin et un désir pour continuer la société de consommation qui est non soutenable dans le cadre d'une économie circulaire. Mais, elle se base sur les besoins des habitants proches et les mines urbaines présentent, actuellement le plastique est un déchet particulièrement abondant dans les villes. De plus, le matériau plastique propose des avantages non négligeables comme la légèreté, de bonnes propriétés mécaniques et une multitude d'applications.

3.2 La proposition d'usine urbaine

Le scénario est le suivant : les habitants et commerçants du quartier apportent de la matière première, c'est-à-dire du plastique. Cette matière est triée selon les différents types de polymères et par couleurs puis lavée. Puis, le matériau est remis sous forme de filament pour permettre l'impression 3D selon la technologie FDM (*Fused Deposition Modelling*) comme le montre l'arbre de processus. Cette usine urbaine se base principalement sur l'idée de « zero waste lab » à Thessaloniki et du mouvement « Precious Plastic ». Le mouvement « Precious Plastic » est une idée collaborative qui propose des machines et des tutoriels pour créer son atelier de recyclage alternatif. Il peut donner naissance à des lieux comme le « zero waste lab ». Ce « zero waste lab » est un espace à Thessaloniki qui comprend un robot pour imprimer en 3D à partir des matériaux plastiques rapportés par les citoyens. C'est également un lieu d'apprentissage pour comprendre le recyclage du plastique, les notions d'économie circulaire.

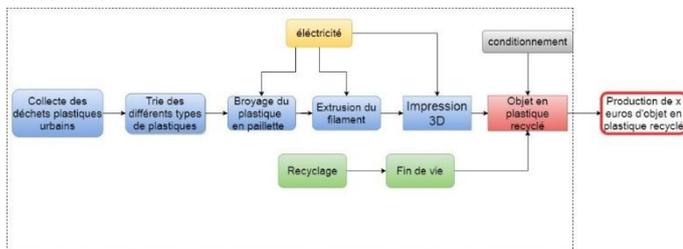


Figure 2 : Arbre de processus de l'usine urbaine modélisée

Ce scénario d'usine urbaine (Figure 2) sera comparé à une usine dite classique qui fabrique des objets en plastique recyclé avec le procédé d'injection moulage.

3.3 Étude économique et choix des produits fabriqués

3.3.1 Coût de fonctionnement de l'usine urbaine

L'usine urbaine aussi bien que l'usine classique afin d'exister, se doit de produire suffisamment pour être viable économiquement. Le coût de fonctionnement global de l'usine urbaine a été estimé à 30 000 € en prenant en compte les dépenses (Tableau 1).

Tableau 1 : Estimations du coût de fonctionnement de l'usine urbaine.

Dépense	Prix d'achat	Durée de vie (an)	Prix/an	Spécifications
Location bâtiment	84 K€		84 K€	13,9 €/mois/m ² à Grenoble S = 500 m ²
Imprimante 3D	44 K€	8	5 56 €	Funmat pro HT
Broyeuse	2 K€	8	285 €	Shredder Pro de Precious Plastic
Extrudeuse	5 K€	8	731 €	JM 35
Électricité /kg			59€	0,17 €/kWh en 2020 à Grenoble
PET vierge pellets/kg			2 €	Si besoin d'un complément
1 employé			30 K€	Salaire brut moyen en France
Supplément			10 K€	
Total /kg de plastique			59 €	
Total fixe			130 K€	

Pour atteindre ce coût de fonctionnement, les usines devraient produire un nombre assez élevé d'objets comme l'explique le tableau ci-dessus. Le prix unitaire indiqué dans le tableau est celui trouvé comme prix de vente pour des objets similaires fabriqués par injection moulage en plastique et en France. De plus, le coût par unité en fonction de la taille du groupe pour les deux technologies montre que pour un groupe de production de 32 928 objets (Tableau 2), le moulage par injection est bien moins cher avec un coût par unité d'environ 30 euros selon les données de la base du logiciel CES EDUPAC. Les objets devaient répondre au critère d'être produit grâce aux 2 procédés (impression 3D et injection moulage). Les objets ont été choisis de sorte à avoir une valeur ajoutée suffisamment grande pour la production en petite quantité. Ils devaient être en majorité imprimables en une seule pièce. Cette liste évoluera avec les besoins des habitants grenoblois.

Tableau 2 : Nombre d'objets à produire.

Vente	Prix unitaire	Nombre items	Prix total	Source
Monture Lunette	85 €	16 110	13 691 K€	(Fitzframes, 2020)
Bancs urbain	159 €	5988	952 K€	Banc plastique recyclé extérieur (« Leader equipment, » 2020)
Jardinière	1 348€	1878	2 531 K€	Bac à oranger (« Leader equipment, » 2020)
Table pique-nique urbaine	800€	7953	6 362 K€	Table pique-nique plastique azur (« Leader equipment, » 2020)
Poubelle tri urbaine	959 €	310	297 K€	Poubelle de tri sélectif en plastique recyclé (« Leader equipment, » 2020)
Chaussure empeigne	85 €	689	58 K€	(Cabell, 2020)
Total	3 436€	32 928		

3.3.2 Temps de production et masse de plastique

Ensuite, le temps de production et la masse de plastique nécessaire a été estimé. Le logiciel « Cura » permet d'imprimer des pièces en 3D à partir d'un fichier STL de l'objet trouvé en libre accès. Il a été utilisé pour les tests et les estimations de temps d'impression des objets avec la machine FDM « Anycubic 4 max ». L'estimation de temps et de masse reste dans le même ordre de grandeur pour l'ensemble des imprimantes 3D. L'échelle étant souvent plus petite que la taille réelle de l'objet, la masse et le temps d'impression ont été ajustés avec une règle de trois sur la face la plus grande (Tableau 3).

Tableau 3 : Masse de plastique et temps de production pour l'usine urbaine.

Objets	Masse plastique /an [T]	Nombre bouteille plastique /an	Temps FDM/ item [h]
Banc	42	1 249 052	118
Table pique nique	90	2 643 931	4
Poubelle de tri	0,3	827	3
Lunette	0,2	7088	1
Jardinière	2	56 988	6
TOTAL	135	4 millions	131

La source principale de PET est les bouteilles plastiques. Sachant que la consommation de bouteilles plastiques est de 96 bouteilles par personne par an en 2018 selon l'association « Zero Waste France » et il y a 443 000 habitants en 2016 dans l'agglomération grenobloise, le nombre de bouteilles consommées est de 42 millions de bouteilles. Avec l'hypothèse qu'il faut 22 bouteilles PET pour produire 750 g de filaments d'après la société « Better Future Factory » (« Refil, le filament recyclé, »), ce qui donne un rendement de quasi 100 %, car les pertes sont négligeables lors du broyage du plastique et l'extrusion du filament. La production de cette liste d'objets demanderait 4 millions de bouteilles. L'agglomération grenobloise fournit donc un gisement largement suffisant pour la production des objets sur un an. La quantité avec les objets cités ci-dessous est de 135 tonnes de plastique PET, sachant qu'actuellement il y a 1 300 T de plastiques (tous types confondus) recyclés dans l'agglomération grenobloise. Les objets imprimés en FDM sont majoritairement vides donc ils sont plus légers que ceux fabriqués par injection moulage.

Les temps d'impression FDM sont estimés long avec les machines disponibles actuellement avec un fil d'1,75mm qui est inadapté à l'impression de grands objets. C'est une des limites de la technologie par exemple le tableau 3 montre que pour un banc, il faudrait 118h d'impression.

Le temps de broyage et d'extrusion du plastique, respectivement : 8,8 kg PS/h et 23 kg/h, sont plus rapides par rapport à l'impression FDM. Le temps d'impression en 3D, qui est le temps limitant pour l'usine urbaine par rapport aux autres étapes, est difficile à estimer sans tests. Les robots d'impression 3D FDM à grande échelle sont encore peu répandus sur le marché. Ainsi, un parc machine de 87 imprimantes 3D serait nécessaire pour répondre à la quantité nécessaire de production pour être rentable. Ce nombre prend comme hypothèse un fonctionnement continu de l'usine soit 365 jours par an, 24 h/24 h. Des machines plus performantes sont nécessaires sans quoi cette technologie ne pourrait pas répondre à la problématique. L'autre solution est de produire moins d'objets, mais des items à encore plus grandes valeurs ajoutées. Pour les autres machines nécessaires à la production urbaine, il faudrait uniquement 2 broyeuses et 1 extrudeuse performante pour produire le filament nécessaire si le fonctionnement de ces 2 machines est continu dans l'année. Le procédé d'injection moulage est un procédé qui

permet de produire plus rapidement des objets une fois le moule fait par rapport à la technique de fabrication additive FDM. Pour l'estimation basse de la vitesse de production qui est de 60 objets/heure, le temps de production totale par injection moulage des 32 928 objets nécessaire est de 22 jours avec une seule machine en fonctionnement en continu

4 EVALUATION DES IMPACTS ENVIRONNEMENTAUX

4.1 Objectifs et champ de l'étude

La modélisation des usines a été faite avec le logiciel d'analyse cycle de vie «SimaPro». La méthode d'impact principal choisi est la méthode «Impact 2002 +». La base de données mondiales «Ecoinvent» a été utilisée. L'unité fonctionnelle est : «produire pour 130 milles euros (le coût de fonctionnement de l'usine urbaine) en 2020 d'objets en plastique (PET) recyclé pour la ville de Grenoble».

L'étude va du berceau à la tombe en excluant la phase d'utilisation qui est jugée identique que l'objet soit produit de manière urbaine ou classique. Soit, la récupération du plastique, sa transformation en pellet ou filament, sa mise en forme, l'emballage et la distribution de l'objet et pour finir sa fin de vie ont été considérés.

La frontière géographique est celle des déchets plastiques produits à Grenoble, un centre de tri à Grenoble, une production en France pour l'usine classique et à Grenoble pour l'usine urbaine, une livraison de l'objet à Grenoble. Les déchets plastiques sont ramassés en vrac sans sacs plastiques supplémentaire soit par le système de collecte sélective de l'agglomération grenobloise qui utilise des camions au gaz naturel et au pétrole, soit pour l'usine urbaine : apport volontaire des citoyens et collecte à vélo. Le type de plastique étudié est le polyéthylène téréphthalate (PET).

L'étape du lavage des plastique avant de le réutiliser n'a pas été modélisée. L'impact environnemental n'est pas négligeable, car cette étape nécessite de l'eau, du détergent et une machine à laver. Il a été pris comme hypothèse que pour les 2 scénarios modélisés, l'impact lié au lavage serait sensiblement le même. Il est difficile de faire des économies d'échelles pour cette étape. Ces raisons ont mené à ne pas modéliser le lavage du plastique. Notons également, qu'il est aussi possible de demander au Grenoblois pour l'usine urbaine de demander de laver leur plastique avant le recyclage dans ce cas l'étape du lavage devrait être modélisé pour comparer un lavage individuel versus industriel. L'impact environnemental du lavage individuel sera peut-être encore plus grand à cause de consommation d'eau plus élevée ou de l'utilisation de détergent non écologique.

La production d'équipements comme les machines sauf l'imprimante 3D ou la construction du site a été exclue de l'étude. Ces investissements sont du long terme donc leur impact est négligeable. De plus, les données pour leur modélisation sont manquantes.

Tableau 4 : Les hypothèses prises pour réaliser l'ACV

Etape du cycle de vie	Usine urbaine	Usine classique
Recyclage du plastique	- Matériaux : PET - Phase de lavage exclus	
	-Transport vélo ou à pieds	-Camion EURO4
Fabrication des objets	-Mix énergétique Grenoblois - FDM	-Mix énergétique Français -Injection moulage
Distribution	-Emballage : cartons -Retrait à l'usine	-Emballage : cartons -Transport Camion
Fin de vie	-Traitement des déchets Français	

4.2 Résultats de l'évaluation des impacts environnementaux

Le graphique représente les impacts relatifs en pourcentage, en fonction des étapes du cycle de vie. Il montre que les impacts sont plus importants dans le cadre de l'usine classique avec une différence plus ou moins importante selon les catégories de dommages. La Figure 3 montre que pour les 4 catégories de dommages (santé humaine, qualité des écosystèmes, changement climatique et ressources), le procédé de transformation, les barres en diagonales sur le graphique, contribue à l'impact principal. Le packaging qui a été modélisé pour les 2 modèles comme peu invasifs ne représente qu'un faible impact. Le transport n'apporte pas un grand impact dans les différentes catégories pour l'usine classique. En effet, les distances prises sont françaises. Le transport aurait sûrement des impacts plus importants sur les changements climatiques, les ressources ou la santé si l'étude avait considéré dans ces hypothèses une production en Chine pour l'usine classique par exemple.

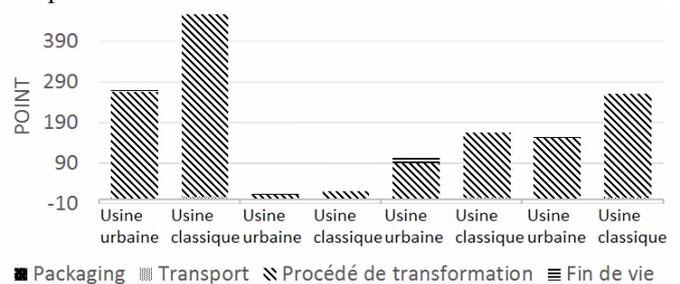


Figure 3 : Impacts environnementaux en point selon les étapes du cycle de vie

Pour permettre de faire ressortir les étapes du cycle de vie avec un impact élevé après les procédés de transformation, un autre graphique a été construit sans cette étape. Il est possible de remarquer que, la fin de vie joue un rôle important dans l'usine urbaine puisque le transport n'a aucun impact comme il est effectué soit à pieds ou à vélo cargo. Cependant, il sera difficile de diminuer l'impact de la fin de vie puisqu'elle dépend des consommateurs et des villes. De plus, la fin de vie a été modélisée par un processus déjà présent dans la base de

données Ecoinvent. Ce n'est pas forcément actualisé aux normes actuelles.

La production de PET a la contribution la plus élevée dans les impacts de production. En effet, dans Ecoinvent le PET recyclé n'existe pas en matière première, le PET vierge a été pris, alors que recycler permet de diviser les impacts de la production et l'extraction de la matière. Ensuite, le processus d'injection-moulage contribue fortement aux impacts environnementaux. Ainsi, produire de manière urbaine permettrait bien une production plus durable. Il semble justifié de chercher de nouveaux modes de production. Pour les deux scénarios, c'est l'électricité et le traitement des déchets des objets fabriqués qui contribuent majoritairement aux impacts. En effet, l'impression 3D est un procédé consommateur en énergie avec une consommation comprise entre 23-246 kWh/kg pour une imprimante FDM (Cerdas, Juraschek, Thiede, & Herrmann, 2017). Une machine d'injection moulage consomme moins d'électricité entre 0,11-5,82 kWh/Kg (Cerdas et al., 2017). L'utilisation d'hydroélectricité pour l'usine urbaine permet déjà de limiter les impacts par rapport à des énergies non renouvelables comme le charbon ou le nucléaire. Différents calculs d'impacts ont été faits en changeant le mix énergétique dès lors, les résultats varient grandement. L'usine urbaine avec le même mix énergétique français que celui de l'usine classique donne des impacts environnementaux bien plus élevés. Les impacts environnementaux entre les 2 scénarios deviennent similaires. L'utilisation de panneaux solaires sur le toit de l'usine pour permettre une usine indépendante des structures électriques non urbaines a été pensé. D'après, l'outil Métrosoleil conçu par la métropole de Grenoble, la production électrique maximum sera de 115 160 kWh/an. Avec la surface de l'usine qui est par hypothèse de 500 m² située au 46 avenue Viallet, soit l'emplacement du laboratoire G-SCOP. La consommation annuelle électrique pour la production a été estimée à 53 197 297 kWh. Les panneaux solaires sur l'ensemble du toit ne suffiront largement pas à produire l'ensemble des besoins énergétiques. Les petites éoliennes urbaines sont actuellement seulement en recherche et développement. La géothermie n'est pas suffisante à Grenoble pour la production d'électricité puisqu'elle est de type « basse énergie donc utilisable seulement pour le chauffage d'un bâtiment ou d'un circuit d'eau seulement.

5 CONCLUSION ET PERSPECTIVES

Cette étude est une recherche exploratoire d'un modèle d'usine urbaine qui permettrait de réduire les impacts environnementaux de produits usuels par une production et l'utilisation de matières premières et secondaires locales. L'évaluation des impacts environnementaux a montré que ce mode de production locale pourrait contribuer à s'approcher de la soutenabilité. Il est à noter que cela dépendrait également de la demande des consommateurs et du modèle de consommation d'usage employé. Les usines urbaines proposent une autre méthode qui serait de produire proche des lieux d'habitations et de petites quantités adaptés aux besoins des citoyens et à la zone de production. En effet, les technologies de production de masse, telle que l'injection moulage, sont impossibles d'être reproduites techniquement

dans une zone réduite comme les villes, voir des zones peu denses en population. Les procédés additifs, comme l'impression 3D, peuvent être une solution durable de production selon certaines conditions. Son impact environnemental est néanmoins non négligeable comme le montre l'analyse de cycle de vie réalisée. Cependant, la technologie reste lente et énergivore actuellement ce qui est une limite à son développement pour un usage industriel. Ainsi en ville, la production de masse est exclue, les objets à haute valeur ajoutée sont plus adaptés au terrain. Le plastique peut être considéré aujourd'hui comme une véritable mine urbaine d'où son étude dans ce projet. Néanmoins avec la directive européenne de la fin du plastique à usage unique adopté en mars 2019, cela pourrait rapidement évoluer. Malgré ce changement, les déchets plastiques resteront longtemps avant de disparaître complètement et d'autres matériaux sont utilisables avec la technologie d'impression 3D. Il est à noter que les propriétés mécaniques du plastique recyclé sont potentiellement altérées par rapport au plastique vierge. Il faudrait donc rajouter du plastique vierge, mais les études à ce sujet sont peu nombreuses et peu précises. Des tests mécaniques sur différentes pièces en plastique vierge et recyclé seraient intéressants à effectuer. L'ACV du cas modélisé montre que la contribution majoritaire est l'étape de production pour l'usine classique ou urbaine. Mais, l'usine urbaine émet moins d'impact sur l'environnement avec les hypothèses et les frontières choisies. Une modélisation plus précise, en termes de mix énergétique, de distances de transport et de l'impact des finitions des objets, serait intéressante. Une usine urbaine avec un mix de machines divers pour produire un plus grand champ d'objet permettrait une plus grande rentabilité et une meilleure réponse aux besoins larges d'une communauté. Enfin, l'acceptabilité sociale des usines urbaines reste à prouver. Cette étude appliquée à Grenoble avec les quantités de plastique et la densité de population disponible serait sans doute transposable à d'autres villes ; l'aspect le plus important à étudier étant de déterminer les caractéristiques-clefs qui conditionneraient son implantation. D'ailleurs, au-delà de l'aspect urbain, l'aspect local semble être l'objet de recherche déterminant et des usines locales dans des endroits moins densément peuplés pourraient être pertinentes, même si le périmètre géographique serait alors étendu.

6 REFERENCES

- Better_future_factory. (2020). Let's make change. Tiré de <https://www.betterfuturefactory.com/>
- Brunner, P. H. (2011). Urban Mining A Contribution to Reindustrializing the City. *Journal of Industrial Ecology*, 15(3), 339-341. doi:10.1111/j.1530-9290.2011.00345.x
- Burggräf, P., Dannapfel, M., Uelpenich, J., & Kasalo, M. (2019). Urban factories: Industry insights and empirical evidence within manufacturing companies in German-speaking countries. *Procedia Manufacturing*, 28, 83-89. doi:10.1016/j.promfg.2018.12.014
- Cabell, O. (2020). Phoenix. Tiré de <https://olivercabell.com/pages/phoenix>

- Cerdas, F., Juraschek, M., Thiede, S., & Herrmann, C. (2017). Life Cycle Assessment of 3D Printed Products in a Distributed Manufacturing System. *Journal of Industrial Ecology*, 21(S1), S80-S93. doi:10.1111/jiec.12618
- Cossu, R., & Williams, I. D. (2015). Urban mining: Concepts, terminology, challenges. *Waste Manag*, 45, 1-3. doi:10.1016/j.wasman.2015.09.040
- Fitzframes. (2020). Tiré de <https://www.fitzframes.com/>
- Gutberlet, J. (2015). Cooperative urban mining in Brazil: Collective practices in selective household waste collection and recycling. *Waste Manag*, 45, 22-31. doi:10.1016/j.wasman.2015.06.023
- Juraschek, M., Becker, M., Thiede, S., Kara, S., & Herrmann, C. (2019). Life Cycle Assessment for the comparison of urban and non-urban produced products 26, 65, 32-37. doi:10.1016/j.procir.2017.04.009
- Leader equipment. (2020). Tiré de <https://www.leader-equipements.com/>
- Lumières-de-la-ville. (2019). INDUSTRIES ET VILLES SONT-ELLES INCOMPATIBLES ? Tiré de <https://lumièresdelaville.net/industries-villes-incompatibles/>
- Make it precious, start a business from plastic waste (2020). Tiré de <https://preciousplastic.com/>
- Refil, le filament recyclé.). Tiré de <https://www.happy3d.fr/fr/refil-le-filament-recycle/>
- Savini, F. (2019). The economy that runs on waste: accumulation in the circular city. *Journal of Environmental Policy & Planning*, 21(6), 675-691. doi:10.1080/1523908x.2019.1670048
- The_new_raw. (2020). Tiré de <https://thenewraw.org/>

(This paper is not included in this version, please contact authors for more information)

Évaluation des pratiques de gestion de l'énergie dans l'industrie forestière

THOMAS HERNÁNDEZ, NADIA LEHOUX, LOUIS GOSSELIN

DEPARTEMENT DE GÉNIE MÉCANIQUE, UNIVERSITÉ LAVAL
Québec, Canada

thomas-antonio.hernandez-gutierrez.1@ulaval.ca

Nadia.Lehoux@gmc.ulaval.ca, Louis.Gosselin@gmc.ulaval.ca

Résumé – Le secteur forestier est l'une des industries les plus énergivores dans plusieurs pays et bien qu'elle produise parfois sa propre énergie à partir de biomasse, elle continue d'émettre des quantités importantes de gaz à effet de serre. Pour cette raison, cette étude a analysé les pratiques de gestion de l'énergie (GE) pour cette industrie. Une revue de littérature a permis d'identifier différents cadres d'analyse de la GE et d'établir une structure d'évaluation basée sur ces méthodologies. L'étude du profil énergétique de l'industrie forestière a ensuite été réalisée à travers cinq dimensions : utilisation de l'énergie dans le secteur forestier, contrôle de la transition énergétique, incitatifs, structure organisationnelle et GE au niveau opérationnel. Les résultats de l'étude montrent que les programmes de GE et les investissements dans les technologies écoénergétiques ont le potentiel d'accélérer la transition énergétique de l'industrie forestière, visant à établir une économie sobre en carbone dans les années à venir. Ceux-ci se sont également avérés être des mécanismes importants pour réduire la consommation d'énergie de plusieurs entreprises de ce secteur.

Abstract – The forest sector is one of the most energy intensive consumers in many countries and although it often produces its own energy from biomass, it continues to emit significant amounts of greenhouse gas emissions. For this reason, this study analyzed energy management practices in the forestry sector. A literature review first led to the identification of different energy management frameworks and to the establishment of an evaluation structure based on these methodologies. The energy profile of the forest industry was then carried out through five dimensions: energy use in the forestry sector, energy transition control, incentives, organizational structure and energy management at the operational level. This study has shown that energy management programs and investments in energy efficient technologies have the potential to accelerate the energy transition of the forest industry that aims to establish a low carbon economy in the years to come. These have also proven to be important mechanisms for reducing the energy consumption of several companies in this sector.

Mots clés – Gestion de l'énergie, industrie forestière, pratiques de gestion de l'énergie, systèmes de production

1 INTRODUCTION

Partout dans le monde, les stratégies énergétiques sont au cœur des intérêts des gouvernements et des grandes organisations, car leur développement contribue à la lutte contre le réchauffement planétaire. Depuis 1980, les sources d'énergie sont devenues plus accessibles et interchangeables, mais combiné à la croissance démographique et économique, l'utilisation de l'énergie a contribué à une augmentation des émissions de carbone de la planète de 52% au cours des 25 dernières années (National Energy Board, 2019).

Dans ce contexte, l'industrie forestière apparaît comme une grande consommatrice d'énergie, en raison du sous-secteur des pâtes et papiers qui est l'une des cinq industries les plus énergivores au monde (Andersson & Thollander, 2019). Pour réduire son empreinte écologique, ce secteur doit donc s'appuyer sur des stratégies et des politiques énergétiques lui permettant d'atteindre les objectifs de sa transition énergétique, qui un peu partout dans le monde progressent vers une réduction de l'utilisation des sources d'énergie fossiles et la valorisation du potentiel énergétique de la biomasse forestière. Toutefois, l'industrie forestière doit aussi concevoir des alternatives lui permettant d'améliorer son efficacité

énergétique et sa rentabilité à court terme, en réduisant ses coûts d'exploitation liés au gaspillage d'énergie. À cet égard, la gestion de l'énergie (GE) et l'amélioration de l'efficacité énergétique sont des composantes qui doivent continuer à être développées (Ressources Naturelles Canada [RNC], 2019a) pour améliorer la compétitivité de l'industrie forestière et ainsi tendre vers un avenir énergétique durable.

Cet article vise à analyser dans une perspective d'ensemble, les pratiques énergétiques de l'industrie forestière et à établir les critères d'évaluation nécessaires pour permettre à cette industrie d'explorer ses enjeux énergétiques, ses meilleures pratiques de GE et ses opportunités d'amélioration. Pour atteindre cet objectif, une revue de la littérature a permis d'examiner la chaîne de valeur de l'industrie forestière, ses secteurs d'activités, sa situation énergétique actuelle et l'importance de la GE dans cette industrie. Par la suite, sur la base de différentes méthodologies et cadres d'analyse de la GE trouvés dans la littérature, des axes d'évaluation pertinents ont été établis et ont permis d'analyser le contexte énergétique de l'industrie forestière dans différents pays du monde. Les résultats de l'étude montrent que le niveau de maturité énergétique des industries forestières varie considérablement d'un pays à l'autre. De même, cet article identifie différents

vecteurs stratégiques importants favorisant la transition énergétique de cette industrie. Cette recherche révèle aussi que les technologies écoénergétiques, les systèmes de GE et la maintenance des équipements sont des facteurs clés permettant aux entreprises de ce secteur d'augmenter leur potentiel d'économie d'énergie de manière significative et durable.

Cet article décrit, à la section 2, la revue de la littérature pertinente à l'étude. La méthodologie de recherche est ensuite détaillée à la section 3. La section 4 présente l'analyse énergétique comparative de l'industrie forestière dans le monde. L'article se termine par une brève conclusion.

2 REVUE DE LA LITTÉRATURE

Cette recherche s'intéressant à l'industrie forestière, il apparaît d'abord nécessaire de se familiariser avec son mode de fonctionnement. D'un point de vue général, selon Sandberg *et al.* (2014), le secteur forestier peut être classé en deux activités principales, soit l'industrie forestière et l'industrie des produits du bois. D'une part, l'industrie forestière comprend la foresterie et l'exploitation forestière, la fabrication de pâtes et papiers, la fabrication de produits en bois et l'exploitation de la biomasse forestière. D'autre part, l'industrie des produits du bois est liée aux panneaux, aux meubles, aux poutrelles, etc. D'un autre côté, bien qu'il s'agisse d'un secteur à forte intensité énergétique, principalement en raison de l'industrie des pâtes et papiers, sa capacité à produire sa propre électricité à partir de la bioénergie a considérablement réduit sa dépendance aux combustibles fossiles, ce qui a entraîné une baisse constante de ses émissions de GES ces dernières années (RNCAN, 2020). Malgré tout, son rôle comme émetteur de GES reste majeur.

2.1 Industrie des pâtes et papiers

Les entreprises de ce sous-secteur transforment les matières végétales ligneuses en une grande variété de pâtes, papiers et cartons (Adès, 2010). De ce fait, les usines de pâtes et papiers intègrent une grande variété de procédés différents lors de la fabrication de leurs produits (voir Bajpai, 2016b). Ces procédés consomment une grande quantité d'énergie et les plus énergivores sont la mise en pâte et le séchage du papier. D'un autre côté, l'ère des médias électroniques a provoqué un effondrement rapide des marchés du papier journal, du papier d'impression et d'écriture (Ministère des Ressources naturelles et de la faune [MRNF], 2012). À l'échelle mondiale, plusieurs fermetures d'usines de papier d'impression et d'écriture ont eu lieu (RNCAN, 2019b). Pour améliorer sa rentabilité et sa compétitivité, l'industrie doit poursuivre ses efforts pour réduire ses coûts de fabrication. L'amélioration de l'efficacité énergétique est un moyen important pour y arriver, puisque l'énergie représente jusqu'à 15% de ses coûts opérationnels totaux (Bajpai, 2016a). Bien que des progrès en matière de conservation de l'énergie soient observés, la GE dans les systèmes de production demeure un élément clé à mettre en valeur. À cet égard, la GE doit être une composante présente dans les stratégies de développement de cette industrie.

2.2 Fabrication de produits du bois

Le sous-secteur de la fabrication des produits du bois est composé de différents domaines d'activité visant à transformer le bois en produits finis. Les principales activités de ce sous-secteur sont les scieries, les usines de panneaux et les usines de deuxième et de troisième transformation (MRNF, 2012). Les scieries sont l'axe central de ce sous-secteur, puisqu'elles sont en charge d'effectuer la première transformation de la matière première pour la transformer en bois de sciage ou bois d'œuvre. Après la transformation, la matière première traitée

est fournie aux usines de panneaux et aux usines de deuxième et de troisième transformation. Les scieries fournissent également des copeaux à l'industrie des pâtes et papiers et des sous-produits à l'industrie de la cogénération (MRNF, 2012). D'autre part, les scieries transforment les billes de bois en bois d'œuvre à l'aide de diverses opérations qui nécessitent des équipements qui consomment beaucoup d'énergie. De plus, l'énergie électrique utilisée représente une part importante de la consommation totale d'énergie et des coûts d'exploitation. Dans leur étude, Gopalakrishnan *et al.* (2012) mettent en relief que le manque de connaissances sur les avantages qu'une efficacité énergétique accrue pourrait apporter aux scieries entraîne des coûts opérationnels élevés.

2.3 GE dans l'industrie

La GE se concentre principalement sur l'amélioration de la performance énergétique d'une organisation, dans le but d'améliorer la productivité, de réduire l'empreinte carbone et d'éliminer les coûts associés au gaspillage d'énergie, tandis que les systèmes de GE sont des instruments permettant d'établir des politiques énergétiques et des objectifs énergétiques, ainsi que des processus et des procédures pour atteindre ces objectifs (ISO, 2018). Malgré le fait que les systèmes de GE sont peu coûteux, les entreprises se concentrent souvent uniquement sur la mise en œuvre de technologies écoénergétiques (Johansson et Thollander, 2018), oubliant que la GE dans les systèmes de production permet de conscientiser le personnel sur l'utilisation de l'énergie (Thollander *et al.*, 2020a). Dans ce contexte, Bunse *et al.* (2011) établissent qu'au niveau opérationnel, la GE est constituée de différentes phases où l'efficacité énergétique doit être mesurée, surveillée et fréquemment évaluée à l'aide d'indicateurs de performance énergétique (KPIs). De même, les auteurs soulignent l'importance des technologies de l'information et des normes (e.g. ISO 50001) comme incitatifs. Dans la littérature, plusieurs études dressent le portrait énergétique de différents secteurs industriels en se basant sur des dimensions clés de la GE. Hossain *et al.* (2020) ont analysé les meilleures pratiques énergétiques de l'industrie du ciment au Bangladesh, à travers un questionnaire qui a pris en compte plusieurs éléments clés de la GE. Une étude similaire a été menée par Hasan *et al.* (2018) dans l'industrie du fer et de l'acier dans le même pays. De leur côté, Abdelaziz *et al.* (2011) ont défini que l'efficacité énergétique industrielle peut être améliorée grâce à des pratiques de GE, la mise en œuvre de nouvelles technologies et l'élaboration de politiques énergétiques. Cependant, dans la littérature, il existe très peu d'études analysant les pratiques et les enjeux énergétiques de l'industrie forestière. À titre d'exemple, Lawrence *et al.* (2019) ont récemment analysé l'état de maturité énergétique de l'industrie suédoise des pâtes et papiers, en mettant l'accent sur les moteurs de changement et les barrières de la GE. Pour cette même industrie, Andersson et Thollander (2019) ont analysé le niveau de développement des indicateurs de performance énergétiques, où les résultats ont montré que le faible développement de ces indicateurs est dû au manque de ressources, de connaissances et de compétences. Dans la prochaine section, une description des grandes étapes qui ont permis de mener à terme la recherche est présentée.

3 METHODOLOGIE

Afin d'évaluer le profil énergétique de l'industrie forestière, le présent travail a été divisé en trois étapes. De façon générale, la recherche a suivi la stratégie de recherche du cercle herméneutique, car elle permet d'explorer de manière flexible

et itérative les aspects les plus importants du sujet à l'étude et d'ajuster le contenu de la recherche au fur et à mesure de la découverte de nouvelles informations. Par conséquent, les bases de données *Engineering Village*, *Web of Science*, *ScienceDirect* et *IEEE Xplore* ont servi de support à la revue scientifique. Dans la phase de collecte des données, l'étude a tenu compte des articles publiés de 2010 à 2020. Les références pertinentes tirées de ces articles, même datant des années précédentes, ont également été prises en compte.

La première étape a consisté à effectuer une revue de la littérature décrivant la chaîne de valeur de l'industrie forestière, analysant le contexte énergétique de ses sous-secteurs les plus importants et discutant de l'importance de la GE dans les systèmes de production. Par la suite, la deuxième étape a consisté à identifier les critères d'évaluation nécessaires à la réalisation de la recherche et à développer un cadre d'analyse qui a permis de dresser l'étude comparative énergétique de l'industrie forestière. Afin de déterminer les axes du cadre d'analyse utilisé dans le présent travail, l'équation de recherche a été définie comme suit : ("*energy management practices*" or "*energy management*") and "*industries*" and "*manufacturing*" and "*framework*". L'étude des articles trouvés s'est ensuite faite de manière inductive, c'est-à-dire que les axes du cadre d'analyse ont été déterminés après avoir examiné de manière exhaustive les articles sélectionnés. Par conséquent, à l'issue de cette analyse et en se basant sur les méthodologies et cadres de la GE utilisés par Hossain *et al.* (2020), Abdelaziz *et al.* (2011) et Bunse *et al.* (2011), il a été possible de déterminer cinq axes clés d'évaluation qui ont par la suite permis de dresser le portrait énergétique de l'industrie forestière : utilisation de l'énergie dans le secteur forestier, contrôle de la transition énergétique, incitatifs, structure organisationnelle et GE au niveau opérationnel. Dans la troisième étape de la recherche, en fonction du cadre d'analyse dégagé, l'étude comparative de l'état énergétique de l'industrie forestière a été effectuée en évaluant les deux sous-secteurs les plus énergivores (l'industrie des pâtes et papiers et l'industrie de la fabrication de produits du bois). Étant donné que les bases de données mentionnées précédemment contiennent peu d'information sur les meilleures pratiques de la GE dans l'industrie forestière, des rapports et publications du gouvernement, des études de cas de même que des sites web d'organisations et d'agences spécialisées dans la GE ont également été consultés. Au final, toutes ces étapes de la revue littéraire ont permis de dresser le portrait énergétique de l'industrie forestière par l'entremise d'un cadre d'analyse rassemblant les concepts clés de la GE.

4 PORTRAIT ENERGETIQUE DE L'INDUSTRIE FORESTIERE

Dans cette section, l'analyse du profil énergétique de l'industrie forestière est présentée, basée sur cinq axes d'évaluation dégagés d'une revue de la littérature. La figure 1 dévoile les axes d'évaluation de cette étude sous la forme d'un cadre d'analyse dimensionnel qui montre la manière dont ceux-ci sont interreliés. Ces axes sont détaillés ci-après.

4.1 Situation énergétique de l'industrie forestière : Utilisation de l'énergie

Dans l'industrie des pâtes et papiers, l'électricité et le gaz naturel sont les sources d'énergie les plus utilisées, selon le type de procédé pour produire la pâte. Le procédé chimique nécessite une grande quantité d'énergie thermique sous la forme de vapeur, tandis que le procédé mécanique nécessite de grandes quantités d'électricité (Thollander *et al.*, 2020b). À son tour, l'électricité alimente tous les équipements des usines, tandis que le gaz naturel produit la vapeur et l'air chaud

nécessaire aux processus de production (Bajpai, 2016a). D'autre part, de l'électricité peut être produite à partir de la biomasse résiduelle générée dans les processus de production. Il s'agit principalement de boues, de liqueur noire, de déchets de bois et d'écorces. L'industrie des pâtes et papiers produit également de l'électricité et de la chaleur à partir de systèmes de cogénération et certaines usines possèdent leur propre centrale hydroélectrique (Adès, 2010). Étant donné que la biomasse est considérée comme neutre en carbone (CO₂), sa valorisation énergétique contribue à réduire les émissions de GES émises par l'industrie. En revanche, dans la fabrication du papier, le séchage est une activité très énergivore. D'énormes quantités de vapeur sont utilisées dans ce processus et le peu d'électricité utilisé sert à alimenter des pompes. Les processus les plus énergivores dans les usines de pâtes et papiers sont la production de pâte et le séchage du papier (Bajpai, 2016b). Il est important de mentionner que la consommation d'énergie dans les différentes industries de pâtes et papiers à travers le monde peut varier en fonction des procédés utilisés, de la taille de l'usine, de la technologie utilisée, des prix des combustibles, des pratiques de GE, etc. (voir Bajpai, 2016b). Au Canada, 62% de la consommation totale d'énergie de ce sous-secteur provient de la biomasse (Gouvernement du Canada, 2020). Ailleurs dans le monde, cette consommation issue de la biomasse représente plus de 65% aux États-Unis (EIA, 2011), 55,2% en Autriche (Posch *et al.*, 2015) et 96% en Suède (Swedish Forest Industries Federation [SFID], s.d.). D'autre part, dans le sous-secteur canadien de production de bois d'œuvre, la biomasse représente 54,4% de l'énergie totale utilisée, tandis que l'électricité représente 25,6% et le gaz naturel 14,2%. (Nyboer, J., et Bennett cité dans Quesada-Pineda *et al.*, 2015). En revanche, en 2016, des scieries suédoises étaient quasi exemptes d'hydrocarbures, ce qui montre que ce pays est en voie de devenir l'un des premiers à éliminer les combustibles fossiles de ses processus de production (SFID, s.d.).

4.2 Contrôle de la transition énergétique : Les moteurs de changement et les barrières

De manière générale, les moteurs de changement représentent les motivations des entreprises à vouloir améliorer leur efficacité énergétique, tandis que les barrières bloquent ces aspirations. Dans la littérature, une analyse réalisée en Suède dans l'industrie des pâtes et papiers a analysé en détail les principaux moteurs et barrières de ce sous-secteur à travers un questionnaire réalisé auprès de 29 entreprises du pays (Lawrence *et al.*, 2019). Dans leur étude, les auteurs ont déterminé que les facteurs économiques axés sur « la réduction des coûts énergétiques » sont les principaux moteurs de la GE dans ce sous-secteur, tandis que l'une des barrières les plus importantes était liée à la connaissance, dont la « grande complexité des processus de production ». D'autre part, dans une entreprise de pâtes et papiers de l'ouest du Canada, selon l'étude de cas menée par *Clean Energy Ministerial* [CEM] (2016a), « l'urgent besoin de l'usine de réduire les coûts a été un facteur de motivation essentiel » qui coïncide avec les résultats présentés par Lawrence *et al.* (2019). Aussi, la collaboration externe avec le gouvernement, le désir d'être plus compétitif, l'ambition du personnel et les audits externes ont été d'autres moteurs qui ont soutenu cette entreprise canadienne lors de la mise en place de son système de GE. Ses barrières étaient principalement liées à la faible disponibilité des capitaux et au manque d'intérêt pour l'efficacité énergétique étant donné que les cadres supérieurs ignoraient les bénéfices à tirer de l'implantation d'un système de GE. Par

ailleurs, dans une entreprise chilienne de pâtes et papiers évaluée par CEM (2018), les motivations de cette usine étaient plutôt de nature réglementaire, car l'entreprise cherchait à avoir une stratégie énergétique à long terme qui lui permettrait d'être plus compétitive et d'avoir une image plus écologique. L'une des barrières les plus difficiles pour cette entreprise sud-américaine fut la complexité et la diversité de ses processus. En Autriche, les usines de pâtes et papiers sont également confrontées à une barrière qui pourrait être qualifiée d'énergie renouvelable limitée, car la topographie du pays les empêche d'élargir leur production d'énergie au-delà de la biomasse (Posch *et al.*, 2015).

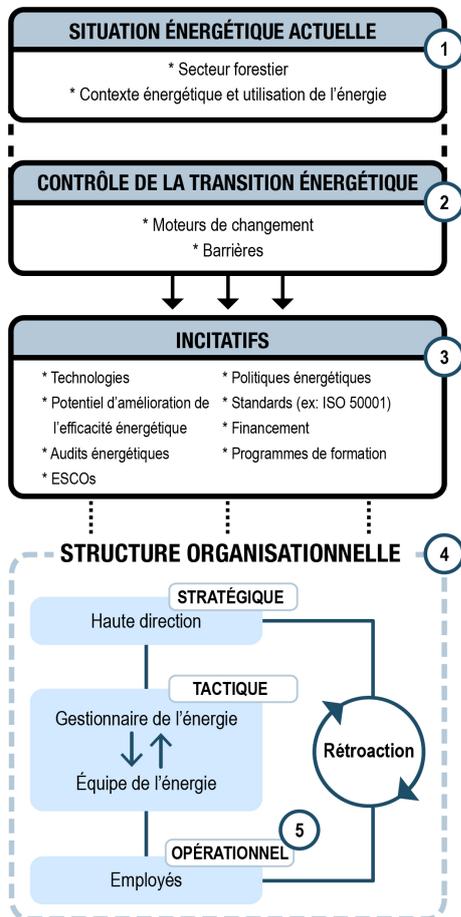


Fig 1. Cadre d'analyse dimensionnel des pratiques de GE dans l'industrie forestière

4.3 Incitatifs

Toute la documentation consultée autour de la GE semble par ailleurs mettre en relief la nécessité de prendre en compte la présence d'incitatifs pour favoriser un virage durable et de tels leviers peuvent prendre plusieurs formes.

4.3.1 Technologie

L'utilisation de la biomasse est devenue une pierre angulaire de l'industrie forestière, car elle offre des avantages énergétiques que les autres secteurs manufacturiers ne possèdent pas. Les industries de ce secteur qui disposent d'une infrastructure moderne leur permettant d'être plus autonomes énergétiquement se concentrent actuellement sur la réduction de leur consommation d'énergie et l'augmentation de la production d'énergie (Bajpai, 2016b). Dans les usines des pâtes et papiers, des améliorations technologiques permettent d'atteindre ces deux objectifs. Dans une analyse composée de divers scénarios, Fleiter *et al.* (2012) ont évalué 17

technologies différentes pour les pâtes et papiers et concluent que les technologies présentant le plus grand potentiel d'économie d'énergie s'avèrent la récupération de chaleur dans les usines de papier et l'utilisation de technologies innovantes de séchage du papier. Cependant, Bajpai (2016a) mentionne qu'après quelques tentatives initiales pour introduire des nouvelles technologies de séchage, la R&D dans ce domaine n'a plus progressé. Dans l'analyse de Fleiter *et al.* (2012), après un sondage auprès de certains représentants allemands de ce sous-secteur, les coûts élevés liés à l'acquisition de nouvelles technologies se sont révélés être une barrière économique pour l'industrie. Les industriels ont établi à deux ans le délai maximal de récupération acceptable pour les investissements liés à l'amélioration de l'efficacité énergétique. D'autre part, en Suède il y a des initiatives en cours de développement visant à éliminer le peu d'énergie fossile utilisé par l'industrie forestière, qui sert dans les fours à chaux des usines de pâtes et papiers et en moindre quantité dans la production de papier. Plusieurs usines ont déjà investi dans des alternatives leur permettant d'utiliser des granules de bois et d'autres biocarburants au lieu de l'huile fossile dans les fours à chaux (SFID, s.d.). Dans l'industrie forestière canadienne, il existe actuellement de nombreuses technologies nouvelles qui permettent la création d'une grande variété de biocarburants à partir de la biomasse forestière, tels que les procédés de gazéification, de carbonisation et de torréfaction, ainsi que la pyrolyse et les liquides de transport. De même, les systèmes de cogénération permettent de produire de l'électricité et de la chaleur simultanément, ce qui rend les installations moins énergivores. Toutefois, certaines de ces technologies n'ont pas encore atteint un stade mature de commercialisation (MRNF, 2012). D'autre part, l'avenue de la bioénergie vise à développer des bioraffineries intégrées et diversifiées produisant de la pâte, du papier, des produits chimiques, de la bioénergie, des biocarburants, etc. (Bajpai, 2016a). À cet égard, pour l'industrie du sciage et des pâtes et papiers, le développement de ces technologies émergentes est l'occasion de transformer les résidus de bois pour les utiliser à des fins énergétiques, afin d'améliorer l'efficacité énergétique de plusieurs entreprises du secteur (MRNF, 2012).

4.3.2 Potentiel d'amélioration de l'efficacité énergétique

La première étape du déploiement de toute initiative d'amélioration énergétique commence par la création d'un programme de GE stratégique et bien ciblé. Comme mentionné précédemment, un système de GE permet aux organisations d'augmenter leur efficacité énergétique et leur compétitivité grâce à une surveillance continue de la consommation d'énergie (Bajpai, 2016c). Également, en plus de produire des résultats rapidement, les coûts de mise en œuvre ont régulièrement des retours sur investissement de moins de deux ans tout en générant des économies continues (RNCAN, 2019a). À titre d'exemple, dans une usine de pâtes et papiers située dans l'ouest du Canada, le coût de mise en œuvre seul du système de GE basé sur le cadre ISO 50001 a eu un retour sur investissement de moins d'un an, et ce, sans tenir compte des coûts des projets d'économie d'énergie (CEM, 2016a). De même, les systèmes de GE correctement implantés peuvent générer des économies d'énergie considérables. À titre d'exemple, dans une usine indonésienne de pâtes et papiers CEM (2016b), suite à l'implantation d'un système de GE, l'usine a obtenu des économies d'énergie totales d'environ 15,2%, ce qui équivaut à une économie d'énergie de 428,834 GJ. Grâce à son système de GE, l'entreprise a réalisé des économies d'énergie annuelles de 1,6 millions USD. Par

ailleurs, selon l'étude de cas CEM (2016a), l'usine canadienne de pâtes et papiers sous étude a déclaré avoir eu une réduction de 5,6% de sa consommation d'électricité après avoir mis en place son système de GE basé sur la norme ISO 50001, ce qui s'est traduit par une économie d'énergie cumulée de 181 GWh en trois ans. Les systèmes de GE semblent donc faire leur preuve. Néanmoins, sans un investissement important dans des technologies écoénergétiques, sans l'existence de bonnes pratiques de GE (Johansson & Thollander, 2018) et sans l'entretien adéquat des équipements et des installations, seules des économies d'énergie limitées peuvent être réalisées. Dans ce contexte, les investissements dans des systèmes de cogénération ont permis à des entreprises européennes de pâtes et papiers de produire près de la moitié de l'énergie électrique qu'elles consomment (Bajpai, 2016c). En Suède, Thollander *et al.* (2020b) illustrent quelques exemples de mesures d'efficacité énergétique mises en œuvre dans l'industrie des pâtes et papiers en rapport avec les systèmes de pompage. D'autres moyens d'économiser de l'énergie dans cet industrie sont le recyclage du papier et l'utilisation de papier recyclé, l'utilisation de moteurs et de systèmes d'éclairage plus efficaces, l'utilisation de technologies innovantes de séchage du papier, l'installation de variateurs de fréquence, la mise à jour des compteurs d'acquisition de données énergétiques, ainsi que les mesures de conservation de l'énergie axées sur la modernisation des anciennes usines et sur l'utilisation de processus et machines moins énergivores (voir Bajpai, 2016c ; Fleiter *et al.*, 2012 ; CEM, 2018 ; Bajpai, 2016a). Du côté de l'industrie de la fabrication des produits du bois, les opérations des scieries nécessitent l'utilisation de moteurs électriques lourds et d'équipements au gaz naturel tels que des chaudières. Les moteurs électriques sont utilisés pour alimenter la plupart des équipements des installations et sont donc les plus grands utilisateurs finaux d'énergie dans une scierie, ce qui en fait des candidats parfaits pour évaluer leur potentiel d'économie d'énergie (Gopalakrishnan *et al.*, 2012). Pour leur part, Wengert et Meyer (1992) ont révélé dans leur étude plus de 75 moyens économiques de réduire la consommation d'énergie dans les scieries sans affecter la production. Certaines de ces recommandations se concentrent sur les moteurs électriques, les systèmes d'éclairage, les compresseurs, le sciage, les chaudières, les fours, les systèmes CVC et l'utilisation de la cogénération pour produire de l'électricité interne. De leur côté, Gopalakrishnan *et al.* (2012) ont développé un outil d'analyse permettant aux scieries de comparer leur consommation d'énergie par rapport à une référence énergétique indiquant la consommation d'énergie optimale requise dans les opérations. La différence entre la consommation d'énergie actuelle et la valeur de référence représente les économies d'énergie potentielles. De ce fait, l'identification des possibles opportunités d'amélioration est la première étape pour évaluer la maturité énergétique d'une organisation.

4.3.3 Audits énergétiques

En tant que supports fondamentaux de tout système de GE, les audits énergétiques périodiques assurent l'amélioration continue de l'efficacité énergétique d'un procédé, d'une organisation ou d'un groupe industriel (Thollander *et al.*, 2020a). Ces audits se concentrent sur l'analyse de la consommation d'énergie et identifient les domaines où des améliorations potentielles sont nécessaires et où les objectifs n'ont pas été atteints. L'évaluation de la performance énergétique et le suivi des activités d'efficacité énergétique nécessitent que l'utilisation finale de l'énergie soit mesurée. Le

moyen le plus efficace pour y parvenir consiste à utiliser des systèmes de mesure et de surveillance de l'énergie (Johansson & Thollander, 2018). Les grandes entreprises ont parfois besoin que des audits soient réalisés par un auditeur énergétique certifié (Thollander *et al.*, 2020a). Dans le secteur forestier, des audits internes et externes ont permis à plusieurs entreprises du sous-secteur des pâtes et papiers de différentes régions du monde d'identifier des opportunités d'amélioration, de mettre en place des contrôles opérationnels, de mieux utiliser les actifs existants, d'analyser leur consommation d'énergie, d'établir des projets d'investissement et de vérifier si leurs activités étaient toujours consistantes avec les exigences de leurs systèmes de GE (voir CEM, 2016a ; CEM, 2016B ; CEM, 2018). D'un autre côté, dans le sous-secteur de la fabrication de produits du bois, Wengert et Meyer (1992) soulignent que l'audit énergétique est aussi la première étape pour générer des économies d'énergie dans les scieries.

4.3.4 Entreprises de services énergétiques

Les entreprises qui ont besoin d'avoir une vue d'ensemble plus objective et concrète de leur situation énergétique actuelle, ainsi que celles qui n'ont pas les compétences et le personnel nécessaires pour réaliser des audits énergétiques, peuvent recourir à la sous-traitance de services externes proposés par les entreprises de services énergétiques appelées ESCOs (Johansson & Thollander, 2018). Ces entreprises peuvent offrir divers services-conseils en énergie, tels que des audits énergétiques et l'identification de mesures d'efficacité énergétique (Thollander *et al.*, 2020a). Grâce à ces entreprises de services énergétiques, différentes industries du secteur forestier ont pu avoir une vue d'ensemble réaliste et complète de leur profil énergétique, former leur personnel interne et identifier les potentiels d'amélioration énergétique de leurs installations (voir CEM, 2016a ; CEM, 2016B ; CEM, 2018).

4.3.5 Politiques énergétiques, standards et financement

Les politiques énergétiques déterminées par une entité gouvernementale sont souvent axées sur la résolution des problèmes de changement climatique et sur le développement énergétique. Les politiques énergétiques mises sur pied à l'interne sont quant à elles plus axées sur le déploiement de plans stratégiques permettant aux organisations d'améliorer leur efficacité énergétique et de réduire leurs émissions de GES à long terme (Abdelaziz *et al.*, 2011). Au Canada, il existe un programme de GE industrielle promu par RNCAN qui vise à accélérer la transition énergétique industrielle du pays. Ce programme gouvernemental offre des systèmes d'information de GE, fait la promotion de la GE industrielle par la mise en œuvre de la norme ISO 50001 et fournit un financement aux entreprises souhaitant mettre en œuvre cette norme et réaliser des évaluations énergétiques ou d'autres projets de GE (RNCAN, 2019a). De plus, le gouvernement supporte jusqu'à 50 % du coût, jusqu'à concurrence de 40 000\$ CA. À titre d'exemple, l'entreprise canadienne de pâtes et papiers évaluée dans le cas d'études (CEM, 2016a) a réussi à mettre en œuvre son système de GE basé sur la norme ISO 50001 avec le soutien et le financement de ce programme gouvernemental. En Suède, toutes les usines de pâtes et papiers ont été certifiées avec un système de GE normalisé en 2005, appelé le programme pour améliorer l'efficacité énergétique dans les industries à forte intensité énergétique (PFE). Parmi les usines suédoises de pâtes et papiers, la plupart ont par la suite mis en œuvre un système de GE basé sur la norme ISO 50001 (Andersson & Thollander, 2019). Les objectifs énergétiques actuels du secteur forestier suédois résident dans l'élimination progressive des sources d'énergie fossiles et dans

la croissance de sa bioéconomie (Swedish Forest Industries Federation, s.d.), tandis que le secteur forestier canadien, en divisant ses objectifs, se concentre sur la valorisation de la biomasse forestière, cherche à stimuler l'innovation de nouveaux produits et procédés, ainsi qu'à exploiter des nouvelles technologies numériques (MRNF, 2012). D'un autre côté, les politiques internes doivent avoir des objectifs clairs et transparents puisqu'elles expriment l'engagement d'une entreprise à accroître son efficacité énergétique et à réduire son impact environnemental (Gordic *et al.*, 2014). Dans ce contexte, il semble que la réussite de ces politiques dans l'industrie forestière dépendra des cadres supérieurs, puisque ceux-ci seront chargés d'assurer la disponibilité des ressources et d'aligner l'ensemble de l'organisation avec les objectifs énergétiques établis (voir par exemple Gordic *et al.*, 2014).

4.4 Structure organisationnelle

La coordination des activités d'un programme de GE au sein d'une organisation nécessite un leader qui identifie les besoins énergétiques actuels et qui poursuit sans cesse les objectifs énergétiques établis dans la politique interne de l'entreprise. Le gestionnaire d'énergie est donc appelé à développer les stratégies, à gérer les projets énergétiques et à garder motivées les troupes. Gordic *et al.* (2014) décrivent dans leur étude une liste détaillée des rôles que joue régulièrement un gestionnaire d'énergie lors de la mise en œuvre d'un système de GE. Selon les auteurs, dans l'industrie du meuble, étant donné qu'il s'agit d'une industrie non énérgivore composée régulièrement de petites et moyennes entreprises, un seul coordinateur énergétique suffit. À l'inverse, selon un entretien réalisé par Stenqvist *et al.* (2011) avec 8 entreprises suédoises appartenant au sous-secteur des pâtes et papiers, la complexité des procédés des usines, la technologie utilisée pour convertir et distribuer l'énergie, les différentes sources d'énergie utilisées dans les différentes opérations et le grand nombre d'employés qui peut être présent (généralement de 500 à 1000 employés), sont des facteurs qui mènent plutôt à un comité ou à une équipe d'énergie multifonctionnelle, composée de plusieurs représentants de différents départements de l'entreprise qui soutiennent activement le gestionnaire d'énergie (Stenqvist *et al.* (2011) ; Gordic *et al.* (2014) ; CEM (2016B) ; CEM (2018)). La dernière position dans la structure organisationnelle est occupée par les employés qui connaissent mieux que quiconque le fonctionnement des équipements, leurs problèmes et la manière dont ceux-ci peuvent économiser de l'énergie. Leurs idées doivent toujours être soutenues et mises en valeur par le gestionnaire de l'énergie. Leur motivation et leur engagement envers les objectifs de l'entreprise en dépendront (Gordic *et al.*, 2014). De plus, un bon moyen de stimuler le changement dans toute la structure organisationnelle est par l'entremise de formations et de cours de sensibilisation axés sur la GE. Selon Abdelaziz *et al.* (2011), il existe généralement deux cours de GE industrielle. Le premier est axé sur les ingénieurs de l'entreprise et les personnes directement impliquées dans le programme énergétique et le second est donné de manière générale dans toute l'organisation à travers des programmes universitaires internes. Selon CEM (2016a), pour l'entreprise canadienne de pâtes et papiers à l'étude, « Energy University est un élément important de la formation des employés » et « comporte un volet obligatoire de sensibilisation à l'énergie pour tous les employés et entrepreneurs ». D'autres entreprises de pâtes et papiers ont également adopté des programmes de formation et de communication dans leurs installations, comme l'illustrent Stenqvist *et al.* (2011) et CEM (2016b).

4.5 GE au niveau opérationnel

Après la mise en œuvre des mesures énergétiques définies dans le plan d'actions, l'étape suivante consiste à évaluer la performance énergétique des améliorations à travers un système de mesure, de contrôle et d'analyse des données qui permet de traiter les informations afin de définir des actions correctives si nécessaire (Gordic *et al.*, 2014). Selon l'étude de Stenqvist *et al.* (2011), 5 des 8 usines de pâtes et papiers interrogées disposaient de compteurs à lecture automatique pour traiter les données en temps réel, à partir d'appareils de mesure chargés d'évaluer constamment la consommation d'électricité et de vapeur des équipements. Les données étaient transférées sur l'intranet des entreprises. Selon la taille de chaque usine, entre 20 et 80 appareils de mesure étaient nécessaires pour mesurer la consommation d'électricité et de vapeur. Une fois les données énergétiques collectées, celles-ci étaient combinées avec les données de production pour établir la consommation spécifique d'énergie (qui est un indicateur de performance permettant d'évaluer l'efficacité énergétique des systèmes). Selon Thollander *et al.* (2020a), la SEC est définie comme l'utilisation d'énergie primaire divisée par le nombre de produits pendant une période donnée. Dans l'étude de cas CEM (2016a), pour surveiller et suivre la performance énergétique des systèmes, l'équipe énergie a utilisé un tableau de bord qui leur a permis, avec l'utilisation de différents indicateurs de performance, de comparer la performance énergétique de l'usine de pâtes et papiers mois par mois, d'années en années. Par contre, ils ont surveillé la SEC uniquement au niveau de l'usine. À titre d'exemple, ils ont mesuré la consommation spécifique d'énergie de l'usine en MWh / tonne et celle des procédés de pâte et papier en kWh / tonne. Ils ont également évalué l'énergie nette consommée dans toute l'usine (MWh). Les indicateurs économiques qu'ils ont utilisés évaluaient les coûts de l'électricité et des combustibles fossiles utilisés. Dans le cas de l'usine de pâtes et papiers indonésienne (CEM, 2016b), à travers un diagramme de Pareto, l'équipe énergie a d'abord identifié les principaux consommateurs d'énergie de l'usine. Après les avoir identifiés, ils ont analysé les données historiques et ont déterminé que la corrélation entre le niveau de production et l'énergie consommée dans les processus était suffisamment élevée pour prévoir la consommation d'énergie future. Par conséquent, pour analyser la performance énergétique de leurs opérations, ils ont comparé la consommation d'énergie prévue avec la consommation d'énergie réelle. L'écart entre le réel et l'estimation correspondait aux économies d'énergie. Ainsi, au total, ils ont utilisé deux indicateurs de performance pour évaluer la performance énergétique de leurs opérations, le premier étant l'indice d'intensité énergétique (EII) pour analyser la différence entre la consommation d'énergie réelle et la consommation prévue, et le second étant la SEC, c'est-à-dire la consommation d'énergie divisée par le niveau de production (GJ / Tonne). Par ailleurs, dans une étude réalisée en Suède par Andersson & Thollander (2019), après avoir mené une enquête auprès de 28 usines de pâtes et papiers, les auteurs ont constaté que 93% disposaient d'indicateurs de performance d'énergie pour mesurer l'efficacité énergétique totale de leurs sites, et 93% avaient également des indicateurs de performance pour surveiller les processus de production les plus énérgivores. Cependant, la plupart des usines ne semblaient pas surveiller les indicateurs de performance au niveau des processus. Alors que des études comme celle d'Ammara *et al.* (2016) apportent une contribution importante au niveau du développement des indicateurs de performance dans les usines de pâtes et papiers, il semble qu'il y ait tout de même encore du chemin à faire

puisque le développement de ceux-ci est régulièrement entravé par le manque de ressources, qui dans l'étude d'Andersson & Thollander (2019) est considéré comme la plus grande barrière énergétique. Finalement, du côté du sous-secteur de la fabrication de produits du bois, Johnsson *et al.* (2019) fournissent une liste d'indicateurs de performance pour surveiller l'efficacité énergétique d'une scierie entière et de ses deux processus les plus énergivores (le sciage et le séchage du bois). Les auteurs soulignent que l'indicateur de performance le plus utilisé dans une scierie est la quantité d'énergie utilisée par m³ de produits sciés (kWh / m³).

4.6 Discussion des résultats

Suite à l'étude de la situation énergétique de l'industrie forestière, il a été possible de constater que le degré de maturité énergétique de cette industrie variait considérablement d'un pays à l'autre. En Europe, la Suède dépend de moins en moins des sources d'énergie fossiles pour alimenter ses opérations, étant un pays quasiment exempt d'hydrocarbures (SFID, s.d.), alors que l'industrie forestière canadienne, malgré ses politiques énergétiques qui cherchent à tirer de plus en plus profit de la biomasse forestière, continue à recourir aux combustibles fossiles (MRNF, 2012). À cet égard, selon Bajpai (2016c), le Japon, la Corée, la France et la Suède semblent être les pays les plus avancés en termes d'économie d'énergie dans ce secteur industriel. Concernant le contrôle vers la transition énergétique du secteur forestier, dans les exemples analysés, il semble que la réduction des coûts énergétiques soit l'une des principales motivations pour gérer l'énergie des opérations (Lawrence *et al.*, 2019), tandis que le manque de capital et la complexité des processus sont les principales barrières dans le sous-secteur des pâtes et papiers (voir CEM, 2016a ; CEM, 2018). En ce qui concerne les incitatifs, l'utilisation de la biomasse forestière résiduelle et les systèmes de cogénération sont des vecteurs stratégiques qui mobilisent la transition énergétique de cette industrie dans le monde (Fleiter *et al.*, 2012). Ainsi, ces derniers sont des technologies présentant un grand potentiel énergétique pour l'industrie forestière. Néanmoins, l'acquisition de nouvelles technologies nécessite des investissements, ce qui représente une barrière économique pour cette industrie. D'un autre côté, parmi les différentes manières de générer des économies d'énergie, les systèmes de GE se sont distingués pour avoir été des instruments offrant des avantages environnementaux, économiques et sociaux aux entreprises du secteur forestier qui les ont mis en œuvre (voir CEM, 2016a ; CEM, 2016B ; CEM, 2018). De même, les technologies écoénergétiques émergentes seront des éléments clés à privilégier afin d'avoir des systèmes énergétiques durables au sein de cette industrie (MRNF, 2012). Le recours à des audits dans l'industrie forestière s'est par ailleurs avéré une approche porteuse, car ceux-ci ont permis à plusieurs entreprises du secteur d'identifier leurs opportunités d'amélioration et de s'assurer que leurs systèmes de GE restaient conformes à leurs politiques et objectifs énergétiques établis. Aussi, des politiques énergétiques consistantes et bien définies ont permis à l'industrie forestière d'augmenter son efficacité énergétique et de réduire son empreinte écologique. Dans le cas de l'Amérique du Nord et de l'Asie, il semble nécessaire de mettre en œuvre des politiques énergétiques plus strictes et consistantes leur permettant d'atteindre leurs objectifs énergétiques à long terme (Bajpai, 2016c). Dans le cas du Canada, les programmes offerts par son gouvernement symbolisent déjà des soutiens importants qui encouragent et financent les initiatives énergétiques de l'industrie forestière de ce pays (RNCAN, 2019a). Lorsqu'on jette un œil à la dimension

de la structure organisationnelle, le rôle joué par le gestionnaire d'énergie a été déterminant pour toutes les entreprises forestières ayant mis en place un système de GE, car ce joueur important a favorisé positivement l'atteinte des objectifs énergétiques et le succès des politiques énergétiques internes. Les programmes de formation se sont également révélés essentiels pour sensibiliser le personnel des organisations. Au niveau opérationnel, le manque d'indicateurs de performance pour les processus, particulièrement dans les usines de pâtes et papiers, semble avoir freiné un traitement efficace des données énergétiques. De même, bien que la consommation de certains produits dans ce secteur ait considérablement diminué ces dernières années et du fait que ces circonstances l'ont rendu un secteur moins productif, moins énergivore et moins polluant, son rôle comme émetteur de GES reste encore majeur. Pour cette raison, le moyen le plus tangible d'améliorer la compétitivité, la rentabilité et l'empreinte écologique de cette industrie, sans affecter davantage sa productivité et sa croissance économique, repose sur le recours à des approches de GE offrant des avantages au niveau de la dimension sociale, économique et environnementale de chaque organisation. Finalement, bien que le cadre d'évaluation développé dans cette recherche ait servi à analyser le profil énergétique de l'industrie forestière, il apparaît suffisamment générique pour être exploité comme outil d'analyse diagnostic dans d'autres secteurs industriels.

5 CONCLUSION

L'industrie forestière, en tant que grande consommatrice et productrice d'énergie, doit poursuivre ses efforts pour améliorer son efficacité énergétique et réduire encore plus ses émissions de GES. Bien que ses progrès en matière de conservation de l'énergie soient dignes de mention, la GE dans les systèmes de production est un élément clé qu'il faut que l'industrie sache mettre en valeur afin d'améliorer sa rentabilité et son efficacité énergétique sur le long terme. Grâce à l'utilisation d'un cadre d'analyse dimensionnel de la GE développé dans cette recherche, il a été possible de mener à bien l'étude du profil énergétique de l'industrie forestière dans le monde. Les résultats de cette analyse ont permis de constater que cette industrie est en pleine transition énergétique, que sa bioéconomie se développe à grande vitesse, que sa transition vers des sources d'énergie plus neutres en carbone sera possible grâce aux technologies émergentes visant à exploiter tout le potentiel énergétique de la biomasse forestière. Toutefois, sans l'existence de bonnes pratiques de GE et d'un programme de GE bien défini, son potentiel d'efficacité énergétique risque de demeurer limité. Enfin, le fait qu'il y ait des pays dans cette industrie plus développés d'un point de vue économie d'énergie que d'autres, montre que de nouvelles mesures, politiques et stratégies énergétiques seront encore nécessaires pour garantir un avenir plus durable.

Dans une recherche future, il serait pertinent de réaliser des études statistiques plus approfondies qui suivent les mêmes bases d'analyse que ce travail, afin de créer une image plus précise du degré de maturité énergétique des différentes industries forestières dans le monde. De cette façon, il deviendrait possible de mieux préciser les raisons pour lesquelles les pratiques énergétiques de cette industrie sont hétérogènes et mettre en lumière la direction vers laquelle orienter la transition énergétique de ce secteur. Une telle étude permettrait également d'analyser de quelle façon l'industrie forestière évolue en termes d'exploitation numérique de

l'information, puisqu'il semble y avoir un besoin de collecter, traiter et gérer plus efficacement les données énergétiques.

6 REFERENCES

- Abdelaziz, E. A., Saidur, R., & Mekhilef, S. (2011). A review on energy saving strategies in industrial sector. *Renewable and Sustainable Energy Reviews*, 15(1), 150-168.
- Adès, J. (2010). L'impact du cadre réglementaire fédéral de réduction d'émissions de gaz à effet de serre de 2007 sur les choix des sources d'énergie de l'industrie québécoise des pâtes et papiers, Université Laval, Québec.
- Ammara, R., Fradette, L., & Paris, J. (2016). Equipment performance analysis of a Canadian Kraft mill. Part I: Development of new key performance indicators (KPI). *Chemical Engineering Research and Design*, 115, 160-172.
- Andersson, E., & Thollander, P. (2019). Key performance indicators for energy management in the Swedish pulp and paper industry. *Energy Strategy Reviews*, 24, 229-235.
- Bajpai, P. (2016a). General Background. In *Pulp and Paper Industry* (pp. 1-8).
- Bajpai, P. (2016b). Pulp and Paper Production Processes and Energy Overview. In *Pulp and Paper Industry* (pp. 15-49).
- Bajpai, P. (2016c). Future Perspectives. In *Pulp and Paper Industry* (pp. 253-256).
- Bunse, K., Vodicka, M., Schönsleben, P., Brühlhart, M., & Ernst, F. O. (2011). Integrating energy efficiency performance in production management – gap analysis between industrial needs and scientific literature. *Journal of Cleaner Production*, 19(6-7), 667-679.
- Clean Energy Ministerial (2016a). Catalyst paper case study. Accès le 16/09/2020, https://www.cleanenergyministerial.org/sites/default/files/2018-12/Catalyst_Canada.pdf
- Clean Energy Ministerial (2016b). PT Indah Kiat Pulp & Paper case study. Accès le 16/09/2020, https://www.cleanenergyministerial.org/sites/default/files/2018-12/PT%20IKPP-%20Tangerang_Indonesia.pdf
- Clean Energy Ministerial (2018). CMPC case study. Accès le 16/09/2020, <http://www.cleanenergyministerial.org/initiative-clean-energy-ministerial/2018-energy-management-leadership-awards>
- EIA, 2011. International Energy Outlook 2011. U.S. Energy Information Administration (EIA). American Forest & Paper Association (AF&PA). 2012.
- Fleiter, T., Fehrenbach, D., Worrell, E., & Eichhammer, W. (2012). Energy efficiency in the German pulp and paper industry – A model-based assessment of saving potentials. *Energy*, 40(1), 84-99.
- Gopalakrishnan, B., Mardikar, Y., Gupta, D., Jalali, S. M., & Chaudhari, S. (2012). Establishing Baseline Electrical Energy Consumption in Wood Processing Sawmills for Lean Energy Initiatives: A Model Based on Energy Analysis and Diagnostics. *Energy Engineering*, 109(5), 40-80.
- Gordic, D., Babic, M., Jelic, D., Koncalovic, D., & Vukasinovic, V. (2014). Integrating energy and environmental management in wood furniture industry. *ScientificWorldJournal*, 2014, 596958.
- Gouvernement du Canada (2020). "Forest bioeconomy, bioenergy and bioproducts." Accès le 16/09/2020, <https://www.nrcan.gc.ca/our-natural-resources/forests-forestry/forest-industry-trade/forest-bioeconomy-bioenergy-bioproducts/13315>
- Hasan, A. S. M. M., Hoq, M. T., & Thollander, P. (2018). Energy management practices in Bangladesh's iron and steel industries. *Energy Strategy Reviews*, 22, 230-236.
- Hossain, S. R., Ahmed, I., Azad, F. S., & Monjurul Hasan, A. S. M. (2020). Empirical investigation of energy management practices in cement industries of Bangladesh. *Energy*, 212.
- ISO. (2018). ISO 50001: 2018, Energy management systems.
- Johansson, M. T., & Thollander, P. (2018). A review of barriers to and driving forces for improved energy efficiency in Swedish industry– Recommendations for successful in-house energy management. *Renewable and Sustainable Energy Reviews*, 82, 618-628.
- Johnsson, S., Andersson, E., Thollander, P., & Karlsson, M. (2019). Energy savings and greenhouse gas mitigation potential in the Swedish wood industry. *Energy*, 187.
- Lawrence, A., Nehler, T., Andersson, E., Karlsson, M., & Thollander, P. (2019). Drivers, barriers and success factors for energy management in the Swedish pulp and paper industry. *Journal of Cleaner Production*, 223, 67-82.
- Ministère des Ressources naturelles et de la Faune (2012). Stratégie 2012-2017 pour transformer l'industrie québécoise des produits forestiers. Accès le 16/09/2020, <https://mffp.gouv.qc.ca/documents/forets/entreprise/strategie-developpement-2012-2017.pdf>
- National Energy Board, 2019. Canada's Energy Transition. Récupéré le 16/09/2020 sur <https://www.cer-rec.gc.ca/en/data-analysis/canada-energy-future/canada-energy-transition/cndsnrgrtrnst-eng.pdf>
- Posch, A., Bruderer, T., Braschel, N., & Gabriel, M. (2015). Strategic energy management in energy-intensive enterprises: a quantitative analysis of relevant factors in the Austrian paper and pulp industry. *Journal of Cleaner Production*, 90, 291-299.
- Quesada-Pineda, H., Wiedenbeck, J., & Bond, B. (2015). Analysis of electricity consumption: a study in the wood products industry. *Energy Efficiency*, 9(5), 1193-1206.
- Ressources naturelles Canada (2019a). Energy Efficiency in Canada: Report to parliament under the energy efficiency act. Accès le 16/09/2020, <https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/www/pdf/publications/emmc/parliament17-18.pdf>
- Ressources naturelles Canada (2019b). Cahier d'information sur la forêt 2018-2019. Accès le 16/09/2020, <https://scf.nrcan.gc.ca/publications?id=39506>
- Ressources naturelles Canada. (2020). L'état des forêts au Canada : Rapport annuel 2019. Accès le 16/09/2020, <https://scf.nrcan.gc.ca/publications?id=40085>
- Sandberg, D., Vasiri, M., Trischler, J., & Öhman, M. (2014). The role of the wood mechanical industry in the Swedish forest industry cluster. *Scandinavian Journal of Forest Research*, 29(4), 352-359.
- Stenqvist, C., Nilsson, L. J., Ericsson, K., & Modig, G. (2011). Energy management in Swedish pulp and paper industry - the daily grind that matters. Paper presented at the 10th eceee summer study - energy efficiency first: the foundation of a low-carbon society, France.
- Swedish Forest Industries Federation (s.d.). "Roadmap for fossil free competitiveness Forest Sector." Accès le 01/10/2020, <https://www.forestindustries.se/news/publications-and-reports/>
- Thollander, P., Karlsson, M., Rohdin, P., Wollin, J., & Rosenqvist, J. (2020a). Energy management. In *Introduction to Industrial Energy Efficiency* (pp. 239-257).
- Thollander, P., Karlsson, M., Rohdin, P., Wollin, J., & Rosenqvist, J. (2020b). Industrial energy systems. In *Introduction to Industrial Energy Efficiency* (pp. 33-47).
- Wengert, G. & Meyer D. (1992). Energy at the sawmill: Conservation and cost reduction. *Forestry Facts*, 61.

Prédiction du degré d'obsolescence en utilisant des techniques d'apprentissage automatique

IMEN TRABELSI^{1,2}, BESMA ZEDDINI³, MARC ZOLGHADRI^{1,4}, MAHER BARKALLAH², MOHAMED HADDAR²

¹ Laboratoire Quartz, Supmeca
3 Rue Fernand Hainaut, 93407 Saint-Ouen, France
Imen.Trabelsi@Supmeca.Fr

² Laboratoire LA2MP
ENIS, ROUTE SOUKRA KM 3.5, 3038 SFAX, TUNISIE

³ LABORATOIRE SATIE CNRS - UMR 8029
CYTEH ENS PARIS-SACLAY, CERGY FRANCE

⁴ LAAS - CNRS
7 AVENUE DU COLONEL ROCHE, 31400 TOULOUSE, FRANCE

Résumé – La prédiction précise du risque d'obsolescence des composants critiques est un défi important dans la gestion du cycle de vie des produits afin d'améliorer la durabilité du système. Traditionnellement, de nombreuses techniques mathématiques ont été proposées pour traiter ce problème. Des études récentes ont également montré que les méthodes d'apprentissage peuvent être un moyen efficace d'améliorer la qualité des prédictions. Cependant, ces techniques évaluent l'obsolescence de manière statique sans tenir compte de son évolution dans le temps. Ce document cherche à répondre à cette problématique et propose une méthode basée sur des techniques d'apprentissage automatique pour prédire le degré d'obsolescence dans le temps. Cette technique est testée dans une étude de cas réel qui concerne l'obsolescence des smartphones.

Abstract – Accurate prediction of the risk of obsolescence of critical components is an important challenge in product life cycle management in order to improve the sustainability of the system. Traditionally, many mathematical techniques have been proposed to address this problem. Recent studies have also shown that learning methods can be an effective way to improve the quality of predictions. However, these techniques assess obsolescence in a static way without taking into account its evolution over time. This paper seeks to address this issue and proposes a method based on automatic learning techniques to predict the degree of obsolescence over time. This technique is tested in a real-life case study that concerns the obsolescence of smartphones.

Mots clés - Obsolescence, Prédiction, Apprentissage automatique.

Keywords – Obsolescence, Prediction, Machine Learning.

1 INTRODUCTION

L'obsolescence est un concept désormais très courant dans notre société où tout change constamment [Trabelsi, I. et al. (2021)]. Elle se définit comme la transition de l'état de disponibilité à l'état d'indisponibilité d'une entité auprès de son fabricant conformément à la spécification d'origine, [IEC 62402, 2019]. Certaines des raisons qui rendent les produits obsolètes comprennent les progrès technologiques, l'épuisement des ressources naturelles, la perte de fournisseurs, les conflits géopolitiques et les réglementations, etc. L'obsolescence peut avoir des conséquences sur l'environnement, la société, les exigences du produit (maintenabilité, testabilité, supportabilité...) et même le fonctionnement des entreprises.

Afin de réduire ces impacts, l'obsolescence doit être gérée d'une façon efficace. Afin de réduire les implications négatives de l'obsolescence, plusieurs solutions existent qui peuvent être classées [Sandborn, P. (2013)] selon 3 manières de gestion à savoir:

- La gestion réactive qui consiste à prendre des mesures lorsque l'obsolescence est déjà survenue.
- La gestion proactive qui est mise en place pour les composants critiques qui risquent de devenir obsolètes.

- Et la gestion stratégique qui complète la gestion proactive et réactive et comprend la détermination de la combinaison optimale des approches d'atténuation et des rafraîchissements de la conception.

Le type de gestion le plus utilisé est la gestion réactive, car elle est plus facile à mettre en œuvre mais il est recommandé de ne l'utiliser que si le coût associé à l'obsolescence d'un composant est faible [Pingle, P. (2015)]. Toutefois, si la probabilité d'obsolescence et les coûts associés sont élevés, il est recommandé d'appliquer des stratégies de gestion proactive pour minimiser le risque d'obsolescence et les coûts associés [Rojo et al., (2010)].

L'obsolescence peut alors être étudiée selon deux points de vue complémentaires à savoir la gestion de l'obsolescence et la conception de produits résilients à l'obsolescence. Ce travail se focalise sur le premier aspect et plus précisément sur la prédiction d'obsolescence.

Le reste de ce document est structuré comme suit : la section 2 présente une analyse documentaire relative aux différents concepts d'obsolescence et d'apprentissage automatique. Les différentes étapes de la méthode proposée sont présentées dans la section 3. La section 4 décrit les résultats de simulation et d'expérimentation d'un exemple illustratif sur les smartphones pour évaluer l'efficacité de la méthode proposée. Enfin,

quelques conclusions et travaux futurs sont présentés dans la section 5.

2 CONCEPT D'OBsolescence

Aujourd'hui, l'obsolescence est devenue un problème de plus en plus courant dans tous les secteurs, que ce soit dans l'aéronautique, l'automobile, la domotique, etc. Ce problème est principalement lié aux systèmes ayant un long cycle de vie [Zolghadri, M., et al. (2018)]: le système est conçu pour fonctionner pendant une longue période alors que certains de ses composants sont exposés à une certaine forme d'obsolescence et, par conséquent, la vulnérabilité et l'exposition à l'obsolescence du système augmentent. L'obsolescence ou la diminution des sources de fabrication et des pénuries de matériaux (DMSMS) est définie comme "la perte ou la perte imminente des fabricants d'origine des articles ou des fournisseurs d'articles ou de matières premières" [Sandborn, P. (2007)]. Dans [SD-22. (2016)], les auteurs ont clarifié la différence entre ces deux concepts en expliquant que "l'obsolescence ne conduit pas toujours à des problèmes de DMSMS, et tous les problèmes de DMSMS ne sont pas dus à l'obsolescence". Tant que le produit est encore en production ou disponible pour répondre à des demandes futures, il ne pose pas de problème de DMSMS. Mais, si les pièces de rechange sont obsolètes ou si le savoir-faire pour effectuer la réparation et la possibilité de tester le système (par exemple un équipement de test) après la réparation ne sont pas disponibles, cela entraîne un problème de DMSMS. Les auteurs dans [Pecht, M. G., & Das, D. (2000)] ont expliqué la différence entre obsolescence et interruption. En précisant que l'obsolescence se produit à un niveau technologique, alors que l'arrêt de production se produit lorsque le fabricant cesse de produire un article. Cependant, dans le domaine industriel, l'obsolescence comprend l'arrêt de la production d'un composant s'il n'y a pas d'autres fabricants pour ce composant spécifique. Le présent document traite le problème de l'obsolescence qui peut être dû à l'évolution technologique et/ou à l'absence de soutien de la part des fabricants. L'objectif est de minimiser son impact sur les coûts du cycle de vie, les performances du système, la disponibilité, la maintenabilité et la sécurité. Il est donc nécessaire de se concentrer sur la gestion de l'obsolescence tout au long du cycle de vie du produit.

2.1 Formes d'obsolescence

L'obsolescence est considérée comme un changement qui peut affecter le produit. Elle peut être considérée comme volontaire ou involontaire. Lorsque le changement est effectué par le fabricant lui-même pour promouvoir de nouveaux produits, augmenter la part de marché et les ventes [Déméné, C., & Marchand, A. (2015)], l'obsolescence est considérée comme volontaire. Sinon, l'obsolescence est considérée comme involontaire [Bartels, B., et al. (2012)]. Ce que nous appelons l'obsolescence naturelle. Dans la littérature scientifique [Bartels, B., et al. (2012), Sandborn, P. (2007), Mellal, M. A. (2020)], plusieurs typologies d'obsolescence ont été identifiées, dont les suivantes :

- Obsolescence technologique se produit lorsqu'il existe une nouvelle technologie qui peut remplacer l'ancienne.
- Obsolescence fonctionnelle est liée au fait qu'une ou plusieurs fonctions ne sont plus nécessaires.
- Obsolescence esthétique est liée aux effets de la mode et à la psychologie des consommateurs.

- Obsolescence logistique signifie que le produit n'est plus disponible en raison de la diminution des sources de fabrication et des pénuries de matériaux.
- Obsolescence économique est liée au coût élevé de l'utilisation, de la réparation et/ou de l'entretien du produit.

2.2 Conséquences d'obsolescence

L'obsolescence devient aujourd'hui un sujet d'actualité avec la croissance des équipements numériques avec des conséquences environnementales, sociales et économiques, voir Figure 1. En fait, elle peut avoir des conséquences sur exigences du système fonctionnelles (aptitude du composant à remplir sa fonction dans son environnement); les exigences non fonctionnelles (maintenabilité, viabilité, vulnérabilité, testabilité...); les exigences de performance (efficacité, temps d'exécution, vitesse de transport...); et les contraintes (dimensions, tensions d'alimentation...) [Trabelsi, I. et al. (2020)].

L'obsolescence a de fortes répercussions sur l'environnement. En effet, les produits obsolètes génèrent un énorme volume de déchets dangereux qui intensifient les émissions de gaz à effet de serre et la pollution. Ces déchets doivent être recyclés autant que possible. Mais pas tous les déchets sont recyclables, ceci conduit à un impact négatif sur l'écologie, en particulier dans les pays où le système de recyclage n'est pas correctement installé.

Évidemment, elle peut avoir une conséquence économique, car l'obsolescence fait partie intégrante du modèle économique. Il peut être considéré à la fois positif et négatif. L'impact positif est lié au fait qu'il fait fonctionner l'économie, grâce à la production de nouveaux produits, la consommation augmente et, par conséquent, l'économie prospère. En revanche, elle peut avoir un impact négatif car une consommation massive nécessite une exploitation intense des ressources naturelles.

Exigences du système	Environnementale	Sociale	Économique
<ul style="list-style-type: none"> • No respect de: • Exigences fonctionnelles • Exigences non-fonctionnelles • Performance • Contraintes 	<ul style="list-style-type: none"> • Déchets dangereux • Les émissions de gaz à effet de serre • la pollution • Épuisement des ressources non renouvelables 	<ul style="list-style-type: none"> • Problème de santé publique • Dette et exclusion sociale • Inégalités sociales 	<ul style="list-style-type: none"> • Produit à coût élevé • Coût de réparation élevé

Figure 1. Les différentes conséquences d'obsolescence

2.3 Causes d'obsolescence

Pour élaborer un plan efficace de lutte contre l'obsolescence d'une pièce ou d'un composant, il est essentiel de comprendre comment l'obsolescence peut se produire. Les raisons possibles de l'obsolescence sont nombreuses, comme montre la Figure 2. Certaines d'entre elles sont liées à l'évolution technologique ou à l'innovation qui rend un composant obsolète même s'il fonctionne encore et peut être produit et acheté. D'autres raisons sont liées au fournisseur, comme la disparition des fournisseurs d'origine du marché pour diverses raisons [Atterbury, R. (2005)], ou le fabricant de composants ou d'équipements d'origine n'est pas disposé à continuer à produire une pièce pour des raisons économiques, par exemple lorsque la demande d'un composant ou d'une technologie diminue et que le fabricant considère qu'il n'est pas rentable de poursuivre la production. L'obsolescence peut être causée par des directives, des règles et d'autres lois imposées par les gouvernements, telles que la directive relative à la limitation de l'utilisation de certaines substances dangereuses (RoHS)

[Brewin, A. (2005)], comme l'interdiction de l'utilisation du fréon en raison de ses caractéristiques d'appauvrissement de la couche d'ozone. Les produits peuvent également devenir

obsolètes s'ils ne sont plus durables en raison du processus de vieillissement des pièces stockées, qui rend leur utilisation impossible [Bartels, B., et al. (2012)].

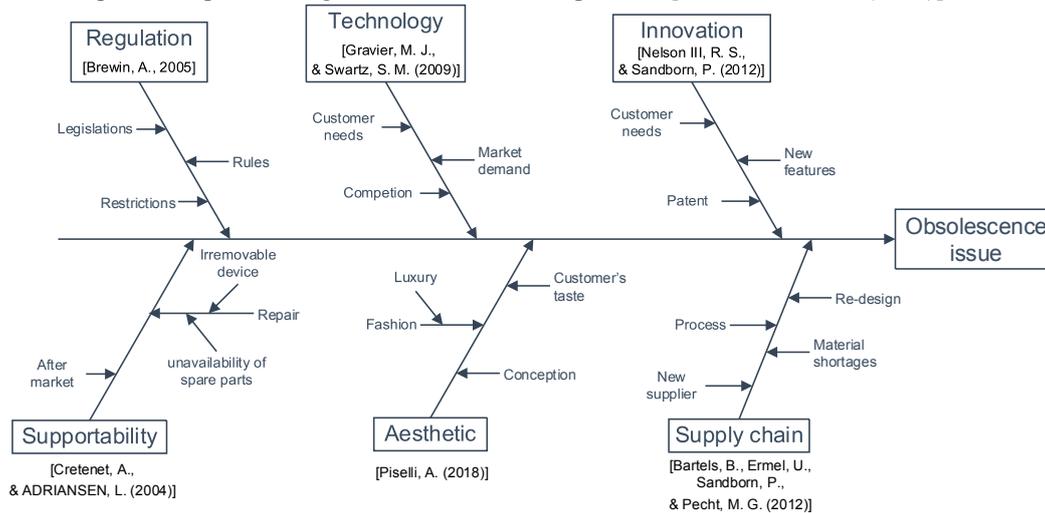


Figure 2. Les causes possibles d'obsolescence

Les causes d'obsolescence sont multiples. Dans ce travail, l'obsolescence est étudiée d'un point de vue technologique en étudiant les avancées technologiques qui rendent un produit obsolète.

2.4 Apprentissage automatique dans la prévision d'obsolescence

Pour réduire les problèmes causés par l'obsolescence, de nombreux chercheurs ont déjà travaillé sur la prédiction de l'obsolescence. La première méthode de prédiction de l'obsolescence a été mise au point dans [Solomon, R., et al. (2000)] en utilisant la méthode gaussienne pour estimer les phases du cycle de vie d'un produit et déterminer s'il est proche de l'obsolescence. Plus tard, les auteurs dans [Sandborn, P., et al. (2011)] ont proposé une technique basée sur la régression linéaire pour prédire l'obsolescence dans le cas de pièces électroniques sans facteurs paramétriques évolutifs clairs. D'autres méthodes statistiques ont été utilisées telles que la régression [Gao, C., et al.(2011)], les séries temporelles [Ma, J., & Kim, N. (2017)].

Selon [Sandborn, P., et al. (2011)], les techniques statistiques sont limitées à la prévision de petits changements futurs. Dans ce contexte, les auteurs de [Jenab, K., et al. (2014)], présentent un modèle qui inclut un concept de graphique de flux pour calculer le temps moyen jusqu'à l'obsolescence. Il est donc nécessaire de modifier ces approches afin d'améliorer les prévisions en cas de changements majeurs [Leclerc, A., & Nystrom, E. (2019)].

Ces dernières années, un grand intérêt a été accordé à l'utilisation d'apprentissage automatique pour améliorer la capacité d'analyse des données dans l'organisation. Dans ce contexte, les chercheurs se sont intéressés à appliquer des techniques d'apprentissage automatique pour la prévision d'obsolescence. Dans [Jennings et al. (2016)], les auteurs présentent deux méthodes basées sur l'apprentissage automatique : l'une appelée Obsolescence Risk Machine Learning (ORML) pour prédire l'état (obsolète ou disponible) des pièces et l'autre appelée Life Cycle Machine Learning (LCML) pour prédire la date d'obsolescence. L'étude de cas du smartphone a montré que l'algorithme de Random Forest est efficace en termes de précision, d'interprétabilité et de maintenabilité / flexibilité pour prédire l'état d'obsolescence. La même technique a été combinée avec l'algorithme

génétique méta-heuristique [Grichi, Y. et al. (2018)] et avec une optimisation méta-heuristique de l'essai de particules [Grichi, Y. et al. (2019)] pour optimiser le processus de prédiction d'obsolescence.

Les techniques existantes permettent de prédire l'obsolescence sans tenir compte de son évolution dans le temps. Afin de surmonter cette limitation, nous proposons une nouvelle approche basée sur l'apprentissage automatique pour prédire le degré d'obsolescence ainsi que son évolution dans le temps.

3 APPROCHE PROPOSÉE

De plus en plus d'entreprises disposent de grandes quantités de données qui constituent des ressources précieuses pour la gestion de l'obsolescence. Cependant, comme ces ressources ne peuvent pas être suffisamment analysées et évaluées, elles sont sans valeur pour l'entreprise. Pour surmonter ce problème, des techniques d'apprentissage automatique sont en cours de développement. Dans cet article, nous proposons une nouvelle approche qui vise à prédire l'évolution de l'obsolescence, voir Figure 3.

L'approche se compose de trois étapes principales:

1. Constitution d'une base de données de caractéristiques techniques.
2. Détermination de la valeur du degré d'obsolescence chaque année.
3. Proposition de modèles de prédiction.

La première phase dans toute méthode basée sur l'apprentissage automatique consiste à préparer la base de données nécessaire pour prédire l'évolution d'obsolescence. La base de données comporte des informations sur le produit étudié comme la date d'introduction dans le marché et les différentes caractéristiques techniques.

La base de données est ensuite subdivisée en deux sous ensembles. Un ensemble d'entraînement permettant d'estimer le modèle de prédiction afin d'être testé sur les nouvelles données de la seconde sous-base.

L'approche proposée nécessite l'utilisation des techniques d'apprentissage automatique. En effet, les techniques d'apprentissage automatique peuvent être classées selon deux catégories, à savoir l'apprentissage automatique supervisé ou non-supervisé. Les techniques d'apprentissage supervisé tentent de discerner les résultats et d'apprendre en essayant de

trouver des modèles prédictifs dans un ensemble de données étiquetées afin de prédire une étiquette. Une interaction humaine est nécessaire pour étiqueter les données. Par exemple, au fil du temps, alors que les utilisateurs marquent de plus en plus de courriels comme étant du spam et marquent d'autres courriels comme étant importants, les

clients de messagerie prédisent l'étiquette d'un courriel entrant en fonction des attributs des courriels passés comme certains comptes de mots. Les techniques d'apprentissage non supervisées ne reposent pas sur des données étiquetées et tentent de trouver des modèles descriptifs dans un ensemble de données sans interaction humaine.

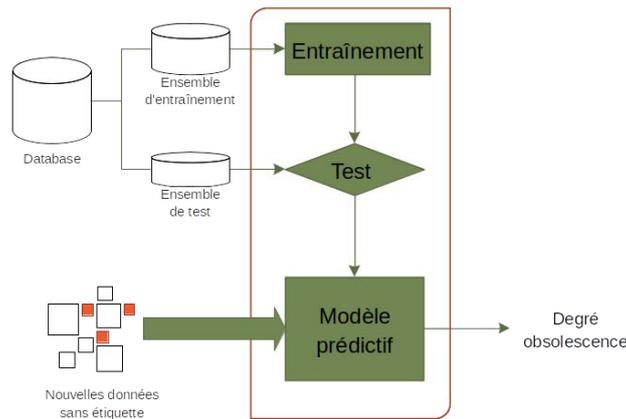


Figure 3. Proposition d'un modèle de prédiction du degré d'obsolescence

Product Name	Announced	Battery (mAh)	max battery Height (mm)	max Height Larg (mm)	Max Larg thickness	minthickness/Resolution	ppi	maxResolution	Internal	max internal	Bluetooth	maxbluetooth	Talk time	max talktime	pb_obsco			
Apple iPhone 5	2012	1440	3150	123.8	153.9	58.6	76.9	7.6	6.7	326	441	64	4	4	8	21	0.00	
	2013	1440	3610	123.8	170.7	58.6	88	7.6	6.2	326	441	64	4	4	8	28	0.00	
	2014	1440	4100	123.8	191.8	58.6	99.6	7.6	4.9	326	577	64	128	4	4.1	8	38	0.00
	2015	1440	4360	123.8	182.7	58.6	93	7.6	4.9	326	577	64	128	4	4.2	8	38	0.02
	2016	1440	5000	123.8	186.9	58.6	108.8	7.6	6.6	326	577	64	256	4	4.2	8	33	0.08
	2017	1440	5300	123.8	174.1	58.6	88.7	7.6	6.8	326	570	64	256	4	5.1	8	40	0.27
	2018	1440	5900	123.8	177.6	58.6	87.4	7.6	7.3	326	570	64	512	4	5	8	23	0.59
	2019	1440	6000	123.8	174.6	58.6	117.9	7.6	6.9	326	550	64	512	4	5.1	8	20	0.88
	2011	1432	1900	115.2	135.5	58.6	71	9.3	8.5	330	330	64	32	4	4	14	22	0
Apple iPhone 4s	2012	1432	3150	115.2	153.9	58.6	76.9	9.3	6.7	330	441	64	64	4	4	14	21	0
	2013	1432	3610	115.2	170.7	58.6	88	9.3	6.2	330	441	64	64	4	4	14	28	0.0144
	2014	1432	4100	115.2	191.8	58.6	99.9	9.3	4.9	330	577	64	128	4	4.1	14	38	0.1011
	2015	1432	4360	115.2	182.7	58.6	93	9.3	4.9	330	577	64	128	4	4.2	14	38	0.3547
	2016	1432	5000	115.2	186.9	58.6	108.8	9.3	6.6	330	577	64	256	4	4.2	14	33	0.7306
	2017	1432	5300	115.2	174.1	58.6	88.7	9.3	6.8	330	570	64	256	4	5.1	14	40	0.9598
	2018	1432	5900	115.2	177.6	58.6	87.4	9.3	7.3	330	570	64	512	4	5	14	23	0.9989
	2019	1432	6000	115.2	174.6	58.6	117.9	9.3	6.9	330	550	64	512	4	5.1	14	20	1
	2010	1420	2000	115.2	125	58.6	73	9.3	9.3	330	330	32	32	2.1	3	14	22	0
Apple iPhone 4	2011	1420	1900	115.2	135.5	58.6	71	9.3	8.5	330	330	32	32	2.1	4	14	22	0
	2012	1420	3150	115.2	153.9	58.6	76.9	9.3	6.7	330	441	32	64	2.1	4	14	21	0
	2013	1420	3610	115.2	170.7	58.6	88	9.3	6.2	330	441	32	64	2.1	4	14	28	0
	2014	1420	4100	115.2	191.8	58.6	99.6	9.3	4.9	330	577	32	128	2.1	4.1	14	38	0.0148
	2015	1420	4360	115.2	182.7	58.6	93	9.3	4.9	330	577	32	128	2.1	4.2	14	38	0.242
	2016	1420	5000	115.2	186.9	58.6	108.8	9.3	6.6	330	577	32	256	2.1	4.2	14	33	0.9513
	2017	1420	5300	115.2	174.1	58.6	88.7	9.3	6.8	330	570	32	256	2.1	5.1	14	40	1
	2018	1420	5900	115.2	177.6	58.6	87.4	9.3	7.3	330	570	32	512	2.1	5	14	23	1
	2019	1420	6000	115.2	174.6	58.6	117.9	9.3	6.9	330	550	32	512	2.1	5.1	14	20	1
2009	1219	1500	115.5	124	62.1	63	12.3	10.8	165	301	32	32	2.1	2.1	14	14	0	
2010	1219	2000	115.5	125	62.1	73	12.3	9.3	165	330	32	32	2.1	3	14	22	0	
2011	1219	1900	115.5	135.5	62.1	71	12.3	8.5	165	330	32	32	2.1	4	14	22	0.0016	
2012	1219	3150	115.5	153.9	62.1	76.9	12.3	6.7	165	441	32	64	2.1	4	14	21	0.0439	

Figure 4. Exemple de la base de données.

Dans cette approche, l'apprentissage supervisé est utilisé pour prédire le degré d'obsolescence d'un produit. Les attributs qui seront utilisés comme entrées sont donc les différentes caractéristiques techniques du produit ainsi que son degré d'obsolescence.

Pour cela, l'outil Google Trends est utilisé comme indicateur pour évaluer le degré d'obsolescence du produit. En fait, cet outil est en accès libre qui permet aux utilisateurs d'interagir avec des données de recherche sur Internet, qui peuvent fournir des informations détaillées sur le comportement des consommateurs. Il permet également de suivre l'intérêt des internautes dans le temps. Nous utilisons ici l'outil Google Trends pour mesurer l'intérêt des consommateurs pour chaque smartphone par an et supposons que le smartphone peut être considéré comme obsolète lorsqu'il n'attire plus les utilisateurs.

Ensuite, une étude comparative des différentes techniques est présentée pour identifier la technique la plus adoptée pour prédire le degré d'obsolescence d'un produit.

4 ÉTUDE DE CAS : L'OBSOLESCENCE DES SMARTPHONES

La problématique de l'obsolescence est particulièrement illustrée par les composants électroniques, notamment les

smartphones. Les smartphones sont sujets à un renouvellement rapide générant ainsi une obsolescence naturelle, qu'elle soit technique, logicielle ou esthétique. Par exemple, lorsque les fabricants proposent encore de nouvelles versions ou des mises à jour incompatibles avec les modèles précédents, on parle d'obsolescence logicielle. Mais lorsque l'utilisation de la batterie est réduite en raison d'un petit nombre de cycles ou que la réparation devient de plus en plus compliquée (avec des modèles presque impossibles à démonter et des pièces de rechange indisponibles), on parle d'obsolescence technique. À cette fin, notre approche est illustrée par l'étude de l'obsolescence des smartphones.

Cette étude comporte principalement les smartphones de haute gamme de la marque Apple, Samsung, Huawei, Oppo, etc. Pour collecter les caractéristiques techniques de chaque smartphone, nous avons utilisé le Web scraping sur l'un des forums de téléphonie mobile les plus populaires, GSM Arena [https://www.gsmarena.com]. La base de données contient plusieurs modèles de smartphones dont on connaît l'année de mise sur le marché et d'autres spécifications techniques. Il existe 57 caractéristiques techniques telles que la technologie numérique du téléphone portable, la capacité de la batterie,

les dimensions et le poids du smartphone, les caractéristiques de l'écran, etc., cf. Figure 4.

Le ratissage des données a été automatisé avec l'aide de la librairie selenium pour python et les données directement sauvegardées sous format CSV.

Pour obtenir des résultats de prévision plus précis, une étape cruciale concerne la sélection des caractéristiques pertinentes avant de démarrer le modèle de prédiction. À ce stade, un sous-ensemble des fonctionnalités les plus pertinentes est sélectionné.

Par conséquent, nous avons proposé d'utiliser une stratégie de vote par le biais de plusieurs algorithmes de sélection de variables tels que la sélection de caractéristiques Tree-based feature selection, L1-based feature selection, GenericUnivariateSelect, SelectPercentile, Random Forest, and SelectFwe. Ainsi, 8 de ces 57 caractéristiques peuvent être considérées comme les plus importantes affectant l'obsolescence des smartphones, comme le montre le Tableau 1.

Tableau 1: Les caractéristiques techniques les plus pertinentes choisies par les algorithmes de sélection de variables.

Caractéristiques	Vote
Épaisseur	5
Résolution	5
Stockage interne	5
Batterie (mAh)	5
Largeur	5
Longueur	5
Talk time	5
Bluetooth	5

Pour construire la base de données, on propose donc d'étudier l'évolution des caractéristiques techniques en fonction du temps, autrement dit, dès la date du lancement on détermine l'évolution de chaque caractéristique en la comparant au max (ou au min) pour les années suivantes. Par exemple, pour l'iPhone 3G, la capacité de stockage interne est de 8 Go au lancement en 2008, l'évolution de cette caractéristique est étudiée en la comparant au maximum pour les années de 2009 à 2018.

La détermination du degré d'obsolescence nécessite les données de ventes. La disponibilité des données est un problème majeur qui limite de nombreux domaines de recherche. Dans ce contexte, une discipline émergente telle que l'obsolescence souffre d'un manque de données utiles. Pour surmonter ce problème, nous proposons dans ce document de se référer aux données de Google Trends et de les considérer comme des données de vente.

La dernière étape dans la construction de la base de données est de déterminer les étiquettes des instances, ce qui correspond dans notre cas à la détermination du degré d'obsolescence. Pour ce faire nous avons utilisé les données du Google Trends. Les données de recherche des différents smartphones sont collectées en fichier CSV afin d'être analysées et interprétées. Nos recherches ont permis de démontrer que la distribution des données peut être considérée comme une distribution de Weibull. Pour des raisons de simplicité, nous n'avons pas inclus ces éléments dans cet article. Ainsi, en effectuant un ajustement de cette distribution, le degré d'obsolescence chaque année est déterminé par la fonction Cumulative de Weibull. Le processus de détermination du degré d'obsolescence est décrit dans la Figure 5.

Un exemple de la base de données est fourni par la Figure 4 ; les entrées du réseau sont les évolutions des différentes caractéristiques (les colonnes bleues) et la sortie désigne le degré d'obsolescence (colonne verte). Ensuite, l'ensemble de données a été divisé en deux groupes : un sous-ensemble pour l'entraînement et un sous-ensemble pour le test (voir Figure 3).

Plusieurs méthodes de régression supervisée ont été testées pour prédire le degré d'obsolescence d'un nouveau smartphone à partir du jeu de données étiqueté des smartphones antérieurs.

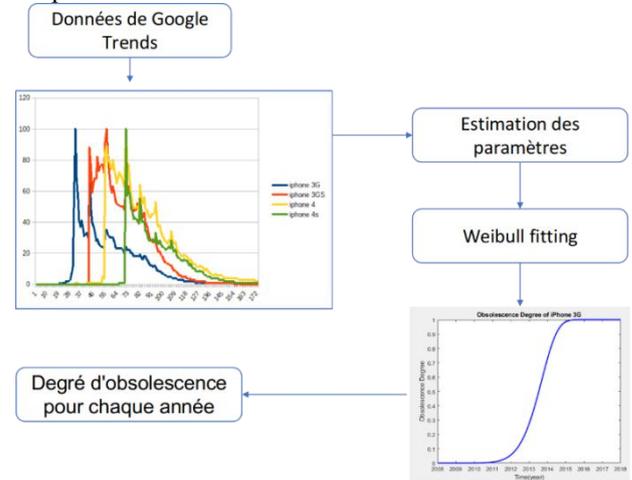


Figure 5. Processus de détermination du degré d'obsolescence.

Dans cette étude de cas, cinq algorithmes ont été utilisés pour créer des modèles de prédiction, les réseaux de neurones artificiels (ANN), les machines à vecteurs de support (SVM), la forêt aléatoire (RF), Méthode des k plus proches voisins (KNN) et la régression linéaire [Hastie, T., et al. (2009), Bonaccorso, G. (2017)]. L'ANN a été sélectionné en raison de son utilisation pour prédire l'obsolescence dans l'industrie de la presse. Dans la liste des "10 meilleurs algorithmes d'exploration de données"[Wu, X., et al. (2008)], RF, SVM, et KNN ont été classés parmi les meilleurs algorithmes d'apprentissage machine et d'exploration de données. La régression linéaire a été choisie pour le but de notre travail. Les performances de ces méthodes ont été mesurées avec l'erreur quadratique moyenne (RMSE) des valeurs prédites avec les valeurs de test.

Les résultats de comparaison entre les cinq techniques choisies (Tableau 2) montrent que la technique de forêt aléatoire (RF) présente l'erreur la plus petite. Cela signifie que la forêt aléatoire est le meilleur algorithme à utiliser pour prédire le degré d'obsolescence sur le marché des smartphones.

Tableau 2: Résultats de comparaison.

Erreur quadratique moyenne	
Méthode	
ANN	0.3749
RF	0.2031
KNN-10	0.2147
Régression linéaire	0.2159
SVM	0.2129

5 DISCUSSION

Dans cet article, nous proposons une approche hybride pour prédire l'évolution de l'obsolescence des systèmes au fil du temps. Cette nouvelle approche permet d'estimer l'évolution de l'obsolescence d'un composant en se basant sur les modèles d'obsolescence des composants en interaction avec celui-ci. Généralement, le modèle d'obsolescence est connu dans le cas de composants soumis à une obsolescence planifiée. Cependant, l'obsolescence planifiée d'un ensemble de composants accélère certainement l'obsolescence d'autres composants. C'est ce qu'on appelle obsolescence induite qui est une obsolescence indirecte causée plus par le comportement de l'environnement que par le comportement du composant lui-même.

Étant donné que les causes d'obsolescence sont différentes comme indiqué dans les sections précédentes, il est nécessaire de contrôler l'environnement du composant (ou du système) étudié pour une gestion efficace de l'obsolescence. Ainsi, une stratégie complète de gestion de l'obsolescence doit aller au-delà de la simple étude des exigences fonctionnelles et non fonctionnelles du système pour étudier son environnement. Ici, un environnement système peut inclure de nombreux autres facteurs tels que les réglementations, les concurrents, l'utilisateur, le climat économique et les fabricants, etc.

Pour aller plus loin dans ce contexte, la gestion de l'obsolescence basée sur les données peut être la solution adéquate pour collecter, structurer et évaluer les données pour une gestion d'obsolescence efficace et précise. À notre connaissance, il n'existe pas de cadre clair décrivant ce processus en normalisant ses étapes. Il est donc important de développer une feuille de route pour standardiser le domaine de la gestion de l'obsolescence basée sur les données.

6 CONCLUSION

La gestion du cycle de vie du produit est largement utilisée dans l'industrie pour guider les décisions et optimiser les coûts d'investissement. La détection et la prédiction de l'obsolescence des systèmes sont considérées comme la tâche la plus critique du processus PLM. Aussi, des modèles mathématiques et des techniques PLM ont été appliquées pour détecter l'occurrence d'obsolescence des systèmes. Toutefois, ces travaux sont limités dans la détection de l'obsolescence puisqu'ils ne tiennent pas compte de son évolution dans le temps. Dans cet article, nous proposons une méthodologie appropriée pour la prévision de l'obsolescence, basée sur la distribution de Weibull et l'apprentissage automatique. Enfin, une étude de cas réel a été utilisée pour valider cette nouvelle approche. La contribution concerne la prédiction de l'obsolescence en fonction de l'évolution des caractéristiques dans le temps. La méthodologie proposée est un premier cadre pour la gestion de l'obsolescence basée sur les données, qui peut être améliorée et complétée par d'autres aspects tels que l'impact économique. Il convient d'envisager l'élaboration d'une stratégie pour gérer le problème de l'obsolescence de manière proactive en tenant compte de l'environnement du système et des interactions qui le dépassent.

7 RÉFÉRENCES

Atterbury, R., Component Obsolescence Group, United Kingdom; *Minefield Obsolescence, A Guide to Tackling Disappearing Products*, Issue 2, 2005.

Bartels, B., Ermel, U., Sandborn, P., & Pecht, M. G. (2012). *Strategies to the prediction, mitigation and management*

of product obsolescence (Vol. 87). John Wiley & Sons.

Bonaccorso, G. (2017). *Machine learning algorithms*. Packt Publishing Ltd.

Brewin, A. (2005). *The Pb-free minefield: a guide to mitigating risk during the transition to RoHS compliance*.

Déméné, C., & Marchand, A. (2015). *L'obsolescence des produits électroniques: des responsabilités partagées*. In *Les ateliers de l'éthique/The Ethics Forum* (Vol. 10, No. 1, pp. 4-32). Centre rech. éthique Université de Montréal.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Gao, C., Liu, X., & Wang, X. (2011). *A model for predicting the obsolescence trend of fpga*. *Proceedings of 2011 9th Conf. Reliability, Maintainability and Safety* (1354-1358)

Grichi, Y., Beauregard, Y., & Dao, T. M. (2018). *Optimization of obsolescence forecasting using new hybrid approach based on the RF method and the meta-heuristic genetic algorithm*. *American Journal of Management*, 18(2).

Grichi, Y., Beauregard, Y., & Dao, T. M. (2019). *An approach to obsolescence forecasting based on Hidden Markov Model and Compound Poisson Process*. *International Journal of Industrial Engineering*, 1(2), 111-124.

Jenab, K., Noori, K., Weinsier, P. D., & Khoury, S. (2014). *A dynamic model for hardware/software obsolescence*. *International Journal of Quality & Reliability Management*.

Leclerc, A., & Nystrom, E. (2019). *Considerations for Obsolescence Prediction*. *Naval Information Warfare Center Atlantic Hanahan United States*.

Mellal, M. A. (2020). *Obsolescence—A review of the literature*. *Technology in Society*, 63, 101347.

Pecht, M. G., & Das, D. (2000). *Electronic part life cycle*. *IEEE Transactions on Components and Packaging Technologies*, 23(1), 190-192.

Pingle, P. (2015). *Selection of obsolescence resolution strategy based on a multi criteria decision model*.

Rojo, F. J. R., Roy, R., & Shehab, E. (2010). *Obsolescence management for long-life contracts: state of the art and future trends*. *The International Journal of Advanced Manufacturing Technology*, 49(9-12), 1235-1250.

Sandborn, P. (2007). *Designing for technology obsolescence management*. In *IIE Annual Conference. Proceedings* (p. 1684). *Institute of Industrial and Systems Engineers (IIE)*

Sandborn, P. (2007). *Software obsolescence-Complicating the part and technology obsolescence management problem*. *IEEE Comp. and Packaging Technologies*, 30(4), 886-888.

Sandborn, P. (2013). *Design for obsolescence risk management*. *Procedia cirp*, 11, 15-22.

Sandborn, P., Prabhakar, V., & Ahmad, O. (2011). *Forecasting electronic part procurement lifetimes to enable the management of DMSMS obsolescence*. *Microelectronics Reliability*, 51(2), 392-399.

SD-22. (2016). *Diminishing Manufacturing Sources and Material Shortages (DMSMS) A Guidebook of Best Practices for Implementing a Robust DMSMS*.

Solomon, R., Sandborn, P. A., & Pecht, M. G. (2000). *Electronic part life cycle concepts and obsolescence forecasting*. *IEEE Transactions on Components and Packaging Technologies*, 23(4), 707-717.

Trabelsi, I., Zolghadri, M., Zeddini, B., Barkallah, M., & Haddar, M. (2020, July). *FMECA-Based Risk Assessment*

Approach for Proactive Obsolescence Management. IFIP Conf. Product Lifecycle Mgt (215-226). Springer, Cham.

Trabelsi, I., Zeddini, B., Zolghadri, M., Barkallah, M., & Haddar, M. (2021, February). Obsolescence Prediction based on Joint Feature Selection and Machine Learning Techniques. In 13th International Conference on Agents and Artificial Intelligence (pp. 787-794). SCITEPRESS-Science and Technology Publications.

Zolghadri, M., Addouche, S. A., Boissie, K., & Richard, D. (2018). Obsolescence prediction: a Bayesian model. *Procedia CIRP*, 70, 392-397.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H. & Zhou, Z. H. (2008). Top 10 algorithms in data mining. *Knowledge information systems*, 14(1), 1-37.

L'ingénierie des systèmes et les réseaux bayésiens pour la conception résiliente à l'obsolescence

AMEL SOLTANE^{1,3}, SID-ALI ADDOUCHE¹, MARC ZOLGHADRI^{1,2}, MAHER BARKALLAH³, MOHAMED HADDAR³

¹ Laboratoire QUARTZ, 3 RUE FERNAND HAINAUT, 93407 SAINT-OUEN, FRANCE

² LAAS-CNRS, 7, AVENUE DU COLONEL ROCHE, TOULOUSE
amel.soltane@supmeca.fr

³ LA2MP LABORATORY, ROUTE SOUKRA KM 3.5, 3038 SFAX, TUNISIA

Résumé – L'obsolescence se produit essentiellement puisque les cycles de vie des composants sont plus courts que le cycle de vie du produit. L'émergence de nouvelles technologies influence les anciennes et les rend obsolètes. Ce phénomène a des conséquences sur l'architecture des produits et se propage dans toute l'architecture du produit. Cet article décrit le phénomène d'obsolescence et ses effets. Nous proposons une méthodologie pour modéliser la propagation de l'obsolescence et leurs conséquences dans l'architecture des produits. Elle contient différents outils pour représenter les dépendances, qui se transforment en canal de propagation de l'obsolescence, par des matrices de dépendances et pour l'analyser par des réseaux bayésiens.

Abstract – Obsolescence mainly occurs because component life cycles are shorter than the product life cycle. The emergence of new technologies influences old ones and makes them obsolete. This phenomenon has consequences on product architecture and propagates throughout the product architecture. This paper describes the phenomenon of obsolescence and its effects. We propose a methodology to model the propagation of obsolescence and its consequences in product architecture. It contains different tools to represent dependencies, which become a channel for obsolescence propagation, by dependency matrices and to analyze it by Bayesian networks.

Mots clés - Obsolescence, Architecture des produits, Dépendance et Réseaux bayésiens.

Keywords – Obsolescence, Product architecture, Dependency and Bayesian networks.

1 INTRODUCTION

Les produits et les systèmes doivent répondre aux besoins en constante évolution des clients. Ces produits et systèmes sont tenus de suivre cette évolution, en subissant parfois des changements dès les premiers stades du cycle de vie. Ces changements peuvent également être guidés par (i) les progrès de la recherche et du développement, (ii) les réglementations imposées par des organismes nationaux ou internationaux qui amènent les entreprises à modifier leur conception, ou (iii) les changements de conception imposés par la pression du marché et de la concurrence. Parmi les différents types de changements, notre travail se concentre sur l'étude de ceux générés par l'obsolescence. Elle correspond au fait qu'un élément devient soit inadapté aux demandes actuelles, soit indisponible auprès du fabricant d'origine, selon le Groupe d'experts [Group E, 2016].

Lorsque l'obsolescence affecte un composant, elle peut avoir un impact sur tous les composants avec lesquels ce dernier interagit. Ce phénomène de propagation peut se manifester par le biais d'échanges de données ou d'énergie par exemple. En d'autres termes, les conséquences de l'obsolescence se propagent par le biais de dépendances qu'il faut identifier. L'analyse de l'architecture du système peut permettre de les modéliser afin de limiter leurs conséquences sur l'ensemble du

système. De manière générale, toute stratégie d'atténuation proactive doit être basée sur l'architecture du système, afin d'être en mesure d'identifier les entités les plus sensibles et de prévoir, voire d'empêcher la propagation d'obsolescence. Souvent, après l'intégration des modifications de conception, il est nécessaire de s'assurer que le système fournit toujours ses services conformément aux exigences et que les nouveaux choix réduisent effectivement le risque d'obsolescence.

Notre étude se concentre sur la résilience d'un système aux changements et notamment ceux liés à l'obsolescence. La résilience à l'obsolescence est le fait que le système ou le produit reste « insensible » à l'occurrence d'obsolescence mais aussi aux changements imposés par la résolution de l'obsolescence. L'étape qui suit l'identification des composants et des fonctions d'un système, consiste à expliciter les dépendances en utilisant la modélisation de l'architecture du système. Nous avons choisi l'une des méthodologies d'ingénierie des systèmes les mieux établies, ARCADIA, développée par Thales. Parmi les raisons de ce choix, ARCADIA fournit une feuille de route claire pour la modélisation des systèmes à différents niveaux. De plus, elle est outillée par un environnement de modélisation open source : logiciel Capella [Roques, 2017]. Elle permet de

contrôler la validité des modèles à travers différentes étapes de modélisation. Sur la base des modèles de systèmes, nous les transformons en modèles de dépendance. Il s'agit d'un des apports à la problématique traitée. Ces modèles seront la base de l'analyse utilisant les réseaux bayésiens pour étudier l'impact des changements au sein de l'architecture d'un produit.

Afin de proposer une méthodologie soutenant une conception résiliente à l'obsolescence, on a recours à trois domaines (Figure 1) : l'obsolescence, l'ingénierie des systèmes (pour la modélisation de l'architecture) et les réseaux bayésiens (pour faire des analyses pronostic et diagnostic).

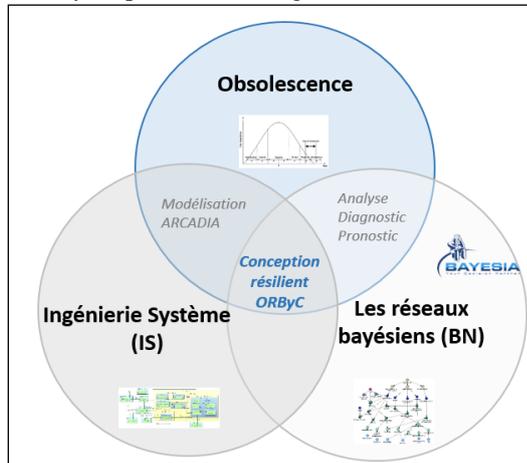


Figure 1. Le contexte de travail de recherche

Ces trois domaines sont représentés en détail dans la partie suivante. La propagation de l'obsolescence est abordée dans la troisième partie. On propose une méthodologie pour déterminer l'effet de cette propagation dans l'architecture produit dans la quatrième partie.

2 ETAT DE L'ART

2.1 Obsolescence

L'Oxford English Dictionary [Fowler et al., 1995] définit l'obsolescence comme "l'état d'être qui se produit lorsqu'un objet, un service ou une pratique n'est plus désiré, même s'il est encore en bon état de fonctionnement". L'obsolescence "se réfère aux matériaux, pièces, dispositifs, logiciels, services et processus qui deviennent non disponibles auprès de leur fabricant ou fournisseur d'origine" [Bartels, 2012]. Selon [SD-22, 2016], elle concerne le processus ou l'état par lequel une pièce d'équipement devient inutile, ou une forme et une fonction qui ne sont plus actuelles ou disponibles pour la production ou la réparation. En d'autres termes, un produit ou un système est obsolète lorsqu'il ne répond plus à la fonctionnalité attendue par les clients. L'obsolescence affecte le matériel et les logiciels.

Les principales raisons de l'obsolescence sont les progrès technologiques [Merola, 2006] (de nouveaux produits apparaissent à la place des anciens), le manque de soutien de la part des vendeurs [Merola, 2006] (l'organisation est obligée de modifier son produit pour obtenir les mises à jour nécessaires), la fusion et l'acquisition d'une entreprise [Bradley et Dawson, 1998] (l'organisation acquise peut être obligée de modifier son système existant) et un produit incompatible. Il est essentiel d'atténuer le risque d'obsolescence afin de minimiser les coûts tout au long du cycle de vie du produit. Il existe plusieurs stratégies d'atténuation. Sandborn [Sandborn 2013] définit 3 types de gestion de l'obsolescence : la gestion réactive, proactive et stratégique.

2.2 L'ingénierie des systèmes (IS) et ARCADIA

Selon l'EIA-ISO 632 [Technical Report, 2012], "L'ingénierie des systèmes est une approche interdisciplinaire qui englobe tous les efforts techniques visant à développer et à vérifier un ensemble de solutions pour les systèmes, les utilisateurs et les processus dans un cycle de vie total et intégré afin de répondre aux besoins des clients". Les méthodes d'ingénierie des systèmes sont utilisées pour modéliser des systèmes complexes. Pour faire face à la conception de systèmes complexes et pour améliorer la qualité et la productivité, l'ingénierie des systèmes se transforme en une approche basée sur les modèles (Model Based System Engineering MBSE) [Sebokwiki, 2019]. L'utilisation de modèles contribue à fournir une source d'information utilisée par l'équipe de développement. Ils facilitent l'intégration de la conception et des spécifications du système. L'ingénierie des systèmes fournit une approche basée sur des processus génériques et un ensemble de concepts (fonction, scénario, élément de système, exigence, etc.). Un processus décrit ce qu'il faut faire et il est composé d'un ensemble d'activités et de tâches organisées autour d'un objectif. Pour décrire le comment faire, on utilise des techniques et des méthodes de modélisation. Les activités et les tâches de l'IS sont des transformations de données génériques appelées entités. Chaque entité est caractérisée par des attributs spécifiques qui peuvent avoir des valeurs différentes. Pour suivre une séquence logique d'opérations, il faut expliciter les relations entre les entités. L'ontologie de l'IS est un ensemble d'entités et de leurs relations. Cette ontologie offre de nombreux avantages (1) un vocabulaire standardisé utilisé dans les méthodes et techniques de modélisation des données (2) l'identification des impacts des modifications dans les entités d'ingénierie système [Sebokwiki, 2019].

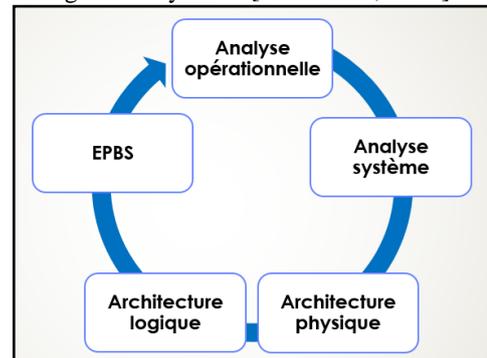


Figure 2. Les 5 niveaux d'ARCADIA

Il existe plusieurs méthodologies d'ingénierie des systèmes. ARCADIA est une méthodologie d'ingénierie basée sur des modèles pour la conception architecturale de systèmes complexes. Elle est apparue en 2007 dans le cadre du projet Thales Airborne Systems [Roques, 2015]. ARCADIA est supporté par Capella. Il contient un atelier de modélisation graphique [Huang et Mak, 1999]. ARCADIA utilise, avec Capella, un langage de modélisation des systèmes, largement inspiré d'UML et SysML. Ils permettent d'obtenir un ensemble de modèles bien informés. ARCADIA contient cinq niveaux (Figure 2). Pour chaque niveau, Capella propose un certain nombre de diagrammes (Tableau 1). Bien qu'ARCADIA soit une méthodologie de modélisation structurée, elle reste flexible. On peut créer des diagrammes dans un ordre "quelconque" et arrêter la modélisation à tout moment en fonction du niveau de compréhension attendu.

Tableau 1. Les diagrammes de Capella

Acronyme	Signification	Réalise
Analyse opérationnelle		
OCB	Operational Capacity Blank	Diagramme permet de décrire les acteurs et les entités opérationnelles.
OAB	Operational Architecture Blank	Diagramme permet de décrire les activités les acteurs et des entités opérationnelles.
Analyse niveau système		
SFBD	System Functional Breakdown diagram	Il permet seulement de créer des fonctions et des sous-fonctions.
SDFB	System Data Flow Blank	Diagramme permet de créer des fonctions et des sous-fonctions et de les relier par des échanges fonctionnels ;
SAB	System Architecture Blank	Diagramme décrire les fonctions du système.
Architecture logique		
LAB	Logical Architecture Blank	Il permet de créer des fonctions et des sous-fonctions.
Architecture Physique		
PAB	Physical Architecture Blank	Le diagramme de l'architecture finale du système avec la détermination de tous les composants du système.
EPBS		
EAB	EPBS Architecture Blank	Il déterminer les conditions que chaque composant doit remplir.

2.3 Les Réseaux bayésiens

Les réseaux bayésiens (RB) font partie de la famille des "Modèles graphiques probabilistes". Ces modèles sont développés pour enrichir le raisonnement des systèmes dont le fonctionnement est basé sur un ensemble de règles incertaines. Un RB est une technique d'encodage d'un ensemble de propositions d'indépendance conditionnelle [Russell et Norvig, 2009].

Il est un graphe dirigé avec deux propriétés : (i) propriété qualitative : le graphe est orienté et acyclique, (ii) propriété quantitative : les probabilités des paramètres représentés par les nœuds sont orientées de manière à respecter la condition markovienne selon laquelle la distribution de probabilité conditionnelle de chaque paramètre ne dépend que de ses parents. L'existence d'un lien orienté de X à Z signifie que l'état de Z dépend de l'état de X. Dans cette construction, le paramètre/nœud X est le "parent" et le paramètre/nœud Z est l'"enfant".

La règle de Bayes (1) nous permet de calculer la probabilité d'occurrence de l'enfant Z en fonction de l'occurrence du parent X et vice versa.

$$P(X|Z) = \frac{P(X \wedge Z)}{P(Z)} = \frac{P(X,Z)}{P(Z)} = \frac{P(Z|X)P(X)}{P(Z)} \quad (1)$$

3 PROPAGATION DE L'OBSOLESCENCE À TRAVERS L'ARCHITECTURE DU PRODUIT

Un système est composé de trois éléments. Le premier élément est le composant (c). Il est défini comme Ulrich dans [Ulrich, 1992] comme une région distincte du produit, permettant l'inclusion d'un sous-programme logiciel. Il possède un ensemble d'attributs (poids, dimension...). Le deuxième

élément est la fonction (f) qui est la transformation qui a contribué à la performance globale du produit [Ulrich et Eppinger, 2004]. L'association de fonctions à des composants donne des fonctionnalités permettant de satisfaire les besoins du consommateur. Cette fonctionnalité (φ) est le troisième élément du système.

Pour représenter l'interaction entre les différents éléments du système, nous pouvons utiliser la matrice *Design Structure Matrix* (DSM) [Browning, 2001]. Elle peut représenter la dépendance des éléments dans un même domaine. En outre, nous pouvons représenter la relation entre deux éléments appartenant à deux domaines différents en utilisant la matrice *Domain Mapping Matrix* (DMM) [Danilovic et Browning, 2004]. Pour représenter la dépendance dans le système, nous utilisons deux DSM et une DMM. Le premier DSM représente la dépendance entre les composants et l'autre représente la dépendance entre les fonctions. Le DMM est utilisé pour connaître les relations entre les fonctions et les composants. Pour représenter l'architecture du système, Gorbea propose une autre matrice MDM (Multiple Domain Matrix) qui est la fusion de DMM et DSM [Gorbea, 2008].

Si un composant devient obsolète (c2 par exemple dans la Figure 3), les autres composants qui ont des liaisons directes ou indirectes avec lui peuvent être affectés. Cependant, chaque composant de l'architecture remplit un ensemble de fonctions. Ainsi, les fonctions des composants obsolètes peuvent également être obsolètes (f2). En outre, l'architecture du système est la combinaison de composants et de fonctions pour répondre aux besoins des clients (fonctionnalités φ2 et φ3). Ainsi, à leur tour, ces fonctionnalités peuvent également être affectées. Le même raisonnement s'applique si l'obsolescence affecte d'abord la fonction. Elle se propage au composant qui la réalise et à la fonctionnalité. Si le client déclenche l'obsolescence, les fonctions et composants qui offrent ce besoin du client seront donc affectés.

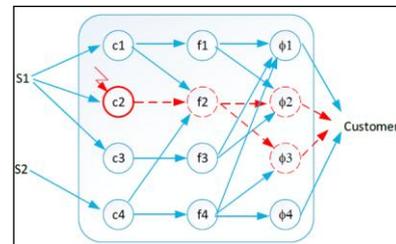


Figure 3. La propagation de l'obsolescence dans l'architecture

4 LA METHODOLOGIE D'OBSO-PROPAGATION

La méthodologie d'ingénierie du système ARCADIA utilise tous les modèles pour pouvoir maintenir le système à jour et pour le valider, voir la boucle (1) de la Figure 4. La méthodologie Obso-Propagation cherche à exploiter les interdépendances identifiées au sein de l'architecture du système de manière à déterminer quels composants, fonctions ou chaînes fonctionnelles pourraient être affectés par l'apparition de l'obsolescence, cf Figure 4.

Ces nouvelles connaissances permettront au concepteur de déterminer les composants du système les plus sensibles en cas d'obsolescence et de proposer ainsi des solutions résilientes. Il s'agit d'utiliser la boucle (2). Plus important encore, cette méthode contient essentiellement un graphe probabiliste pour tester et comprendre l'effet d'un changement causé par l'obsolescence sur l'ensemble de l'architecture. Tout cela est fait en tenant compte des scénarios possibles

d'apparition de l'obsolescence. Ainsi, Obso-Propagation est un outil d'aide à la décision permettant d'identifier les meilleurs scénarios de remédiation de l'obsolescence et de mettre en œuvre les choix de conception possibles. Ces nouvelles connaissances permettront au concepteur de déterminer les composants du système les plus sensibles en cas d'obsolescence et donc de proposer les solutions les plus résilientes. Il s'agit d'utiliser la boucle (2) dans la Figure 4.

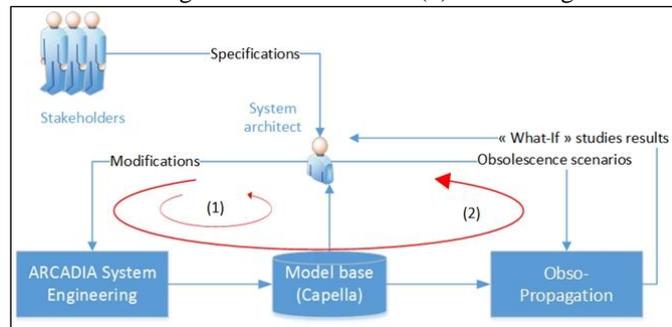


Figure 4. La méthodologie d'Obso-Propagation

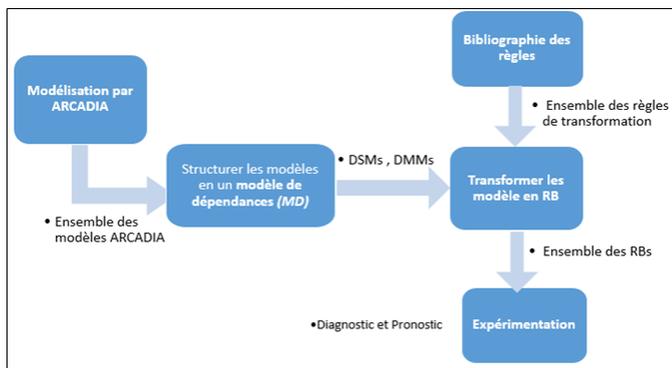


Figure 5. La méthodologie Obso-Propagation et l'ingénierie Système

La méthodologie Obso-propagation est basée sur 4 étapes (Figure 5).

4.1 La modélisation du système par la méthodologie ARCADIA

La première étape est la modélisation du système. Afin d'étudier la propagation de l'obsolescence, nous devons modéliser un système en détail. Dans cette étape, nous déterminons toutes les entités qui constituent un système (fonctions, composants, fonctionnalités,...) et toutes les dépendances entre eux (les relations entre les fonctions, les relations entre les composants et les relations entre les fonctions et les composants).

4.2 Extraction du graphe de dépendance

Après la modélisation du système, l'étape suivante consiste à représenter toutes les dépendances entre les entités du système dans un graphe ou une matrice. Cela se fait en exploitant les données de la première étape et en cartographiant toutes les dépendances possibles dans des matrices du type DSM et DMM.

Les DSMs obtenus contiennent les relations entre :

- Les entités externes : les liaisons entre les éléments de l'environnement du système.
- Les fonctions internes du système : les dépendances fonctionnelles internes.
- Les fonctions externes du système : les relations entre les fonctions des entités externes.
- Les sous-systèmes du système : les relations entre les composants logiques du système.

Les DMMs obtenues relient les relations entre :

- Les fonctions internes et externes du système : ces relations sont les chaînes fonctionnelles aboutissant à la réalisation des fonctionnalités attendues par les clients.
- Les fonctions et les composants du système : l'identification des rôles des composants – qui fait quoi ?
- Les fonctions externes (activités opérationnelles) et les entités externes : l'identification des activités des acteurs et des entités extérieures : qui fait quoi à l'externe ?
- Les composants et les sous-systèmes : appartenance des composants physiques (de l'architecture physique) aux composants logiques (de l'architecture logique) – qui héberge qui ?
- Les composants et ses éléments de la configuration : l'identification des éléments de configurations de tous les composants du système.

4.3 Création des réseaux bayésiens

En se basant sur les graphes de dépendance et une bibliothèque de règles de transformation préétablie, ces matrices sont transformées en Réseaux Bayésiens. Ces réseaux doivent être vérifiés et validés. Ils sont alors aptes à être utilisés pour des expérimentations.

4.4 Expérimentations

Une fois que les RBs sont entièrement définis, on peut déterminer les conséquences d'un changement dans l'architecture. Par exemple, si une fonction d'un composant physique devient obsolète, nous pouvons connaître la propagation de cette obsolescence dans l'ensemble du système. Il s'agit de calculer la probabilité d'un tel événement. L'évolution des probabilités doit être analysée et interprétée afin de tirer des recommandations.

Cette méthode est applicable au système complexe car elle est basée sur des graphiques probabilistes. En outre, elle peut aider les concepteurs à étudier l'impact des changements sur l'architecture du produit et à effectuer des simulations de scénarios diagnostiques (si les probabilités d'un nœud enfant sont modifiées) et pronostiques (si les probabilités d'un nœud parent sont modifiées) afin d'évaluer les risques et les coûts de chaque choix de conception du système.

5 CAS D'APPLICATION

On applique la méthodologie Obso-Propagation sur EOLE qui est un système de ballon à sonde. Sa mission principale est de fournir des données météorologiques à différents utilisateurs scientifiques (Figure 7). Il est constitué de deux sous-systèmes : un sous-système d'acquisition « en l'air » et un sous-système de traitement au sol.

On fait la modélisation d'EOLE en développant les diagrammes pour chaque étape d'ARCAIDA. Pour des raisons de simplification, on représente le diagramme PAB de niveau physique. Il contient tous les spécifications du système (Figure 6).

Après la modélisation du système par la méthodologie ARCADIA, nous pouvons en extraire des données. Ceci permet la création de matrices de dépendance de type DSM permettant d'associer des éléments de nature similaire et DMM permettant d'associer des éléments de domaines différents ; celles-ci sont présentées dans les Figures 10 et 11 en Annexe. Pour ce faire, on a principalement recours à

l'architecture physique qui décrit le système de façon - Pour les composants l'architecture : « Obso » ou « NN-obso » (peuvent être obsolètes ou non) ;

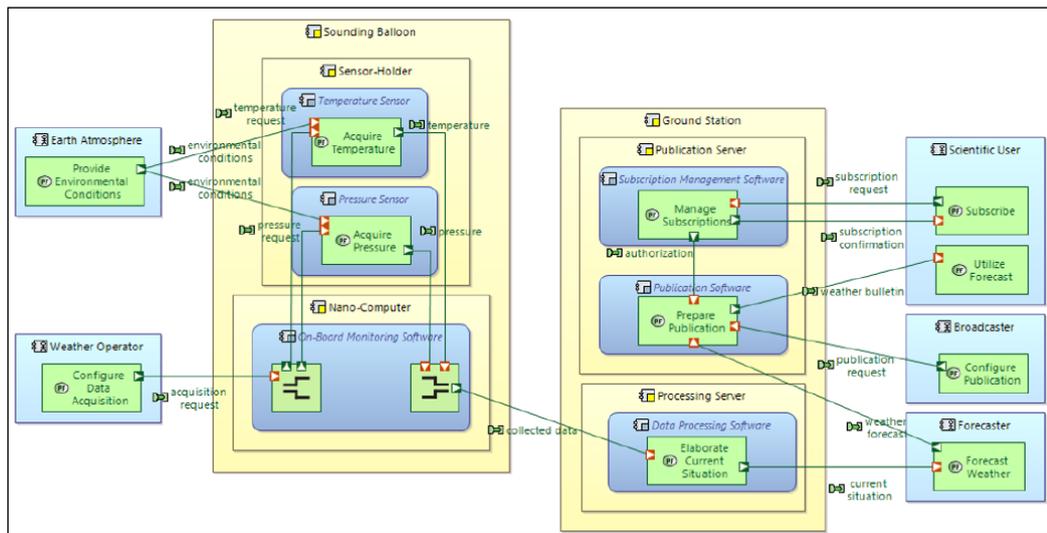


Figure 6. PAB (Physical Architecture Blank) d'EOLE

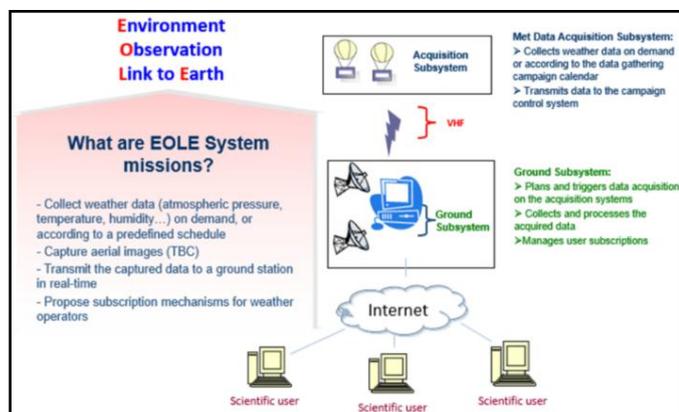


Figure 7. Le système EOLE

Chaque entité de l'architecture (composant, fonction, fonctionnalité) est associée à un nœud du réseau bayésien. Les liaisons de dépendance entre les entités sont représentées par des arcs. Dans l'exemple EOLE, nous essayons de déterminer le réseau bayésien (Figure 8) de la chaîne fonctionnelle d'acquisition des données.

Les nœuds du réseau sont : 6 fonctions (des nœuds en vert), 3 composants (des nœuds en rouge) et un acteur externe (nœud en gris).

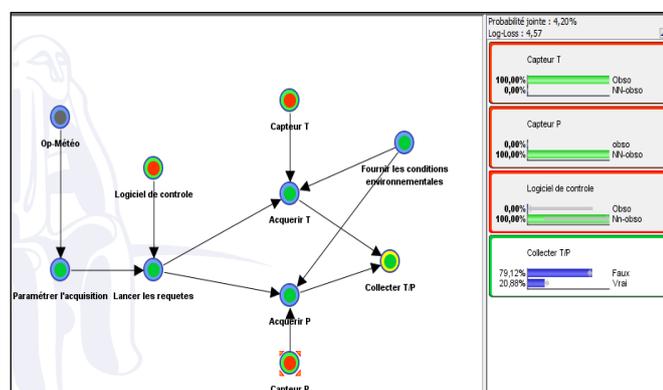


Figure 8. le réseau bayésien d'un sous-système d'EOLE
Pour chaque nœud, on détermine les modularités :

- Pour les fonctions : Vrai ou Faux (peuvent être réalisable ou non)

Les probabilités de distribution pour chaque modularité d'un nœud sont les probabilités d'obsolescence pour les composants ou les probabilités de réalisation pour les fonctions.

À droite de la Figure 8, on peut faire plusieurs observations, par le logiciel BayesianLab, en se basant sur les probabilités déjà définies. Par exemple, si on suppose que le capteur de température est à 100% obsolète et les autres composants (capteur de pression et le logiciel de contrôle) ne le sont pas, on voit alors l'influence de ce changement sur la fonction finale de la chaîne fonctionnelle d'acquisition des données (collecter T/P). On a alors une probabilité de 79.12% que cette fonction ne soit pas réalisable.

Il est important de connaître les caractéristiques des entités d'un système et de disposer d'indicateurs qui définissent la structure d'un graphique. L'outil Gephi nous aide pour cet objectif. On peut représenter les matrices de dépendance (DSMs et DMMs) sous forme d'un graphe tel que représenté dans la Figure 9. Gephi offre une liste d'indicateurs relatifs aux nœuds, aux arcs et au graphe, tels que des indices relatifs à la centralité. Le degré de centralité d'un nœud correspond au nombre de liens incidents à ce nœud et s'interprète, dans notre cas, comme une entité qui est en relation de dépendances avec d'autres entités. Plus le degré d'un nœud sera grand, plus l'entité correspondante sera dépendante d'autres entités.

Gephi permet la représentation des nœuds en fonction des valeurs des indicateurs. Par exemple dans la Figure 9, il représente en taille plus grande les nœuds qui ont des valeurs de centralité plus élevées.

On remarque que le nœud de plus grande centralité dans le graphe est la fonction « préparer la publication ». Il s'agit donc d'une fonction fragile et sensible à la propagation de changement. En effet, elle a plusieurs connexions avec les autres entités. Donc pour protéger le système global, on doit minimiser la probabilité d'affectation de ce nœud par l'obsolescence.

6 DISCUSSION DE RESULTAT ET CONCLUSION

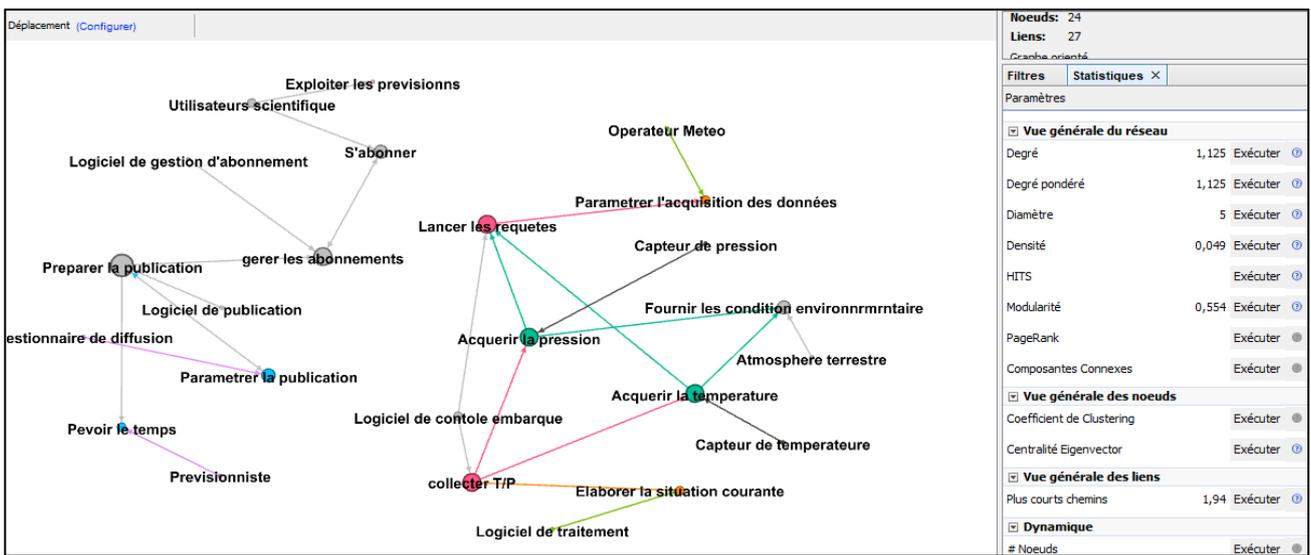


Figure 9. L'analyse de structure du système EOLE par Gephi

Ce travail présente une méthodologie structurée afin de déterminer la résilience des systèmes face aux risques d'obsolescence. L'obsolescence étant inévitable l'étude de la résilience prend tout son sens dès lors qu'on s'intéresse à des systèmes dont la durée de vie utile est supérieure à la durée de vie de leurs composants. L'approche proposée est basée sur l'exploitation des modèles issus de l'application des approches d'ingénierie système. Celles-ci, et en l'occurrence la méthodologie ARCADIA, offrent la possibilité de connaître finement non seulement l'architecture physique du système mais aussi ses capacités, ses fonctions, son architecture logique. La méthodologie ARCADIA conduit par ailleurs les architectes système à répondre à la question de "Make, Buy ou Make-to-Make" (composants fabriqués en interne ou achetés ou développés par des tierces). Ces connaissances profondes du système sont ainsi déterminées et structurées à travers les différentes étapes (i.e. niveaux) de la méthodologie. La résilience à l'obsolescence doit être comprise ainsi comme une caractéristique à étudier à travers tous ces niveaux de description des systèmes.

Le travail présenté dans cet article s'intéresse à l'architecture physique du système. Nous avons proposé de transformer les modèles de cette architecture en matrices DSM et DMM afin de déterminer les dépendances internes système. Ces matrices sont alors transformées en réseaux Bayésiens. L'intérêt de ces réseaux réside essentiellement dans leur capacité à déterminer les impacts potentiels de l'obsolescence des composants ou de leurs fonctions.

La méthodologie d'obso-propagation est une gestion stratégique de l'obsolescence visant à atténuer le risque d'obsolescence dès la phase de conception. Elle est utilisée pour tester le niveau de résilience du système face aux changements imposés par la gestion d'obsolescence.

Avec les changements rapides de technologies, besoins, performance, etc., le risque d'obsolescence augmente et la capacité d'adaptation du système diminue automatiquement. Cette méthode peut donc intégrer des scénarios d'obsolescence grâce à l'utilisation des réseaux Bayésiens. Différents types de preuves probabilistes (exactes, probabilistes, probabilistes et numériques) sont appliqués pour intégrer la probabilité d'obsolescence. L'étude de la résilience face à l'obsolescence permet par conséquent de détecter l'effet des obsolescences éventuelles sur le système en déterminant les sous-ensembles

les plus exposés. Elle aide notamment le concepteur dans le choix des éléments critiques mais aussi dans la construction des barrières techniques et technologiques dès la conception [Soltan et al., 2018].

7 ANNEXE

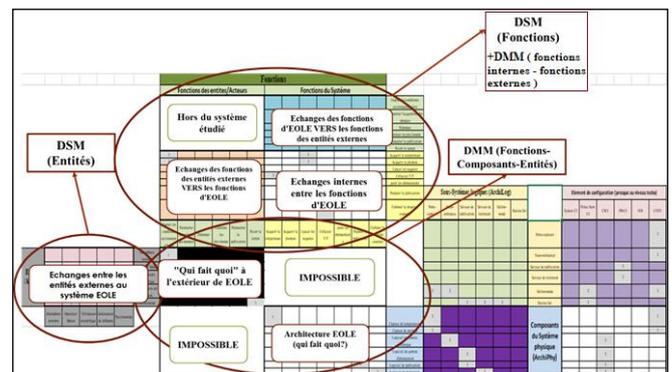


Figure 10. Les DSMs et DMMs d'EOLE 1/2

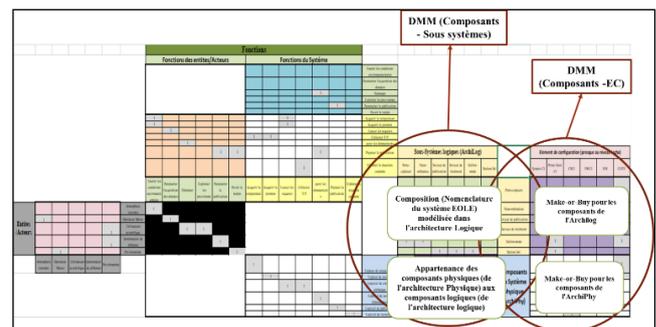


Figure 11. Les DSMs et DMMs d'EOLE 2/2

8 REFERENCES

- Group E. (2016). Obsolescence Management
- Roques, P. (2017). Systems Architecture Modeling with the Arcadia Method: A Practical Guide to Capella. Elsevier Sc.
- Thompson, D., & Fowler, HW., Fowler, FG. (1995) The concise Oxford dictionary of current English.

- Bartels, B., Ermel, U., Pecht, M., & Sandborn, P., Strategies to the Prediction, Mitigation and Management of Product Obsolescence. 2012.
- SD-22 DMSMS., (2016). A Guidebook of Best Practices, Defense Standardization Program Office.
- Merola, L., (2006). The costs software obsolescence threat. In: Fifth International Conference on Commercial-off-the-Shelf (COTS)-Based Software Systems, IEEE, pp 7–pp
- Bradley M, Dawson R (1998). An analysis of obsolescence risk in it systems. *Software Quality Journal* 7(2):123–130
- Sandborn P (2013) Design for obsolescence risk management. *Procedia CIRP* 11:15–22
- Technical Report., (2012). Processes for engineering a system. URL: <https://doi.org/10.4271/eia632>, doi:10.4271/eia632.
- Sebokwiki., (2019). Guide to the Systems Engineering Body of Knowledge (SEBoK).
- Roques, P., (2015). MBSE with the ARCADIA Method and the Capella Tool.
- Huang, G., & Mak, K., (1999). Current practices of engineering change management in uk manufacturing industries. *Int.J. Operations & Production Management* 19, 21–37.
- RUSSELL STUART, J. & NORVIG, Peter., (2009). Artificial intelligence: a modern approach. *Prentice Hall, 2009*.
- Ulrich, K., (1992). The Role of Product Architecture in the Manufacturing Firm
- Browning, TR., (2001). Applying the design structure matrix to system decomposition and integration problems: a review and new directions, *IEEE T.Eng.Mgt*, 48(3), pp 292-306.
- Danilovic, M., & Browning, T., (2004). A formal approach for domain mapping matrices (DMM) to complement design structure matrices (DSM), *Proceeding 6th Design Structure*
- Gorbea, C., Spielmannleitner, T., Lindemann, U., Fricke, E., (2008). Analysis of hybrid vehicle architectures using multiple domain matrices, DSMC
- SOLTAN, A., ADDOUCHE, S., ZOLGHADRI, Marc, & al., (2018). Obsolescence paths through the value chain. *Procedia Manufacturing*, vol. 16, p. 123-130.

Session 3.5 : Modélisation et gestion des entreprises, des connaissances et compétences

Chairman : Hervé Pingaud - Lilia Gzara

MOUNIA FERGUENE^{1,2}, NADIA LEHOUX^{1,2}, CAMÉLIA DADOUCHI^{2,3}

¹ Université Laval

Quebec, Canada

² FORAC research consortium

mounia.ferguene.1@ulaval.ca

Nadia.Lehoux@gmc.ulaval.ca

³ POLYTECHNIC MONTREAL, LID

Montreal, Canada

camelia.dadouchi@polymtl.ca

Résumé – La complexité et l'évolution de l'environnement d'affaires de l'industrie forestière ont soulevé de nombreuses incertitudes pour les décideurs en ce qui a trait aux tendances du marché. Cette recherche vise, à travers une revue systématique de la littérature, à décrire la dynamique du secteur par le biais d'une évaluation systématique des facteurs qui l'impactent, des comportements de marché attendus, ainsi que des opportunités et défis pouvant survenir. Cette analyse a d'abord permis d'identifier huit types de facteurs affectant l'industrie forestière, soit les facteurs i) économiques, ii) démographiques et sociétaux, iii) politiques et réglementaires, iv) environnementaux, v) technologiques, vi) mondiaux, vii) de substitution et viii) organisationnels. Par ailleurs, quoique la demande globale en produits forestiers soit prévue d'augmenter dans les prochaines années, notamment en raison de l'intérêt accru pour la bioéconomie, la substitution des journaux et du papier d'impression et d'écriture par les nouvelles technologies de l'information semble entraîner beaucoup d'inquiétude. La compétitivité de cette industrie semble par ailleurs grandement dépendre des prix des produits et des politiques et incitatifs gouvernementaux mis en place. Enfin, cet article met en lumière certains défis que rencontre l'industrie, tels que la capacité de sécuriser l'approvisionnement et les barrières au commerce international pour certains produits forestiers.

Abstract – The complexity and evolution of the forest industry's business environment have raised many uncertainties for decision makers regarding the progress of market trends. This research aims, through a systematic literature review, to depict the sector's dynamics through a systematic assessment of the factors involved, the expected market trends, and the opportunities and challenges that may arise. This review first led to the identification of eight types of factor affecting the forest industry, namely i) economic factors, ii) demographic and societal values, iii) policies, regulations and incentives, iv) global climate change, v) scientific and technological advances, vi) globalization, vii) substitution and viii) organizational factors. Furthermore, although predictions indicate that the overall demand for forest products should increase in the upcoming years, especially with the growing interest towards a bioeconomy, the growing use of information technologies that has been replacing printed and written documents seems to raise many concerns. The competitiveness of this industry also appears to depend highly on product prices as well as governmental policies and incentives. Finally, this paper highlights some of the industry's challenges such as the ability to secure wood supply and the international trade barriers for various forest products.

Mots clés – Industrie forestière, Produits forestiers, Tendances du marché, Étude prospective.

Keywords – Forest industry, Forest products, Market trends, Outlook study.

1 INTRODUCTION

The forest industry constitutes a major source of economic wealth in many countries. It creates jobs in rural areas, helps mitigate climate change, generates income, but also provides input for other industries [Lebedys, 2004]. After gaining maturity over the past decades, companies operating in this sector have continuously been seeking to improve their efficiency, through cost reduction and process enhancements, to keep up with the competition. However, the evolution of the external environment is reshaping the structure of the sector and creating uncertainties related to future production, consumption

and trade of forest products [Hansen et al., 2013]. To address these uncertainties, various approaches were used in the literature to predict the potential change in demand patterns, as well as the impact of external factors. In this context, Buongiorno (1996) presented the evolution of the quantitative approaches used to model the forest sector, from econometric methods, to linear programming, system dynamics, then complex models combining strengths of previous approaches. Hetemäki and Hurmekoski (2016) reviewed forest sector outlook studies to predict the consequences of the structural changes that occurred in the 21st century and, according to their findings, the current models failed to properly capture all of the changes hence the importance of complementing the current

modelling methods to better meet the needs of policy-makers. In another study, Hurmekoski and Hetemäki (2013) pointed out the scarcity of comprehensive market analyses for the forest sector. They also highlighted the need to increase the academic research in this field, because of the changing environment characterized by a growing competition, the shift of demand for traditional forest products, an evolution of the global economy and the emergence of new forest-based products. Therefore, this research aims to outline the dynamic of the forest sector and to highlight different factors, challenges and opportunities related to its future. To this end, a Systematic Literature Review (SLR) was conducted to define and answer three research questions:

Q1: What are the factors that impact the forest sector and orient the market demand?

Q2: What are the expected market trends in the forest sector?

Q3: What are the challenges and opportunities that may arise for the forest sector?

The SLR protocol involved to define a set of keywords related to the forest sector, the market evolution and the drivers of change so as to identify relevant articles from two different engines: *Engineering Village* and *Web of Science*. The research strategy resulted in 63 relevant articles. Analysis of these articles revealed eight categories of factors affecting this sector: i) economic, ii) demographic and societal values, iii) policies, regulations, and incentives, iv) global climate change, v) scientific and technological development, vi) globalization, vii) substitution, and viii) organizational. The main contribution of this paper is to provide a clear picture of the possible future of the forest sector, while guiding the different stakeholders involved in this industry in their decision-making process.

The remainder of this paper is organized as follows. First, the SLR methodology adopted to select the research papers is explained. A quantitative summary of the results obtained is then presented, followed by a content analysis answering the research questions. Finally, the main findings and limitations of the research are discussed in the conclusion.

2 RESEARCH METHODOLOGY

According to Tranfield et al. (2003), a Systematic Literature Review offers many advantages compared to traditional narrative reviews, by improving the quality of the evidence-based research, through the adoption of a reproducible and transparent process which minimizes bias. Therefore, the methodology suggested by these authors was followed to conduct a review that would be helpful for assessing the dynamic of the forest sector under different external factors, and the expected future demand for forest products. The details of each step are presented in the following sections.

2.1 Research questions

The first step of the methodology consisted of defining the scope of the study and delimiting the perimeter of the literature to be explored. Since the purpose was to explore the overall picture of the forest products market, the factors affecting it, the expected trends, and the main opportunities and challenges that may result, the research questions were formulated as follows:

Q1: What are the factors that impact the forest sector and orient the market demand?

Q2: What are the expected market trends in the forest sector?

Q3: What are the challenges and opportunities that may arise for the forest sector?

2.2 Search approach

To effectively answer the research questions and explore the related literature, a set of keywords and search fields were

defined after multiple trials. During these trials, new terms were extracted from the most relevant resulting papers, then tested in multiple search fields to target similar studies and complement the first keyword set. The final search was performed using three subsets of keywords linked with the operator "AND" on the engines *Engineering Village* (with databases *Compendex*, *Inspec*, *GEODATABASE*, *GeoRef* and *Knovel*) and *Web of Science*.

The first subset of keywords in Table 1 was used to target the forest sector. Then, two additional subsets were introduced to narrow down the scope of the search results to studies that focus on the evolution of the sector.

In *Web of Science*, the three subsets of keywords were used in the topic field. However, to reduce the irrelevant results generated in *Engineering Village*, the second two subsets of keywords were used in the abstract field, as shown in Table 1. It should be noted that more general terms, like "Forests" or "innovation" for instance, were eliminated as they generated a significant number of unrelated results. They were replaced by more specific keywords such as "Forest industry" and "Sector development".

Table 1 Summary of the keywords used for the SLR

Keywords	Search field (<i>Engineering Village</i>)	Search field (<i>Web of Science</i>)
"Wood products" OR "Wood product" OR "Forest industry" OR "Forest product" OR "Forest products" OR "Forest sector" OR "Wood biomass" OR "Wood pellets" OR "Woody biomass" OR "Sawn wood"	Subject/ Title/ Abstract	Topic
Demand OR Market OR Industry OR Markets OR Industries	Abstract	Topic
Foresight OR Outlook OR Forecast* OR Prospect* OR "Scenario analysis" OR "Scenario analyses" OR "Attractive markets" OR Evolution OR Driver OR Drivers OR "Sector development" OR "Sector developments"	Abstract	Topic

Given the importance of the structural changes in the forest sector since the beginning of the 21st century [Hetemäki and Hurmekoski, 2016], only recent papers published in English between 2005 and 2020 were selected. Both journals and conference articles were considered in *Engineering Village*, and articles, proceeding papers and reviews in *Web of Science*.

These evaluation criteria led to the selection of 390 references extracted from *Engineering Village* and 466 references from *Web of Science*. Afterwards, abstracts, titles and (when needed) full papers were examined to determine to which extent the articles were related to the main topics of the research and to the research questions. It is worth mentioning that 317 papers from *Engineering Village* and 424 papers from *Web of Science* were excluded as their focus was more on bioenergy extracted from other resources than forests, other materials than the forest products, the environmental footprint of the industry, the operational enhancement of the forest value chain processes, and the technical characteristics of a new product or a specific technology. While 29 articles from *Engineering Village* and 2 articles from *Web of Science* were excluded as they were not accessible for full paper examination.

As a result, a total of 63 articles were selected, namely 44 papers stemming from *Engineering Village* and 19 additional articles from *Web of Science* (21 papers were found in both databases).

3 RESULTS AND FINDINGS

In this step, a descriptive analysis was performed to identify the

distribution of publications per year and per region of interest, the objective and the scope of the studies, as well as the methodologies used in the literature. Then, the research questions were addressed through a content analysis.

3.1 Descriptive analysis

The number of relevant articles published per year is presented in Figure 1. It indicates a certain increase in 2011, 2017 and 2018, with a drop during recent years. It should be noted that since the review was completed on September 4, 2020, more articles could have been published.

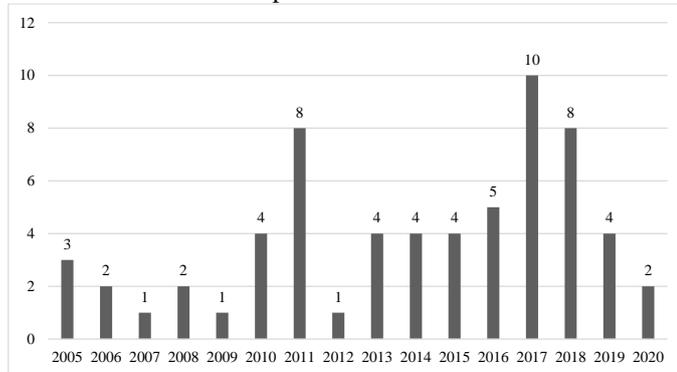


Figure 1 Number of publications per year

Figure 2 classifies the papers based on the research methodology adopted. However, if the authors employed multiple approaches to achieve their goals, the dominant one was assigned for the paper. The model-based category includes all quantitative modeling methods, such as econometrics and partial or global equilibrium models. Descriptive statistics concerns the studies where historical activity, such as production or consumption patterns, were summarized to provide a holistic overview. Case studies are extensive qualitative or quantitative research that focus on a specific phenomenon within an industry or a country. Finally, qualitative research consists mainly of the literature reviews, experts' analyses and the Delphi methodology. As can be seen in figure 2, quantitative model-based and qualitative methodologies are the most popular approaches.

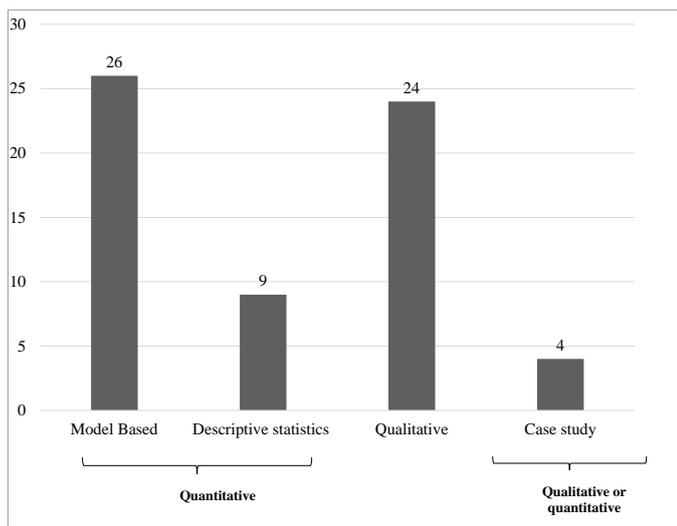


Figure 2 Number of publications per methodology

In addition, the types of studies were also identified and presented in Figure 3. Most of the studies consisted of predictive analyses with the purpose of anticipating the future of the industry or the impact of a specific factor, followed by trend analyses that mainly refer to forest-product market overviews with the aim of explaining the drivers behind these trends. Other

papers tackled the challenges and the ability of the forests to cover the growing demand. Few studies assessed the available modeling methods, while fewer evaluated the potential of the new forest products (besides wood bioenergy products), technologies, or services through innovation analyses.

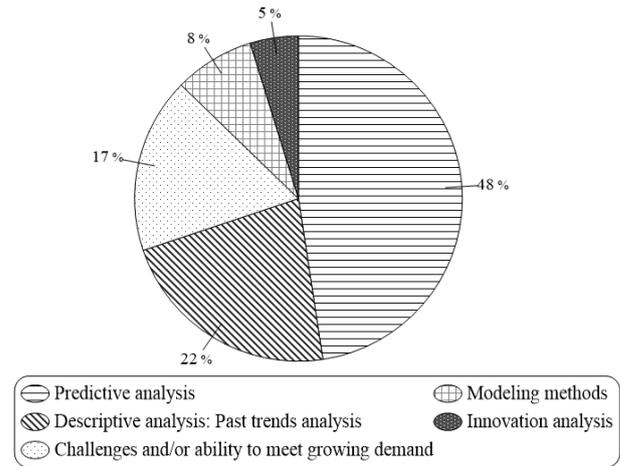


Figure 3 Reviewed articles per type of the study

Figure 4 presents the scope of the selected studies. However, those that did not focus on a specific product were grouped into the forest products category. This analysis shows that most papers focussed on bioenergy or the pulp and paper industry.

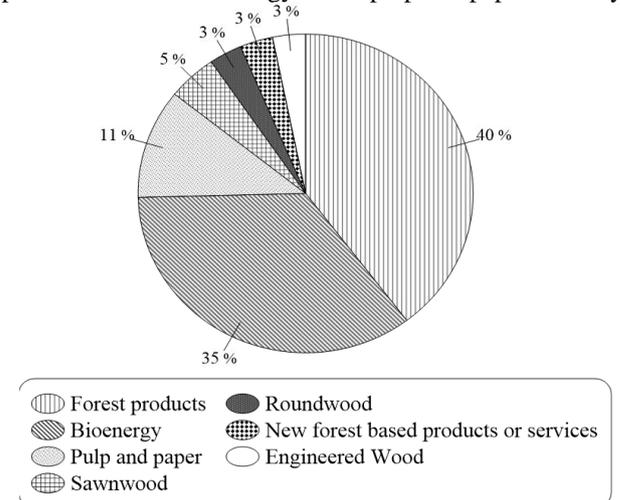


Figure 4 Reviewed articles per scope of the study

Finally, Table 2 classifies the publications per region studied. Those that did not focus on a specific country or region were placed in the global category. As can be seen, the majority of papers focused on European countries.

Table 2 Number of references per region

Region studied	Number of studies
Europe (including Russia)	26
Global	20
North America	11
Asia	4
Australia	2

3.1 Content analysis

Through the content analysis of the 63 selected articles, the previously defined research questions were answered. However, it is important to note that although all the 63 articles were analyzed for the review, due to the limited size of this paper, only the most relevant ones are presented in the following subsections.

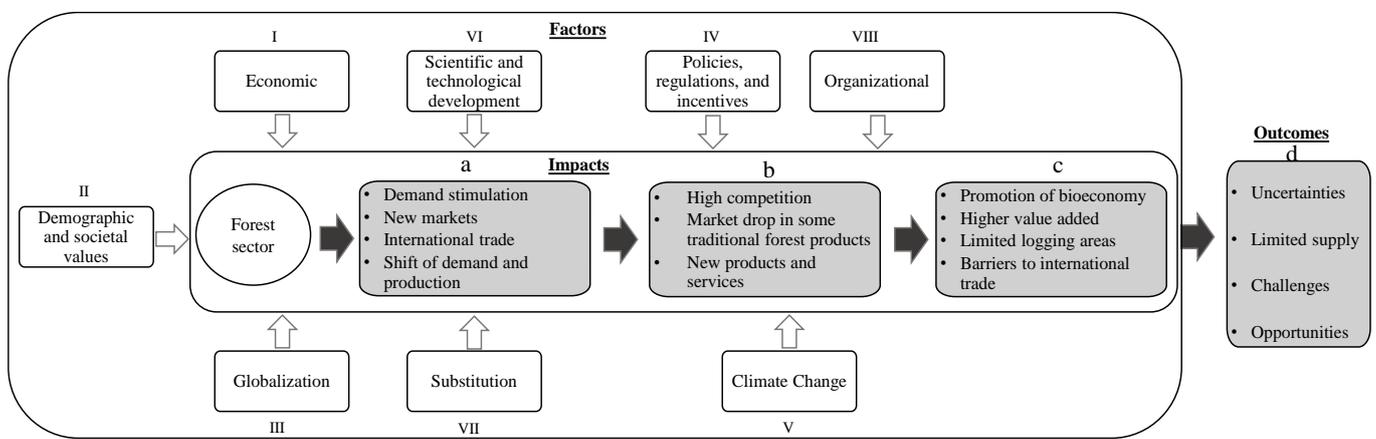


Figure 5 Dynamic of the forest sector under external factors

3.1.1 The factors impacting the forest industry

To answer the first research question “Q1: What are the factors that impact the forest sector and orient the market demand?”, the factors affecting the forest industry were grouped into eight categories: i) economic, ii) demographic and societal values, iii) globalization, iv) policies, regulations, and incentives, v) global climate change, vi) scientific and technological development, vii) substitution, and viii) organizational factors. To obtain these groups, the different factors were assessed from the selected papers, followed by an analysis of the categories previously established by the authors. The factors were then assigned to the categories identified, as well as to new ones in order to make sure that all the factors found would be taken into account.

As an example, Jonsson (2011) presented the factors driving long-term structural changes as being mainly derived from economic development, demographics, scientific and technological developments, globalization, global climate change, policies, regulations and customer preferences linked to climate change, environmental and other policies and regulations. Packalen et al. (2017) reviewed the possible drivers of change in the sawmilling industry through a PESTEC (Political, Ecological, Social, Technological, Economic, Consumer) framework using a Delphi analysis. According to their findings, threats and opportunities mainly arise from political decisions related to land use, building standards and wood-based energy. Technological and innovations also seem to have a positive impact on sawmilling industry.

However, to the best of our knowledge, no previous research has proposed a review of the factors affecting the forest sector through a systematic approach.

The resulting categories of factors as well as their impacts are presented in the next section. Figure 5 summarizes the categories of factors numbered from I to VIII, as well as their impact presented in boxes labelled from a to c and the outcome of these interactions shown in the box d.

Economic

Based on the articles investigated, the economic factors seem to be one of the major determinants of demand for forest products (Figure 5-I-a). The economic growth, usually reflected through the Gross Domestic Product (GDP), as well as the prices for raw materials and final products, can be associated to a large extent to imports, exports, and the consumption of forest products [Hetemaki and Mikkola, 2005; Kolo and Tzanova, 2017]. However, the high economic growth rates in developing countries, mainly located in Asia, compared to other regions particularly in Europe, is shifting demand and creating new markets for forest products [Jonsson, 2011]. Consequently, trade relationships with these countries will likely have strong

implications through the increased trades and overall demand [Erdozain et al., 2019, Joseph Buongiorno and Zhu, 2017].

The exchange rate also has a positive effect on local demand, but fluctuations in the exchange rate and global market prices have a strong impact on profit margins and competitiveness of forest products on international markets [Erdozain et al., 2019; Kolo and Tzanova, 2017]. Prices for substitution products, on the other hand, have a direct impact on the competitiveness of forest products (such as fossil fuel and green electricity for wood bioenergy) [Björheden, 2017; Sjølie et al., 2015].

Trend analyses in the literature highlighted a strong correlation between sawn wood consumption and housing activity, which itself is sensitive to economic crises (for example, the market drop in Finland and the USA, following the great recession) [Hurmekoski et al., 2015; Wear et al., 2016; Woodall et al., 2011]. The pulp and paper industry is similarly affected by these crises through the uncertainties that they create for the manufacturing activity. Since paper and cardboard are required for packaging, their consumption highly depends on the total manufacturing output [Wear et al., 2016; Stephens, 2013].

Demographic and societal values

According to Jonsson (2011), population growth may broaden the market and result in a higher demand for wood products. However, the unequal growth rates between developing and developed countries is shifting demand (Figure 5-II-a). In fact, global population is growing rapidly but is also diversified. Therefore, it is important to segment the forest products market to create customized and innovative products accordingly [Dasmohapatra, 2009].

Societal values and consumer preferences play an important role when it comes to selecting a specific product. Hurmekoski et al. (2015) observed a relation between abundance of forests and the consumption of sawn wood per capita, which can be explained by the societal preferences of the population for the economic utilization of forests. Moreover, society’s growing environmental concerns and recognizing wood as being environmentally friendly are expected to trigger demand for wood products [Jonsson, 2013], particularly building materials that qualify for green credits and bioenergy [Dasmohapatra, 2009].

Globalization

Globalization has stimulated international trade and created a global market for forest products (Figure 5-III-a) [Turner et al., 2005]. Trade policies and tax practices are considered as key elements to facilitate international trade, and they seem to be particularly present in Asian countries and Latin America [Buongiorno and Zhu, 2017]. These new conditions, among other previously discussed factors, can explain the continuously growing shift of demand, supply and production of raw

materials and final products from the Northern Hemisphere to the Southern one, which is seizing the competitive advantage from traditionally dominant countries with abundant forests. The forest industry has become more complex and spatially disconnected, especially with the access to raw materials at lower prices from intensively managed forests in countries with fast-growing trees and lower wages such as China, Vietnam, and the Philippines [Dasmohapatra, 2009].

Policies, regulations, and incentives

Government policies are more likely to promote the use of forest products, given their important positive impact, but they can sometimes present some challenges and inhibit the activity of the sector (Figure 5- IV-c).

The commitment of the European Union's (EU) member states to sustainability and reduction of greenhouse gas emission became the main driver of demand for wood energy. The bioeconomy is also incorporated into the general framework of the EU policy [Šupín et al., 2019]. To this end, action plans are developed to support and encourage the use of bioenergy, such as the taxation of fossil fuel carbon emission and the promotion of bioenergy through specific subsidies [Björheden, 2017]. As well, public initiatives were created to promote and finance businesses operating in the renewable energy field [Trømborg et al., 2008]. On the other hand, the alarming global climate change concerns may lead governments to impose logging bans and expand the protected forest areas [Smeets and Faaij, 2007; Wijewardana, 2005].

Cooperation is another incentive capable of promoting the sector. According to Pristupa and Mol (2015), it creates cross boundary networks between different stakeholders operating in renewable energy fields and stable links between producers and traders.

In addition to lobbying to change regulations and promote forest products may also occur in various groups of interest such as forest owners, service providers and institutions, especially in densely forested countries, to sustain and find new uses for forest products [Hurmekoski et al., 2015]. However, according to the authors, norms and incentives such as construction standards and regulations may limit or guide the use of sawn wood. Tariffs and non-tariff measures, such as government laws, regulations, policies and practices, to favour domestic products on the domestic market or to stimulate exports of a particular domestic product, can also constitute a challenge to the international trade of forest products [Wijewardana, 2005]. A good example encompasses the decisions by US administration which affect the Canadian forest industry, especially with the combined countervailing and antidumping duties on Canadian softwood lumber and newsprint recommended by the US Department of Commerce [Erdozain et al., 2019; Buongiorno, 2018]. Other country-specific decisions can also have strong implications on the forest sector, such as the decision of the United Kingdom to leave the European Union which has an impact on forest product consumption, production and trade [Johnston et Buongiorno, 2017].

Global climate change

Global climate change has a double effect (Figure 5-V-c). On one hand, it is expected to improve the productivity of forests with longer growth seasons and higher atmospheric CO₂. But on the other, natural disturbances such as fires, pest infections and severe drought may have a negative effect in the long and short term [Erdozain et al., 2019; Jonsson, 2011, 2013].

Scientific and technological development

Scientific and technological advances aim to continuously

improve the quality of trees, create new products, enhance the different processes of the forest sector, and increase the use of biomass from forests (Figure 5-VI-b) [Jonsson, 2011; Björheden, 2017]. However, they are also creating substitutes for forest products as explained in the next section.

Substitution

According to Latta et al. (2016), wood products can either replace one another, or they can be substituted by products from other sectors (Figure 5-VII-b). First, the most apparent substitution effect are the information and communication technologies threatening the pulp and paper industry by replacing newsprints and writing and printing paper. Although the effect is stronger in North America and the OECD countries, it will expand throughout the whole in the long term [Johnston, 2016; Latta et al., 2016; Turner et al., 2005].

In the energy sector, people have traditionally depended on fossil fuel, oil, and natural gas as direct substitutes of wood energy. Due to global environmental concerns, new competitors, especially from waste-derived fuels and corn ethanol as well as green electricity are gaining in strength [Šupín et al., 2019; Zhang et al., 2014; Sjølie et al., 2015; Trømborg et al., 2008]. In addition, wood for construction also has a competitive advantage since it is more energy-efficient and environment-friendly than steel frames in concrete-based houses. However, recent innovations in other competitive materials are expected to reduce or to allow no CO₂ emissions compared to traditional building materials [Hurmekoski et al., 2018].

Organizational

Organizational improvements throughout the value chain can increase the value of wood products. An example in the construction sector would be wood suppliers creating a firm specialized in wood construction to promote its use (Figure 5-VIII-c) [Hurmekoski et al., 2018; Hurmekoski and Sjølie, 2018].

3.1.2 Market trends

In this sub-section, the main expected market trends were assessed from the literature to answer our second research question "Q2: What are the expected market trends in the forest sector?". Considering the increasing interest for the pulp and paper as well as the bioenergy sectors in the literature, their main trends will be presented separately.

Global forest products

According to Buongiorno (2015), production, consumption, prices and trade of forest products in a country depend on the global supply and demand. Therefore, they used the Global Forest Products Model (GFPM) to predict the future of the sector under specific assumptions about macroeconomic conditions, particularly economic and demographic growth, and policy analysis. The long-term forecasts from 2011 to 2065 performed by the authors predicted an increase in demand and production of all forest products, namely round wood (industrial round wood, and round wood for energy generation), wood-based panels, paper, and paperboard.

Turner et al. (2005) used the same model to determine the possible future of the international forest products market and its implications for the USA in 2030, within the context of a growing globalization. They argued that forest product imports, as raw materials or final products, would be dominated by China, taking the leading position of Japan. Important new markets should also emerge such as Mexico and the Republic of Korea. As for the export of forest products, Canada should remain dominant and the USA should decrease its share for sawn wood and printing and writing paper but increase its

exports for round wood, paperboard and other types of paper. These results are partly in line with the most recent Asia-Pacific Forest Sector Outlook Studies released by FAO (APFSOS III) that predicted a strong demand for forest products to 2030 and to 2050 in the Asia-Pacific region, in accordance with the population and economic growth [Yasmi et al., 2019].

Sjølie et al. (2015) assessed the experts' expectations regarding the impact of key external factors on the Norwegian forest industry. Based on their observations, they suggested that sawn wood and bioenergy demand should expect to increase, and the society's environmental concerns should trigger demand for wood products, although there is a possibility that it will promote forest preservation and carbon sequestration instead.

Pulp and paper

The global shift in the pulp and paper sector combined with an overall decreasing demand pattern, have recently generated many uncertainties regarding the future of the sector.

Hujala et al. (2013) analyzed the trends and bilateral trade flows of bleached kraft pulp and recovered paper. Through this study, they argued that Asia, and particularly China, are the main drivers of trade in this sector, after being traditionally dominated by Europe and North America. This phenomenon may be amplified with the current context of globalization and an unequal economic growth.

The past trends analysis performed by Latta et al. (2016) showed that the observed decreasing consumption pattern of newsprints and printing and writing paper in the USA, followed by the OECD countries, was accompanied with the development of internet and communication technologies. Therefore, they integrated a new parameter in the demand equation to reflect the impact of internet use on the consumption of printing and writing paper. Based on these findings, Johnston [2016] updated the GFPM demand equation to project the market of pulp and paper products in 2030 under two scenarios: a full internet adoption by 2050 and by 2100. Their forecasts showed that demand for newsprint and printing and writing paper is foreseen to decrease in both scenarios, with a higher rate in the 2050 scenario where the internet adoption per capita is faster. Consequently, the global consumption of mechanical and chemical pulp wood is forecasted to decrease as well but with a relatively lower rate. The production of pulp and paper is similarly forecasted to decrease globally, and shift from North America mostly to Asian countries such as China and the Republic of Korea [Johnston, 2016], and some European countries such as Austria with lower wages [Turner et al., 2005].

Wood for energy use

With the current climate change concerns and the need to ensure a sustainable energy security, biomass as an affordable and practical renewable source of energy has recently known a growing popularity [Junginger et al. 2008]. A market analysis performed by Sikkema et al. (2011) revealed the EU's dominance in the production and consumption of wood pellets, led mainly by Sweden and Germany as major producers. However, while some European countries such as Germany and Austria are able to support their domestic needs, others are more likely to require imports of wood pellets, for instance the Netherlands, Belgium, Denmark and Italy.

Šupín et al. (2019) analyzed and forecasted the development of bioenergy policies in the EU after 2020. According to them, the EU could double the renewable share of its energy mix cost effectively, from 17% in 2015 to 34% in 2030 and biomass could remain a key renewable energy source beyond 2020.

Heinimö et al. (2011) also analyzed the potential of using second-generation biofuel in Finland to meet EU targets. Being

made of lignocellulosic from forest biomass and dedicated energy crops, second-generation biofuel production in the EU requires lower costs and offers a better reduction of CO₂, compared to first-generation biofuel. Yet, investments in this technology depend on economic factors and a high-profitability. As for global adoption of renewable energy sources in the EU, Finland and Sweden are leading the way with forest biomass being the largest energy source in both countries. For example, Finland has achieved a 38.7% of renewable energy share in 2014 while Sweden's renewable energy consumption represented 52.6%, with an even more ambitious target of zero Green-House Gas emission by 2050 [Björheden, 2017]. In contrast, the development of natural gas in North American countries in the last decades, with lower prices, has decreased the domestic market shares of wood energy [Sikkema et al., 2011]. According to Wear et al. (2016), the future potential of wood as an energy source remains unclear, especially with the uncertainty related to government policies regarding the use of wood for energy generation and the future production.

3.2 Opportunities and Challenges

Analyzing the literature to answer our third research question (Q3: What are the challenges and opportunities that may arise for the forest sector?) revealed the main opportunities and challenges presented in this section.

One of the main challenges that the forest industry is facing is the ability of forests to meet the growing demand for raw materials especially for energy use [Eker, 2014]. The availability of resources is limited by many factors: the conversion of forestlands to other uses, illegal logging and natural degradation of forests that limit the number of standing trees, and logging bans being established to cover these losses and preserve biodiversity [Wijewardana, 2005]. Hence, studies have analyzed to what extent forest resources are able to meet the needs for raw materials from forests. For example, using the GFPM, Buongiorno et al. (2011) projected the consequences of an increased demand for wood biofuel on the global forest, by comparing two future scenarios in 2030. Jonsson (2013) demonstrated, through the Wood Resource Balance tool, a possible lack of resources to cover the EU needs from domestic forests. Smeets and Faaij (2007) computed the quantitative results of an extensive literature review to predict the global need for forest products. According to them, the global needs should be met in 2050, but economic and ecological criteria may reduce the supply of wood from natural forests.

Furthermore, wood pellets production and trade are constrained by poor technological knowledge in terms of renewables, inadequate domestic market structure for renewables, the absence of national standards, as well as tariffs and non-tariff barriers [Junginger et al., 2008; Pristupa and Mol, 2015; Wijewardana, 2005]. In fact, tariffs on wood products are usually low, but escalated tariffs on higher value-added products can be considered as a barrier to international trade [Wijewardana, 2005].

On the other hand, the alarming concern related to the decreasing demand for newsprints and writing and printing paper is consistently drawing attention to the future of the pulp and paper industry and ways to overcome this huge challenge. Poudyal et al. (2017) suggested three alternatives that may compensate the major losses caused by the closure of pulp and paper mills and help the recovery of the economy: bio-oil, wood pellets and bioelectricity, and assessed their feasibility through an Input-Output (IO) analysis. Toppinen et al. (2017) analyzed how the pulp and paper industry in Europe should change strategically in 2030 from the experts' point of view. According

to the experts, one of the solutions to ensure the industry's survival is cost reduction, but also the capability to create innovative products and to respond to the growing need for bio-based products.

Hurmekoski et al. (2018) analyzed the potential of the emerging wood-based products: wood-based construction, biochemical, plastics and packaging, and textiles. According to their research results, summed up together, these products have a great potential and can increase the revenue of the forest industry. Besides, with prefabrication, standardization of wood construction and the development of engineered wood, wood can now be used for large scale construction and even compete with steel and concrete [Mallo and Espinoza, 2014; Hurmekoski et al., 2018].

Finally, the fast-growing interest towards bioenergy can open new perspectives for the sawmilling industry and encourage them to use the residues of their activity for energy production [Packalen et al., 2017], but also make use of the great biomass potential from forests particularly with the damages of the mountain pine beetle in Canada [Junginger et al., 2008].

4 DISCUSSION AND CONCLUSIONS

In this research, a systematic literature review was conducted, using specific keywords related to the forest sector and its future developments, to answer three research questions. These questions consist of the main factors affecting the forest sector, the expected market trends and the opportunities and challenges that may arise. Analyzing 63 relevant articles allowed us to assess eight major categories of factors driving or inhibiting the demand for forest products, notably i) economic factors, ii) demographic and societal values, iii) policies, regulations, and incentives, iv) global climate change, v) globalization, vi) scientific and technological advances, vii) substitution, and viii) organizational factors.

The interaction of these factors results in many possibilities for the forest sector. It generates opportunities and challenges but creates uncertainties regarding the future of the forest industry, as shown in Figure 5. Therefore, anticipating future market trends and demand patterns may turn out to be a complex task, requiring a deep analysis of these factors as well as their impact. While concerns are rising from the declining demand for some traditional forest products, new opportunities with promising potentials are flourishing, opening new perspectives for the forest-based sector. Consequently, the pulp and paper industry as well as the bioenergy sector have received considerable attention in the literature. The global shift towards a bioeconomy is foreseen to favour forest products for construction, energy generation and many other uses. In fact, wood bioenergy has great potential and may compensate the decreasing demand for newsprint and printing and writing paper. However, government commitment plays an important role in gaining competitive advantage against substitution products. Despite the emergence of new wood-based products, a few studies have assessed their potential on future markets and the consequences of diversifying the forest products.

This paper also shed light on some challenges that the forest industry may face such as the ability to supply the growing demand, the substitution of some traditional wood products, as well as the barriers to international trade.

Finally, it is worth mentioning that, although this research followed a systematic methodology, it may present some limitations related to the choice of keywords, search engines, the selection process that might result in a bias caused by the perspectives of the authors, in addition to the time scale considered. As the study was completed on September 4, 2020,

more articles could have been published without being included in this research.

The results of this paper offer a set of elements to be considered by policy makers as well as companies involved in the primary wood transformation industry so as to guide their strategic decision making and demand forecasting processes. Further research could also investigate the future of particular forest industries and integrate the results of this research to develop an adapted model, taking into account the highlighted factors and modeling methods.

5 BIBLIOGRAPHY

- Björheden, R. (2017). Development of Bioenergy from Forest Biomass—a Case Study of Sweden and Finland. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 38(2), 259–268.
- Buongiorno, J. (2015). Global modelling to predict timber production and prices: The GFPM approach. *Forestry*, 88(3), 291–303. <https://doi.org/10.1093/forestry/cpu047>
- Buongiorno, J. (1996). Forest sector modeling: A synthesis of econometrics, mathematical programming, and system dynamics methods. *International Journal of Forecasting*, 12(3), 329–343. [https://doi.org/10.1016/0169-2070\(96\)00668-1](https://doi.org/10.1016/0169-2070(96)00668-1)
- Buongiorno, J., Raunihar, R., & Zhu, S. (2011). Consequences of increasing bioenergy demand on wood and forests: An application of the Global Forest Products Model. *Journal of Forest Economics*, 17(2), 214–229. <https://doi.org/10.1016/j.jfe.2011.02.008>
- Buongiorno, J., & Zhu, S. (2017). Potential effects of a Trans-Pacific Partnership on forest industries. *Forest Policy and Economics*, 81, 97–104. <https://doi.org/10.1016/j.forpol.2016.08.001>
- Dasmohapatra, S. (2009). Future marketing drivers for the forest products industry. *BioResources*, 4(4), 1263–1266.
- Eker, M. (2014). Trends in woody biomass utilization in Turkish forestry. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 35(2), 255–270.
- Erdozain, M., Freeman, E. C., Ouellet Dallaire, C., Teichert, S., Nelson, H. W., & Creed, I. F. (2019). Demand for provisioning ecosystem services as a driver of change in the Canadian boreal zone 1. *Environmental Reviews*, 27(2), 166–184. <https://doi.org/10.1139/er-2018-0064>
- Lebedys, A. (2004). Trends and current status of the contribution of the forestry sector to national economies. *Forest Finance: Working Paper (FAO)*.
- Hansen, E., Panwar, R., & Vlosky, R. (2013). *The global forest sector: Changes, practices, and prospects*. CRC Press.
- Heinimö, J., Malinen, H., Ranta, T., & Faaij, A. (2011). Renewable energy targets, forest resources, and second-generation biofuels in Finland. *Biofuels, Bioproducts and Biorefining*, 5(3), 238–249. <https://doi.org/10.1002/bbb.291>
- Hetemäki, L., & Hurmekoski, E. (2016). Forest products markets under change: Review and research implications. *Current Forestry Reports*, 2(3), 177–188.
- Hetemäki, L., & Mikkola, J. (2005). Forecasting Germany's printing and writing paper imports. *Forest Science*, 51(5), 483–497.

- Hujala, M., Arminen, H., Hill, R. C., & Puumalainen, K. (2013). Explaining the shifts of international trade in pulp and paper industry. *Forest Science*, 59(2), 211–222.
- Hurmekoski, E., & Hetemäki, L. (2013). Studying the future of the forest sector: Review and implications for long-term outlook studies. *Forest Policy and Economics*, 34, 17–29. <https://doi.org/10.1016/j.forpol.2013.05.005>
- Hurmekoski, E., Hetemäki, L., & Linden, M. (2015). Factors affecting sawnwood consumption in Europe. *Forest Policy and Economics*, 50, 236–248. <https://doi.org/10.1016/j.forpol.2014.07.008>
- Hurmekoski, E., Jonsson, R., Korhonen, J., Jänis, J., Mäkinen, M., Leskinen, P., & Hetemäki, L. (2018). Diversification of the forest industries: Role of new wood-based products. *Canadian Journal of Forest Research*, 48(12), 1417–1432. <https://doi.org/10.1139/cjfr-2018-0116>
- Hurmekoski, E., & Sjølie, H. K. (2018). Comparing forest sector modelling and qualitative foresight analysis: Cases on wood products industry. *Journal of Forest Economics*, 31, 11–16. <https://doi.org/10.1016/j.jfe.2017.10.002>
- Johnston, C. M. T. (2016). Global paper market forecasts to 2030 under future internet demand scenarios. *Journal of Forest Economics*, 25, 14–28. <https://doi.org/10.1016/j.jfe.2016.07.003>
- Johnston, C. M. T., & Buongiorno, J. (2017). Impact of Brexit on the forest products industry of the United Kingdom and the rest of the world. *Forestry*, 90(1), 47–57. <https://doi.org/10.1093/forestry/cpw062>
- Jonsson, R. (2011). Trends and Possible Future Developments in Global Forest-Product Markets—Implications for the Swedish Forest Sector. *Forests*, 2(1), 147–167. <https://doi.org/10.3390/f2010147>
- Jonsson, R. (2013). How to cope with changing demand conditions — The Swedish forest sector as a case study: An analysis of major drivers of change in the use of wood resources. *Canadian Journal of Forest Research*, 43(4), 405–418. <https://doi.org/10.1139/cjfr-2012-0139>
- Junginger, M., Bolkesjø, T., Bradley, D., Dolzan, P., Faaij, A., Heinimö, J., Hektor, B., Leistad, Ø., Ling, E., Perry, M., Piacente, E., Rosillo-Calle, F., Ryckmans, Y., Schouwenberg, P.-P., Solberg, B., Trømborg, E., Walter, A. da S., & Wit, M. de. (2008). Developments in international bioenergy trade. *Biomass and Bioenergy*, 32(8), 717–729. <https://doi.org/10.1016/j.biombioe.2008.01.019>
- Kolo, H., & Tzanova, P. (2017). Forecasting the German forest products trade: A vector error correction model. *Journal of Forest Economics*, 26, 30–45. <https://doi.org/10.1016/j.jfe.2016.11.001>
- Latta, G. S., Plantinga, A. J., & Sloggy, M. R. (2016). The Effects of Internet Use on Global Demand for Paper Products. *Journal of Forestry*, 114(4), 433–440. <https://doi.org/10.5849/jof.15-096>
- Mallo, M. F. L., & Espinoza, O. A. (2014). Outlook for cross-laminated timber in the United States. *BioResources*, 9(4), 7427–7443.
- Packalen, T., Kärkkäinen, L., & Toppinen, A. (2017). The future operating environment of the Finnish sawmill industry in an era of climate change mitigation policies. *Forest Policy and Economics*, 82, 30–40.
- Poudyal, N. C., Joshi, O., Taylor, A. M., & Hodges, D. G. (2017). Prospects of wood-based energy alternatives in revitalizing the economy impacted by decline in the pulp and paper industry. *Forest Products Journal*, 67(7–8), 427–434.
- Pristupa, A. O., & Mol, A. P. (2015). Renewable energy in Russia: The take off in solid bioenergy? *Renewable and Sustainable Energy Reviews*, 50, 315–324.
- Sikkema, R., Steiner, M., Junginger, M., Hiegl, W., Hansen, M. T., & Faaij, A. (2011). The European wood pellet markets: Current status and prospects for 2020. *Biofuels, Bioproducts and Biorefining*, 5(3), 250–278.
- Sjølie, H. K., Bysheim, K., Nyrud, A. Q., Flæte, P. O., & Solberg, B. (2015). Future Development of the Norwegian Forest Industry, Based on Industry Expectations. *Forest Products Journal*, 65(3–4), 148–158. <https://doi.org/10.13073/FPJ-D-14-00061>
- Smeets, E. M. W., & Faaij, A. P. C. (2007). Bioenergy potentials from forestry in 2050: An assessment of the drivers that determine the potentials. *Climatic Change*, 81(3–4), 353–390. <https://doi.org/10.1007/s10584-006-9163-x>
- Stephens, M. (2013). Pulp and paper prospects in Australia: Looking to the future. *Appita: Technology, Innovation, Manufacturing, Environment*, 66(1), 13.
- Šupín, M., Loučanová, E., & Olšáková, M. (2019). SUSTAINABLE BIOENERGY POLICY FOR THE PERIOD AFTER 2020. DIGITALISATION AND CIRCULAR ECONOMY, 315.
- Toppinen, A., Pätäri, S., Tuppura, A., & Jantunen, A. (2017). The European pulp and paper industry in transition to a bio-economy: A Delphi study. *Futures*, 88, 1–14.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Trømborg, E., Bolkesjø, T. F., & Solberg, B. (2008). Biomass market and trade in Norway: Status and future prospects. *Biomass and Bioenergy*, 32(8), 660–671. <https://doi.org/10.1016/j.biombioe.2008.02.022>
- Turner, J. A., Buongiorno, J., Zhu, S., & Prestemon, J. P. (2005). The U.S. forest sector in 2030: Markets and competitors. 55(5), 10.
- Wear, D. N., Prestemon, J. P., & Foster, M. O. (2016). US Forest Products in the Global Economy. *Journal of Forestry*, 114(4), 483–493. <https://doi.org/10.5849/jof.15-091>
- Wijewardana, D. (2005). Market perspectives for timber from planted. *New Zealand Journal of Forestry*, v 50, n 1, p 34-41.
- Woodall, C. W., Ince, P., Skog, K. E., Aguilar, F. X., Keegan, C. E., Sorenson, C., Hodges, D. G., & Smith, W. B. (2011). An overview of the forest products sector downturn in the United States. *Forest Products Journal*, 61(8), 595–603.
- Yasmi, Y., Arts, B. J. M., & Hoogstra-Klein, M. A. (2019). *Forest Futures: Sustainable pathways for forests, landscapes and people in the Asia-Pacific region: Asia-Pacific Forest Sector Outlook Study III*. FAO.

Vers un processus de changement plus efficient

THIERRY SIEBENBORN¹, JEAN-LUC MAIRE², MAURICE PILLET²

¹ EMP-Consulting
144 zone Scheleck1 – L3225 BETTEMBOURG, Luxembourg
ts@emp-consulting.eu

² SYMME – Université de Savoie Mont Blanc
5 chemin de Bellevue – 74940 ANNECY, France
jean-luc.maire@univ-smb.fr
maurice.pillet@univ-smb.fr

Résumé – L’obligation constante pour les entreprises de s’adapter et de se renouveler mettent aujourd’hui en avant le besoin de disposer d’un processus de changement robuste et performant. Au-delà des méthodes ‘commerciales’ de conduite du changement, nous proposons une nouvelle approche de processus en ajoutant un axe à la représentation matricielle traditionnelle des processus de réalisation et de support, et en identifiant le long de cet axe des points clés de la phase transitoire du changement.

A chacun de ces points, nous équilibrons d’abord les ressources par excès dans le but de maintenir les performances de l’organisation et afin d’atteindre le jalon suivant et, au final, l’état cible après changement.

Nous abordons ensuite la question de la performance du changement en termes d’efficience. Celle-ci est obtenue en optimisant le choix des ressources nécessaires pour atteindre l’état cible. Nous nous appuyons pour cela sur une mobilisation des compétences, existantes mais parfois jusqu’alors non utilisées dans l’organisation dans son fonctionnement avant changement. Notre objectif est de développer un outil d’aide à la décision guidant le pilote du changement vers une transformation plus efficiente de l’organisation.

Abstract – The constant requirement for companies to adapt and reinvent themselves highlights the need for a robust and high-performance change process. Beyond ‘sales’ methods for change management, we offer a new approach that encompasses an axis to the matrix representation of realization and support processes, and that identifies key points of the transitory phase of change.

For each of these points, we first balance resources by excess with the objective of maintaining the company’s performance and reaching the next milestone, until achievement of the targeted state.

In a second phase, we consider change performance from an efficiency perspective. Efficiency is realized by optimizing the selection of resources required to achieve the targeted state, based on the mobilization of existing competences – which are sometimes concealed and unused within the organization.

Our objective is to develop a tool that facilitates decision-making for the change manager towards a more efficient organization.

Mots clés – transformation, changement, phase transitoire, compétences, efficience.

Keywords – transformation, change management, transitory phase, skills, efficiency.

1 INTRODUCTION

Dans le contexte de l'industrie 4.0, une entreprise doit aujourd'hui s'adapter en permanence à un environnement souvent fluctuant : évolution des marchés, mutations technologiques, évolution culturelle du personnel... Cela implique pour elle de passer d'un état de fonctionnement connu, avec ses forces et faiblesses, à un état cible dont les performances ne sont, à l'instant où la transition est envisagée, que théoriques. Incertain et surtout à chaque fois nouveau, le changement n'en est que plus difficile à réaliser et à piloter. Un changement peut s'appliquer à tout ou partie de l'entreprise (division, département, atelier, service, etc.) et peut prendre la forme aussi bien d'une série d'améliorations continues ou la forme d'une rupture radicale dans le fonctionnement de l'organisation. Si la nécessité de gérer des changements n'est pas nouvelle pour l'entreprise, ce qui l'est davantage en revanche, c'est l'augmentation quasi-constante ces dernières années de la fréquence avec laquelle ces changements doivent être menés. C'est cet acte de changement, et plus précisément la manière dont il s'opère ou devrait s'opérer, que nous abordons dans notre travail avec la proposition d'une approche de formalisation d'un processus de changement. Ce processus met tout particulièrement l'accent sur la manière de gérer et piloter la partie transitoire du changement qui sépare le début et la fin du changement, la plupart des démarches focalisant davantage sur le choix de l'objet du changement (processus, pratiques, flux, etc.) et/ou sur le choix des acteurs à impliquer (direction, équipe dédiée, opérationnels, etc.). Nous listons d'abord quelques familles de méthodes utilisées pour réaliser un changement et comparons leurs capacités respectives à gérer la partie transitoire d'un changement. Nous présentons ensuite notre approche du processus de changement qui prend la forme d'une séquence de trois phases et d'un pilotage à effectuer à plusieurs points remarquables. Nous montrons ensuite que le pilotage du changement à ces points s'opère par un équilibrage des ressources à réaliser et par une mobilisation des compétences associées à ces ressources. Enfin, nous faisons le lien entre les compétences mobilisées et la performance du processus de changement.

2 LES METHODES/OUTILS DE CHANGEMENT

2.1 Les familles de méthodes/outils du changement

Les méthodes/outils utilisées pour réaliser un changement ont en commun la caractéristique de structurer ce changement et de le piloter pour qu'il ait des chances d'aboutir. Cela nous permet de distinguer plusieurs types de méthodes :

- les méthodes de type top-down, où le changement est piloté par la "tête" de l'entité. C'est le cas par exemple du reengineering, [Hammer et Champy, 1996] le benchlearning [Karlöf et al., 2001] ou encore la méthode de management par les contraintes [Goldratt et Cox, 1984 et 1986].
- les méthodes de type bottom-up, où les pilotes du changement sont les acteurs directement concernés par le changement. C'est le cas du Kaizen [Imai, 1972], ou de la méthode implicitement associée au système Toyota [Ohno, 1990] à partir du modèle de production introduit en 1973.
- les méthodes caractérisées par un pilotage assuré par une équipe surfonctionnelle dédiée. C'est le cas des méthodes basées sur une gestion par projets [Cleland et King, 1983], des méthodes pilotées par une équipe Qualité comme la TQM (Total Quality Maintenance) [Feigenbaum, 1961], ou encore de la méthode Six Sigma [Harry et Schroeder, 2000] dont la conduite est assurée par une équipe qualifiée (Black Belt, Green Belt, ...).

2.2 Comparaison des méthodes

Le tableau 1 synthétise les particularités des méthodes précédemment citées (cœur de cible, caractéristiques du changement mené, facteurs de succès, vecteur de transversalité, sources éventuelles de difficultés) et complète cette description en fournissant :

- une représentation graphique illustrant l'allure générale de la phase transitoire du changement. Il s'agit de montrer des différences significatives entre les différentes méthodes en ce qui concerne la manière dont le changement est progressivement atteint : incrémental et régulière, par paliers, ...
- une évaluation de la difficulté à gérer et piloter cette phase transitoire du changement : difficulté peu importante (méthodes classées * ou **), difficulté modérée (méthodes classées ***) ou difficulté importante (méthodes classées **** ou *****).

Tableau 1. Analyse comparative des méthodes de changement

	Méthodes de type Top Down			
	BPR	Benchlearning	MPC	Hoshin
Cœur de cible du changement	Reconception des processus existants	Imitation / apprentissage de "best practices"	Maîtrise des goulots, équilibrage des flux	Mobilisation des ressources autour d'objectifs de progrès
Caractéristique du changement	Remise en cause profonde des processus	Substitution des pratiques usuelles par des "best practices"	Changement ciblant uniquement les ressources contraintes	Changement par percées
Facteur de succès	Communication sur l'intérêt d'un changement radical	Appropriation des pratiques par les utilisateurs	Résolution d'un seul problème à la fois	Définition et déploiement des objectifs de progrès
Vecteur de transversalité	Top-management	Acteurs du processus	Utilisateurs de la ressource goulot	Indicateurs de suivi des objectifs
Source de difficultés	Modification de modes de fonctionnement très ancrés dans la culture (processus ou pratiques)		Goulots pas toujours identifiables hors systèmes de production	Objectifs de progrès pas toujours cohérents et parfois antagonistes
Phase transitoire du changement				
Difficulté de transition	*****	***	**	**

	Méthodes de type Bottom-up		Méthode pilotée par équipe dédiée		
	Kaizen	TPS	Gestion projet	TQM	6 Sigma
Amélioration des équipements existants	Réduction/ suppression des gaspillages	Le projet comme support méthodologique au changement	Obtenir la satisfaction des clients par la Qualité	Réduction de la variabilité des processus	
Changements par "petits pas"	Amélioration graduelle à "peu de frais"	Changement structuré autour d'un projet	Amélioration continue	Management du changement basé sur la mesure	
Résultats visibles rapidement	Développement de l'autonomie	Equipe pluri-compétences	Mobilisation des ressources humaines	Maîtrise des outils statistiques de la qualité	
Ateliers "KAIZEN"	Groupes autonomes	Equipe projet	Groupe de progrès	Black Belt	
Insuffisance ou abondance de suggestions d'amélioration des acteurs	Longue durée de mise en œuvre et effort de longue haleine	Manque de communication entre équipe projet et acteurs du terrain	Nécessité d'une grande rigueur dans la gestion des enregistrements qualité	Une culture qualité comme pré requis indispensable à la méthode	
Phase transitoire du changement					
Difficulté de transition	*	**	***	**	****

Cette analyse montre que la partie transitoire du changement est inégalement prise en charge par les méthodes. C'est le cas pour :

- La méthode 6 Sigma — Pour beaucoup d'entreprises, cette méthode est perçue comme réservée à des spécialistes des outils statistiques. Le recours systématique à ces outils et leur nécessaire maîtrise par l'ensemble des acteurs rend en effet parfois difficile la mise en œuvre des actions décidées en vue du changement. De plus, bien que très codifiée (Green Belt, Black Belt, etc.), la structure d'accompagnement de la phase transitoire vers le changement requiert également un nombre important de ressources pas toujours disponibles dans l'entreprise.
- BPR — C'est sans doute avec cette méthode que la phase de transition du changement est la moins bien accompagnée. Les acteurs se trouvent en effet confrontés à la nécessité de s'adapter rapidement et brutalement à un processus totalement nouveau, décidé par le management, avec une production à maintenir alors que tous les points de repère et d'appui ont changé.

L'enjeu de nos travaux est donc de proposer une modélisation du processus du changement intégrant une discrétisation dans le temps de sa partie transitoire, et cela de manière à rendre possible sa gestion et son pilotage. Le paragraphe suivant résume notre approche.

3 NOTRE APPROCHE DU PROCESSUS DE CHANGEMENT

3.1 *Processus de changement et autres processus*

L'objectif de l'approche par processus dans les organisations était au départ de privilégier une vision transversale des activités en positionnant et structurant celles-ci autour d'un flux de réalisation du produit (figure 1a.). Pour représenter les processus qui interviennent en tant que supports aux processus de réalisation, l'approche processus a ensuite ajouté un second axe à la représentation, perpendiculaire au premier, pour structurer l'activité de l'organisation (figure 1b.). Cette approche à deux dimensions, largement reprise dans les cartographies de processus, donne une image de l'articulation des processus devant s'opérer pour garantir un fonctionnement standard de l'organisation.

Cette description des processus répond bien à un besoin de schématiser les interactions entre processus à un instant donné dans l'organisation et dans le cadre d'un fonctionnement bien stabilisé des processus. Elle est en revanche peu adaptée dans le contexte où un changement est conduit sur l'un ou plusieurs de ces processus. Un changement correspondant au passage d'un état stable de l'organisation à un autre, la représentation de ce changement ne peut en effet se faire sur le même plan que celui décrivant l'état initial des processus. Cela nous amène à ajouter un troisième axe perpendiculaire aux deux premiers (figure 1c.). Une graduation de cet axe en unités de temps contribuera, nous le verrons plus tard, à décrire le processus de changement comme une suite continue d'équilibrages de ressources à réaliser pour passer d'un état de l'organisation à un autre.

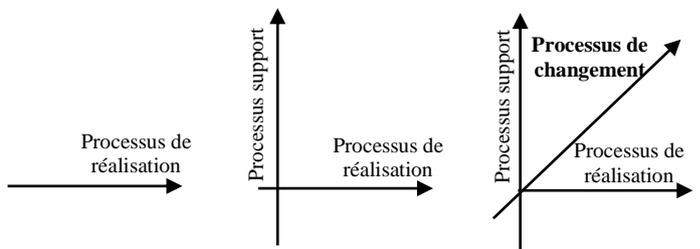


Figure 1a. Figure 1b. Figure 1c.
Cartographie des processus de l'entreprise

3.2 *Le changement : un découpage en trois phases génériques*

La proposition d'un découpage du changement est très ancienne puisqu'elle a été proposée dès les années 1940 [Lewin, 1951]. Rendu célèbre pour son modèle à trois étapes du changement organisationnel "*Dégel – Mouvement – Regel*", l'idée centrale est qu'une organisation doit être d'abord portée hors de ses valeurs et de son mode de fonctionnement courant, doit ensuite être changée, puis replacée sur de nouvelles valeurs et un nouveau mode de fonctionnement.

Avec les travaux plus récents de Alain Vas [Vas, 2002], le changement est perçu comme une décomposition en "*Planification – Négociation – Improvisation*" qui permet d'intégrer à la démarche de changement planifié les ajustements nécessaires en cours de processus.

Nous avons quant à nous pris le parti d'utiliser une décomposition empirique du changement en trois phases sous la forme de "*Préparation – Action – Mise au point*".

Les trois décompositions se rejoignent sur l'idée d'un premier temps dédié à une projection intellectuelle du changement, un second dédié à la mise en œuvre concrète d'un ensemble d'actions et enfin un dernier temps dédié à la stabilisation du nouveau système. Le caractère générique de ce découpage permet d'appréhender chaque changement, par nature unique et complexe, de manière identique et reproductible.

3.3 *Le changement : identification des points remarquables*

Cette approche du changement abordé comme l'enchaînement de trois phases génériques homogènes permet un premier cadrage codifié de l'acte de changement. La question qui se pose maintenant est de savoir comment ce cadrage peut aider à structurer un changement dans le temps. Nous nous interrogeons donc sur l'existence de certains points remarquables situés dans ces phases ou à leurs extrémités, communs à tout changement et pouvant de ce fait être considérés comme des éléments de discrétisation possibles d'un processus de changement.

Certains de ces points sont de nature "proactive", c'est-à-dire planifiables comme, par exemple, les quatre points situés aux extrémités des phases. Ces quatre points ont également un statut de jalon auquel est attaché un ensemble de critères à satisfaire pour passer à la phase suivante du processus : début du projet de changement, début de modification, fin de modification et fin du projet de changement.

Mais les quatre extrémités des phases ne sont pas les seuls points remarquables d'un changement. Une analyse plus détaillée d'un changement permet en effet d'identifier au moins deux autres groupes de points particuliers, révélés ou non selon des changements effectués.

Le premier groupe est lié à l'équilibrage, généralement planifié, qui est réalisé entre les ressources pour le processus

sur lequel est appliqué le changement. Les points recherchés sont des points d'équilibre naturel des ressources ou, plus exactement, des instants où la perception du risque demande une vérification du dimensionnement tant quantitatif que qualitatif de ces ressources. Selon la nature du changement et l'aptitude de son pilote, ces instants délimitent des zones à risque durant lesquelles les ressources sont à surveiller (par exemple, à l'arrivée d'un nouvel équipement, à la fin d'un apprentissage ou d'une formation, lors de la préparation d'une zone en vue d'un déménagement, en phase de test d'un nouveau logiciel, etc.). A ces points planifiés peuvent se rajouter d'autres points de contrôle positionnés en cours de réalisation du changement et non planifiés au début du changement. Il arrive en effet parfois que certaines zones à risques aient été sous-estimées par le pilote du changement. Lorsqu'une dérive est constatée par rapport à l'état cible, il s'agit pour le pilote de fixer un ou plusieurs points supplémentaires d'étude des ressources de manière à garantir le respect des objectifs initialement alloués. Ces points remarquables, bien que non planifiés et induits par les événements, sont de même nature que les précédents puisqu'ils concernent des équilibres de ressources passives. Généralement, et nous l'avons constaté à de nombreuses reprises, ces événements se produisent au cours des deux premières phases d'un processus de changement.

Le deuxième groupe de points particuliers est également constitué d'évènements non planifiables. Ces points coïncident avec les réactions à l'acte de changement. La complexité du facteur humain permet de percevoir le risque ainsi que son positionnement le long de l'axe temps du processus de changement, mais ne permet pas d'en évaluer ni l'instant d'occurrence et encore moins l'ampleur. L'expérience nous permet toutefois de considérer que si une bonne communication est menée dans le but d'éviter une réaction d'annonce, ces évènements se passent, s'ils ont lieu, dans la troisième phase du changement. C'est en effet à ce stade que les acteurs prennent enfin pleinement mesure des conséquences du changement sur leurs pratiques de la situation initiale. Il est important de souligner le caractère non déterministe de ce type de réactions. Une approche sécuritaire pourrait éventuellement conduire à planifier un point en phase de stabilisation.

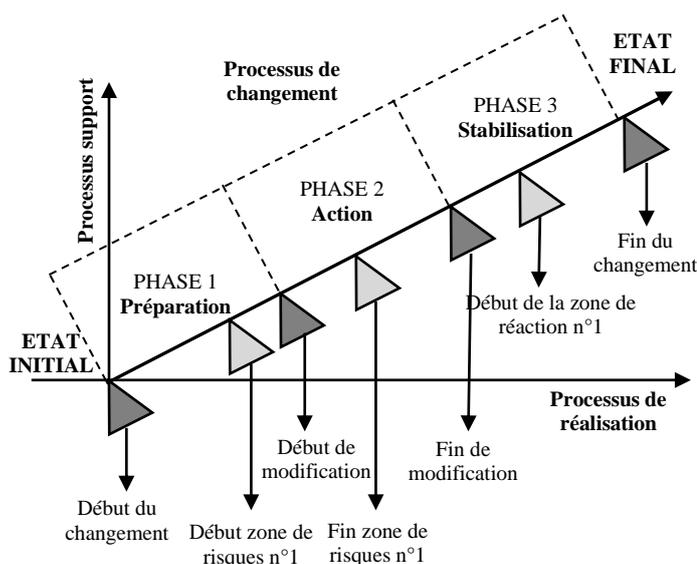


Figure 2 : Modélisation d'un processus de changement

Avec l'identification de ces points particuliers, un changement peut être discrétisé dans le temps. Cela permet au pilote de ce changement de gérer l'affectation de ressources en fonction des différentes phases du processus et de valider que les conditions sont réunies pour passer à la phase suivante. Cela nous amène à évoquer dans le paragraphe suivant le rôle central des ressources et des compétences durant la partie transitoire du changement.

4 LES RESSOURCES ET LES COMPETENCES DANS LE CHANGEMENT

4.1 Le rôle des ressources

Une ressource peut se définir comme un actif tangible ou intangible d'une entreprise [Penrose, 1972]. Elle peut s'apparenter à un moyen disponible ou requis et peut être financière, physique, humaine, organisationnelle ou technologique.

Dans un processus de changement, une approche consiste à s'intéresser aux ressources nécessaires à ce changement et à la manière de les mobiliser. Le niveau d'affectation de ces ressources pour obtenir le résultat de performance visé est généralement connu. En revanche, la manière dont ce niveau peut et/ou doit évoluer durant la partie transitoire du changement est peu abordé dans les méthodes de changement. Notre approche consiste à observer et, au besoin redéfinir, le dimensionnement de ces ressources, en lien avec l'observation de la performance aux points remarquables du processus de changement (Figure 3).

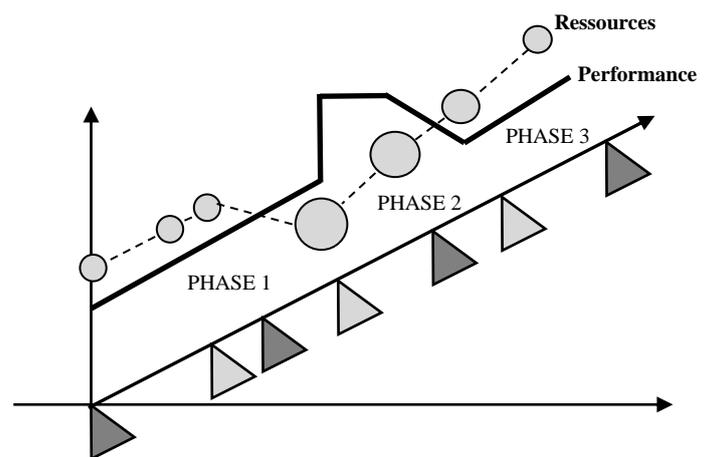


Figure 3 : Performance et affectation des ressources

Le constat d'une dérive peut induire le pilote, soit à accepter cette dérive jusqu'au prochain point, soit à réagir en créant un nouveau point intermédiaire avec, si besoin, une allocation de ressources supplémentaires. Cette allocation à réaliser doit être comprise comme l'ensemble des ressources nécessaires et suffisantes à mobiliser pour atteindre la cible de performance au prochain point et/ou pour maintenir une performance acceptable pour l'entité.

Cette modélisation, basée sur un équilibre des ressources à des points clés de la partie transitoire du changement, présente l'avantage de pouvoir s'intégrer dans les modèles existants de représentation des processus. Elle peut également être utilisée dans la mise en œuvre de l'ensemble des méthodes de changement listées précédemment.

4.2 Le rôle des compétences

Le pilotage de l'affectation des ressources est nécessaire pour passer d'un point du changement à un autre. En revanche, il existe généralement plusieurs affectations de ressources possibles pour faire ce pas et le choix par le pilote de la meilleure affectation est important.

Pour le comprendre, prenons l'exemple simple d'un changement consistant à transformer un système de stockage existant. Durant la partie transitoire d'un changement, des pièces situées dans une zone de stockage doivent être transférées vers un autre emplacement. Plusieurs solutions de mobilisation de ressources sont envisageables comme, par exemple, le recours en interne à une équipe de magasiniers, à un cariste ou encore le recours à un prestataire externe spécialement qualifié pour ce type d'opération. Si ces ressources peuvent toutes être utilisées, la question est de savoir quel sont celles qui se révéleront les plus adaptées pour atteindre la performance visée, et cela dans le contexte du moment où cette question est posée.

Des systèmes, comme par exemple ceux proposés dans [Berrah, 1997] ou [Covès, 2000], permettent de mesurer, a posteriori ou durant le changement, la performance d'une organisation. En revanche, ces systèmes ne donnent pas d'éléments sur la manière d'arbitrer le choix des ressources à affecter, parmi celles possibles, pour conduire le changement.

Les travaux décrits dans [LeBoterf, 2002] donnent quant à eux un éclairage sur les ressources de type homme à utiliser, et surtout, sur les compétences de ces ressources à mobiliser pour atteindre la performance souhaitée en fonction du contexte. Le problème se pose alors de savoir quelles sont les compétences requises et disponibles pour mener un changement. Pour évaluer les compétences, les entreprises utilisent généralement des grilles de notation à partir desquelles les aptitudes des collaborateurs sont comparées avec celles requises pour les fonctions qu'ils assurent. Le résultat est que l'adéquation entre les ressources et les fonctions est aujourd'hui assez bien réalisée dans les entreprises. En revanche, cela est moins le cas dans un contexte de changement où les besoins en compétences sont exprimés sans lien direct avec les fonctions habituelles des collaborateurs.

4.3 Les compétences face au changement de contexte

La notion de compétence comme une notion intermédiaire qui permet de penser les relations entre le travail et les savoirs détenus par des individus n'induit pas la performance comme l'obtention d'un résultat [de Terssac, 1996]. Des travaux récents [Magnani, 2018], axés sur la méthode de changement LEAN, abordent la question du rôle et de l'implication des individus dans le contexte d'une évolution par amélioration continue.

L'implication au travail est définie comme ce qui permet à un individu d'accéder à une partie de ses connaissances non mobilisées jusqu'à présent. Ceci permet d'éviter de formaliser un lien parfois difficile à établir entre compétence et contexte [Getz et Carney, 2012]. Les questions du pouvoir et de la liberté relative de l'acteur dans les organisations se posent comme des barrières à la libération des connaissances des individus dans un contexte d'accroissement des exigences de compétitivité et de perpétuelles transformations auxquelles sont confrontées toutes les organisations [Plane, 2019].

Le succès actuel du concept d'entreprise libérée [Chiquet, 2013] vise au même objectif, celui de ne pas se limiter à une partie contextualisée des connaissances des individus mais de miser plutôt sur une autoadaptation des individus à l'organisation par de nouveaux accès numériques à

l'information. Les effets d'une implication et d'un rassemblement des acteurs autour d'un projet commun sont indéniables, mais ils doivent également être mis en regard avec le risque d'aller trop vite dans le changement et de ne pas formaliser convenablement les actions engagées. Les résultats sont parfois mitigés lorsque le changement est trop dépendant de l'autonomie et de l'interaction des acteurs dans le temps. L'implication des acteurs est toutefois notable dans cette forme nouvelle d'organisation qui échappe aux règles déterministes traditionnelles.

Lorsque l'entreprise fait le constat de ne pas posséder toutes les compétences et les connaissances requises pour conduire un changement, elle est parfois amenée à envisager d'autres approches, comme par exemple celle de devoir travailler avec des partenaires extérieurs ou encore celle de devoir animer des réseaux internes et externes pour faire circuler les connaissances et les rendre disponibles là où elles sont utiles [Weil, 2003].

Toutes ces approches cherchent à éviter de devoir identifier des nouveaux besoins de l'état cible et des phases transitoires du changement, ainsi que la déclinaison de ces besoins en compétences. Malgré d'importantes avancées ces dernières années dans la capacité des entreprises à formaliser l'adéquation de leurs ressources à leurs besoins, un nouveau référentiel mieux adapté à un changement de contexte serait sans doute à construire.

En outre, comme nous l'avons vu précédemment, même si nous parvenions à réaliser une identification des besoins futurs, celle-ci ne donnerait pas d'indications sur la performance atteinte par l'organisation au cours du changement. Ceci nous amène à réfléchir sur le choix de la meilleure approche à utiliser pour qualifier des compétences aptes à maximiser les performances d'un processus de changement.

5 DES COMPÉTENCES AUX PERFORMANCES D'UN PROCESSUS DE CHANGEMENT

Nous avons proposé de modéliser la partie transitoire d'un changement sous la forme d'une séquence de plusieurs phases avec un jalonnement à plusieurs points remarquables. A ces points, le pilote du changement a en charge de dimensionner les ressources pour trouver un équilibre favorable pour atteindre la performance visée à la phase transitoire suivante. Nous allons compléter cette modélisation en intégrant les compétences à mobiliser. En effet, la simple connaissance du niveau de performance visé et des ressources disponibles ne suffisent pas à garantir le succès d'un changement.

5.1 Performance des processus

Avant d'analyser le cas spécifique du processus de changement, faisons un rapide tour d'horizon sur la manière dont les performances des autres processus sont abordées. Comme le précise l'ISO 9001 [ISO, 2015], '*la définition de la performance n'est pas un concept prédéfini. Il s'agit d'un ensemble d'éléments qui permettent la mesure*'. Dans ce même paragraphe, la norme propose également une matrice de notation de maturité dans laquelle sont évaluées les capacités respectives des processus à évoluer et à s'adapter au changement. Cette mesure peut être abordée selon trois critères complémentaires : l'efficacité (rapport entre ce qui a été réalisé et l'objectif initialement défini), l'efficience (rapport entre ce qui a été réalisé et les moyens engagés) [Baudot, 2019] et la valeur créée (vs les inputs) pour le client [Hammer, 1993].

Notre approche de processus de changement (une discrétisation en phases et avec un équilibre de ressources à des points remarquables) peut répondre au critère d'efficacité,

et cela en demandant au pilote de privilégier une affectation en excès des ressources pour atteindre le résultat de performance visé. Mais si nous voulons avoir une approche pertinente de la performance du changement, nous devons également prendre en compte le critère d'efficacité.

5.2 L'efficacité d'un processus de changement

L'efficacité est la capacité d'un individu ou d'un système à obtenir de bonnes performances. C'est également l'aptitude à parvenir au résultat avec une utilisation minimale des ressources. Dans notre processus de changement, l'efficacité peut se résumer comme l'apport minimum de ressources à engager pour passer de l'état initial à l'état cible. Cela nous amène à identifier les différentes solutions possibles d'optimisation des ressources nécessaires à la mise en œuvre d'un changement en fonction du contexte du changement.

A ce stade, il apparaît nécessaire de faire un point quant à la formalisation de la finalité du changement, en termes d'attendus mais également en termes de nouvel environnement à obtenir et de nouveaux besoins à couvrir. Par rapport au cahier des charges traditionnel, les modèles collaboratifs des méthodes dites agiles [Roto, 2010] aident déjà les entreprises à esquisser les finalités d'un changement par une description des nouveaux moyens, des nouvelles connaissances et/ou des nouveaux pilotages à mettre en œuvre. Il est évidemment possible d'intégrer cette description dans notre approche du changement aux points remarquables planifiés, et cela de manière déterministe et avant même que le changement ne débute. En revanche, en ce qui concerne les points remarquables non planifiés, un compromis doit être trouvé entre le temps de formalisation et d'analyse et la vitesse de réaction pour équilibrer le système et continuer sa progression vers la cible. A ces points non planifiés, le choix du meilleur compromis est très lié à l'identification du besoin dans son contexte particulier, de façon à savoir quelles ressources peuvent être et sont à mobiliser dans ce contexte. Par exemple, doit-on mobiliser les connaissances d'un consultant externe à l'organisation ou peut-on mobiliser des connaissances de nos collaborateurs détenues mais jusqu'alors non utilisées ? A-t-on besoin d'une machine spéciale de levage ou juste d'un outil à adapter sur un chariot de notre parc ? Dans notre logique de performance du changement, nous devons nous assurer du juste dimensionnement de ces ressources, de challenger les différentes solutions pour s'assurer de passer d'une stratégie de mobilisation en excès de ressources (recherche de l'efficacité) à une mobilisation au plus juste (recherche de l'efficacité).

Se limiter à qualifier les compétences des ressources en fonction de formations suivies pour les ressources hommes et en fonction de quelques expériences passées pour les ressources matérielles ne suffit pas. Une bonne formalisation des connaissances et savoirs des ressources mobilisables au cours du changement semble pourtant être la clé d'un changement réussi. La combinaison du contexte, des besoins de la cible, des compétences individuelles et de leur association au niveau d'une équipe sont les variables d'une équation exprimant l'efficacité d'un processus de changement. Nous ne savons pas encore écrire cette équation, mais les moyens actuels d'intelligence artificielle pourront contribuer à sélectionner les bonnes associations entre ces éléments selon des critères choisis et à proposer plusieurs chemins possibles pour garantir la performance visée par le changement. Avec cet objectif, nos travaux actuels visent à créer un outil utilisant des mécanismes d'IA et permettant à un pilote du changement de gérer dynamiquement l'affectation des ressources et de mobiliser leurs compétences associées pour atteindre la

performance souhaitée. Pour cela, nous recensons et codifions les bonnes pratiques utilisées par les pilotes expérimentés du changement pour parvenir à un équilibre optimisé des ressources. Nous développons également un guide pour qualifier les ressources et collecter les connaissances associées.

6 CONCLUSION

Cet article visait à baliser un chemin vers une meilleure performance du processus de changement avec, pour objectif prochain, le développement d'un outil logiciel aidant un pilote à gérer la transformation avec un bon niveau d'efficacité.

Pour y parvenir, nous pouvons déjà nous appuyer sur la formalisation du processus de changement détaillée dans [Siebenborn, 2005] avec une approche constituant les phases de la phase transitoire du changement et les équilibres de ressources comme cadre de représentation. Nous pouvons ensuite affiner cette formalisation en y intégrant l'affectation de ces ressources et une mobilisation des compétences associées pour en faire des variables d'efficacité du processus de changement.

Une première limite à notre démarche serait d'aboutir à un équilibre déraisonnable entre là où nous voulons aller et comment y aller. En d'autres termes, si les efforts de formalisation des contextes, des connaissances et des interactions demandent une énergie supérieure au gain possible à en attendre, cela enlèvera tout intérêt à la démarche. Notre équation générique de l'efficacité devra en tenir compte.

Une seconde limite serait de rester dans la gestion du changement en n'utilisant que des outils RH utilisés pour gérer les compétences. Notre démarche devra intégrer d'autres disciplines partageant le même objectif.

Une troisième et dernière limite sera de proposer un outil trop spécifique au changement mené. Notre démarche devra être la plus générique possible pour applicable à des types de changement très différents.

REFERENCES

- Baudot, JY. (2019). <http://www.jybaudot.fr/Management/performance.html>
- Berrah, L. (1997). Une approche d'évaluation de la performance industrielle : modèle d'indicateurs et techniques floues pour un pilotage réactif. Thèse de doctorat, Université de Savoie.
- Cleland, D. I., & King, W. R. (1983). System Analysis and projet management (Vol 3). McGraw-Hill, New York.
- Chiquet, B. M. (2013). Une nouvelle technologie managériale : l'Holocarcy. IGI Partners.
- Covès, C. (2000). Analyse et estimation des performances de processus d'entreprise. Thèse de doctorat, Université de Montpellier 2.
- Feigenbaum, A. V. (1961). Total Quality Control (2^d édition). McGraw-Hill, New York.
- Getz, I., & Carney, B. N. (2012). Liberté et Cie. Ed Fayard.
- Goldratt, E. M., & Cox, J. (1984), (1986). The Goal: excellence in manufacturing. North River Press.
- Hammer, M. (1996). Beyond Reengineering Work. Harper Collins Business, London.
- Hammer, M., & Champy, J. (1993). Reengineering the corporation: a Manifesto for business revolution. Nicholas Brealey, London.

- Harry, M. & Schroeder, R. (2000). Six Sigma: The Breakthrough Management Strategy revolutionizing the world's top corporations. Ed Top Corporations.
- Imai, M. (1972). Kaizen : la clé de la compétitivité japonaise. Ed Eyrolles, Paris.
- ISO. (2015). ISO9001 - 2015, ISO 9001 quality management systems – requirements. & 9.3, International Organization for Standardization, Geneva.
- Karlöf, B., Lunggren, K., & Edenfeldt, M. (2001). Benchlearning: good examples as lever for development. Chichester: Wiley, 2001.
- LeBoterf, G. (2002). Les connaissances : une ressource pour agir avec compétence. Actes du 1^e colloque de CIGI2002, Nantes.
- Lewin, K. (1951). Field Theory in social Science. Ed D Cartwright, Harper & Row, NewYork.
- Magnani, F. (2018). La dimension humaine du lean: le cas du groupe PSA, these en doctorat, ESCP & Panthéon-Sorbonne
- Ohno, T. (1990). L'esprit Toyota. Ed Masson, Paris.
- Penrose, E. T. (1972). The theory of growth of the firm (5^e edition). Basil Blackwelland Mott, Oxford.
- Plane, JM. (2019). Les approches contemporaines du management. Management des organisations, Dunod, 159-296.
- Rota, V.M. (2010). Gestion de projet agile. Ed Eyrolles.
- Siebenborn, T. (2005). Une approche de formalisation du processus de changement dans l'entreprise. Thèse de doctorat, Université de Savoie.
- Terssac, G. de (1996). Savoirs, compétences et travail. Savoirs théoriques et savoirs d'action, PUF, 223-247.
- Vas, A. (2002). Les processus de changement organisationnels à l'épreuve des faits : une approche multi-pragmatique. Actes de la 11^e Conférence Internationale de l'AIMS, Paris.
- Weil, T. (2003). Annale des Mines, Ermits. 30-34

Apports du marketing pour le management de projets innovants collaboratifs en lien avec une transition écologique et responsable

THÉO GALMICHE¹, ELISE MARCANDELLA²

¹ Université de Lorraine - Polytech-Nancy
F-54000 Nancy, France
theogalmiche@orange.fr

² Université de Lorraine – CEREFIGE / IUT Epinal-Hubert Curien
F-54000 Nancy, France
elise.marcandella@univ-lorraine.fr

Résumé – L’objectif de cette communication est de proposer un modèle de management responsable des projets innovants collaboratifs permettant de faire le lien entre les fonctions marketing et gestion de projet. Il s’agit d’une approche théorique visant à assembler et développer des modèles existants.

Les modèles principaux sur lesquels s’appuie cette communication sont le modèle FIRMIN (Forme d’Intégration de la Responsabilité Sociétale dans le Management de l’Innovation) et le modèle P.R.I. (Produit.Relationnel.Institutionnel) orienté marketing responsable et qualité.

Les liens entre management responsable de l’innovation et marketing sont très souvent sous-estimés alors qu’ils sont sources de contradictions, notamment lorsque les organisations impliquées dans un projet innovant collaboratif n’ont pas une vision complète du contexte de développement du projet et des attentes de ses acteurs.

Le modèle FIRMIN-PRI a pour objectif de faciliter une meilleure prise en compte des attentes des parties prenantes impliquées dans un projet innovant collaboratif et de favoriser ainsi une meilleure intégration de celui-ci dans son écosystème.

Abstract – The purpose of this paper is to propose a responsible management model for collaborative innovative projects, making it possible to link the marketing and project management functions. It is a theoretical approach aimed at assembling and developing existing models. The main models on which this communication is based are the FIRMIN model (Form of Integration of Social Responsibility in the Management of Innovation) and the P.R.I. (Product.Relational.Institutional), they were oriented toward responsible marketing and quality. The links between responsible management for innovation and marketing are very often underestimated even though they are sources of contradictions, especially when the organizations involved in a collaborative innovative project don’t have a complete vision of the project development context and the expectations of its actors. The objective of the FIRMIN-PRI model is to ease the consideration of the stakeholders’ expectations involved in an innovative collaborative project and thus promote its integration in the ecosystem.

Mots clés - Innovation, responsable, management, marketing, FIRMIN

Keywords – Innovation, responsible, management, marketing, FIRMIN

1. INTRODUCTION

L’édition 2021 du Congrès CIGI-Qualita, place « L’Humain au cœur de l’Industrie du Futur ». En lien avec cette problématique, nous proposons de présenter un modèle de management responsable de l’innovation appelé « FIRMIN » pour Forme d’Intégration de la Responsabilité sociétale dans le Management de l’INNovation [Marcandella, 2015 ; 2017] avec un focus plus particulier sur les apports du marketing à cette approche responsable de la gestion de projets innovants collaboratifs.

L’industrie du futur, encore plus que l’industrie actuelle, sera confrontée à la nécessité vitale d’intégrer les

enjeux du développement durable (DD) dans sa stratégie de développement. L’intégration de ces enjeux représente un défi pour les entreprises qui sont confrontées à la nécessité de prendre en compte un nouveau modèle de développement ne laissant pas toute la place à la création de valeur économique. Quel rôle peut alors jouer le marketing dans ce management responsable alors qu’à l’origine il s’agit d’une fonction tournée essentiellement vers le développement économique de l’entreprise ?

Le marketing classique en raison de son orientation « profit à tout prix » [Benaceur, 2016] a commencé dès les années 70 à être la « cible de critiques de la part de certains auteurs par rapport à l’incitation permanente à la consommation (inviter à consommer plus) sans prendre en

considération l'intérêt sociétal [Dawson, 1969 ; Feldman, 1971 ; Kotler et Levy, 1971], d'où la notion de Marketing Sociétal, qui facilitera l'émergence des notions de Marketing Durable ou de Marketing Socialement Responsable » [Benhallam, 2016, p.36]. Si autrefois, le marketing était tourné uniquement vers les besoins du client, aujourd'hui la conception de l'offre s'est ouverte à de nouvelles parties prenantes. Ainsi, dans le cadre du marketing responsable [Pastore-Reiss, 2012, p.9] : « la conception de l'offre est le fruit d'une collaboration avec l'ensemble des acteurs du produit : fournisseurs, société civile et collaborateurs aussi bien que distributeurs et clients ». L'auteure souligne également la nécessité pour la fonction marketing de « s'ouvrir à son écosystème, mutualiser les ressources mais aussi les coûts, écouter pour comprendre les freins et les difficultés » et le risque de se retrouver face à des injonctions contradictoires lors de l'intégration des enjeux du développement durable dans la stratégie de développement de l'entreprise.

Les apports de l'ouverture du marketing à son écosystème sont à la base du modèle « P.R.I » (Produit.Relationnel.Institutionnel) orienté marketing responsable et qualité [Barthel, 2013]. L'auteur considère le marketing comme « un ensemble de processus qui consistent à créer, développer et communiquer de la valeur aux clients » [Kotler *et al.*, 2006] au sens large (clients finaux ou autres parties prenantes (PP) exerçant une influence directe ou indirecte sur la qualité organisationnelle et la qualité des produits offerts sur le marché) en vue de les satisfaire durablement. Cette approche du marketing ouverte aux attentes des parties prenantes nous a semblé pouvoir apporter une nouvelle dimension au modèle FIRMIN [Marcandella, 2017]. Comme le souligne Desjeux [Desjeux, 2013, p107] à propos du marketing : « le problème principal vient de la contradiction entre d'un côté l'obsolescence et l'innovation continue et de l'autre côté le développement durable. La question est de savoir si le marketing est capable d'innover sur le processus même d'innovation ». Ces deux modèles ayant fait l'objet d'articles séparés dans le cadre de précédents congrès QUALITA, nous nous intéresserons plus particulièrement dans cet article à ce qui permet de les relier.

Cet article, après une brève présentation des modèles FIRMIN et P.R.I (§2) propose d'étudier les apports du marketing responsable au management responsable de l'innovation, présentés ensuite sous la forme d'un modèle FIRMIN intégrant la fonction marketing (§3) puis discutés en s'intéressant notamment aux apports et limites de cette approche pour aller vers une transition écologique et responsable (§4).

2. MANAGEMENT RESPONSABLE, INNOVATION ET MARKETING

2.1. Le management responsable

La première définition, internationalement reconnue, du concept de développement durable est issue du rapport de Brundtland présenté en 1987 par la Commission mondiale sur l'environnement et le développement de l'ONU [Brundtland, 1987] : « Le développement durable est un développement qui répond aux besoins du présent

sans compromettre la capacité des générations futures à répondre aux leurs ». Cette approche politique a été déclinée au cours du temps pour faciliter son application à différentes échelles géographiques et par différentes parties prenantes. C'est ainsi qu'en 1992 la Conférence internationale de Rio a donné le jour aux Agendas 21, déclinaison adaptée aux enjeux territoriaux du DD, et en 2010 l'organisme international de normalisation ISO a publié la norme internationale ISO 26000 – Lignes directrices relatives à la Responsabilité sociétale [AFNOR, 2010], qui s'adresse à tout type d'organisation (entreprises, organismes publics...) souhaitant intégrer les enjeux du DD dans sa stratégie et ses relations avec les parties prenantes.

De ce fait, le management responsable, compris comme un management intégrant les lignes directrices de la responsabilité sociétale (*cf.* Tableau 1), peut jouer un rôle particulièrement important dans le processus de création et de gestion de projet puisqu'il permet d'impliquer tous les acteurs dans une approche consensuelle facilitant la mise en relation des ressources et connaissances de chaque partie prenante au service du projet et du développement durable. En effet faciliter la collaboration, la mise en commun des visions différentes et les enjeux majeurs d'un projet conduit à une gestion du projet plus efficace et surtout consciente de sa globalité et de ses impacts. Ainsi le management responsable peut rendre obsolète un modèle de management classique qui ne prendrait pas en compte tous les acteurs et parties prenantes et la complexité de leurs relations.

Approche de la RS	Situation ou vision
Approche conception de « produits ou services »	S1 : « Produits Verts ou Equitables »
	S2 : « Production durable » Pensée cycle de vie
Approche Managériale	S3 : « Gestion des Performances durables »
Approche Stratégique [nouveau concept]	S4 : « Management responsabilisant » ou « Management RS » de l'innovation c'est-à-dire intégrant les principes de la responsabilité sociale des organisations selon la NF ISO 26000 et une approche dynamique de l'évaluation

Tableau 1 - Typologie des situations d'intégration des enjeux du DD dans le management de l'innovation selon FDX-273 [AFNOR, 2014]

2.2. FIRMIN - Représentation du management responsable de l'innovation

La figure 1 présente le modèle FIRMIN, Forme d'Intégration de la Responsabilité sociétale dans le Management de l'Innovation (FIRMIN) [Marcandella, 2017]. Il s'agit d'une représentation générique du management responsable de l'innovation qui met en évidence 7 caractéristiques essentielles à prendre en compte : le contexte dans lequel se déroule l'innovation (conditions cadres de l'innovation), l'écosystème social (parties prenantes des projets innovants), les quatre étapes du processus de management intégrant la responsabilité sociétale (exploration / évaluation / management des projets / capitalisation) et la nécessité d'une gouvernance particulière (cercle central incluant une communication, une évaluation et une gestion des risques partagées entre les parties prenantes). Les acteurs directement impliqués

dans la gestion du projet innovant collaboratif constituent ce que nous nommons ici la « communauté de projet ».

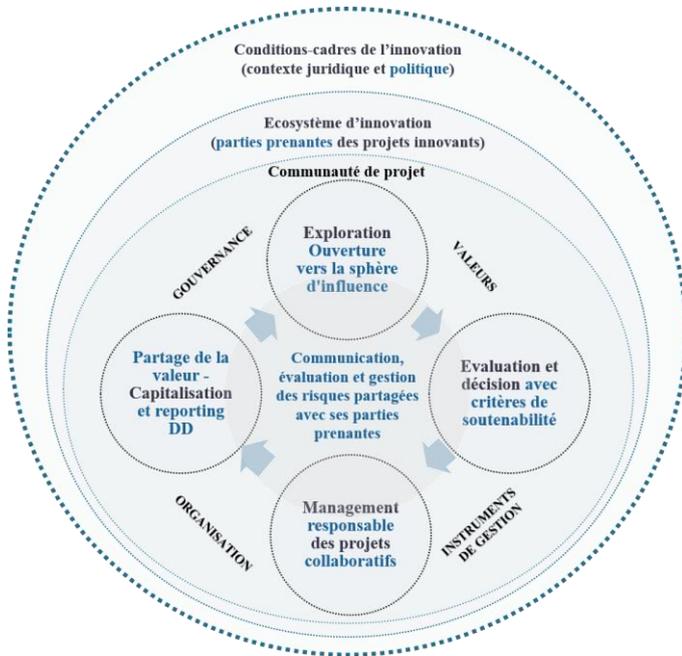


Figure 1. Forme d'Intégration de la Responsabilité sociétale dans le Management de l'INnovation, modèle FIRMIN [D'après Marcandella, 2017]

Ce modèle permet de situer les différents acteurs mais également les ressources et les compétences à mettre au service du projet ainsi que les éléments de contexte qui peuvent l'impacter. La fédération des différents éléments au sein de ce modèle permet donc une vision globale mais néanmoins précise de la responsabilité sociétale au sein de l'écosystème projet. Cette visualisation des informations nécessaires à la gestion d'un projet permet de faciliter la recherche des informations pertinentes et utiles pour le mener à bien. Le but est qu'il serve en quelque sorte de « base de données » de référence, en contenant tous les tenants et aboutissants d'un management efficace permettant de mener à bien ledit projet.

Ainsi, FIRMIN incite la communauté de projet à se questionner sur : les valeurs qui lient les membres de cette communauté, la gouvernance de la communauté et les relations avec les parties prenantes, le mode d'organisation et les instruments de gestion à mettre en œuvre pour gérer le(s) projet(s) innovant(s) collaboratif(s).

Cette représentation nous est également utile en tant que chercheurs pour identifier des travaux, notamment scientifiques, qui pourraient venir nourrir notre réflexion. Ainsi, en tant que chercheurs en sciences de gestion, impliqués dans une recherche-action dans un pôle de compétitivité visant à consolider cette représentation [Marcandella, 2017], nous nous sommes interrogés sur les outils, méthodes, méthodologies, concepts qui pourraient nous aider à jouer le rôle de facilitateur consistant à animer le cercle central « Communication, évaluation et gestion des risques partagées avec ses parties prenantes ». Ce questionnement nous a remis en mémoire les travaux en lien avec le marketing responsable que Patrick Barthel avait présentés au Congrès QUALITA en 2013 et 2015

[Barthel, 2013 ; 2015] et donné l'idée de proposer un agencement entre son modèle P.R.I et FIRMIN [Marcandella, 2015 ; 2017]. Dans sa version initiale, FIRMIN s'appuyant sur le tableau des typologies d'intégration des enjeux du développement durable dans le management de l'innovation proposé par l'AFNOR (cf. Tableau 1), restreignait le rôle du marketing à l'approche « conception de produits ou service » par sa capacité à identifier les attentes des consommateurs et à l'approche « managériale » par sa capacité à communiquer en externe pour valoriser l'image de l'entreprise par le recours à des certifications et labels divers. De ce fait, son rôle ne semblait pas central pour la gestion responsable des projets innovants collaboratifs. Les travaux sur le marketing responsable ont permis de réinterroger la place du marketing dans ce processus.

2.3. P.R.I. – un modèle de construction d'une stratégie marketing responsable de la qualité

Patrick Barthel [Barthel, 2013] à l'origine du modèle P.R.I, faisant ainsi écho à nos interrogations sur la place du marketing dans le processus de management de l'innovation, souligne que « dans le contexte actuel, le marketing a conscience de la nécessité de dépasser les seules exigences inhérentes à la qualité technique (respect de normes techniques), de la qualité appréhendée à travers les labels « Produit » ou encore de la qualité organisationnelle vue au travers des seules normes ISO ». Dans cette perspective, il suggère que le professionnel du marketing devrait privilégier la satisfaction du client sur toute la durée de vie du produit acheté, mieux prendre en compte les intérêts de la collectivité, et enfin se focaliser essentiellement sur trois formes de qualité, correspondant à trois écosystèmes clefs de toute organisation quelle qu'elle soit, à savoir l'écosystème « Produit », l'écosystème « Relationnel » et enfin l'écosystème « Institutionnel ».

Il est nécessaire de préciser ici que les notions de projet et de produit sont conjointes, puisque la mise au point d'un produit passe par un ensemble de processus mettant en commun un certain nombre d'acteurs et de parties prenantes. De ce fait la création et la mise sur le marché d'un produit passe obligatoirement par la gestion efficace d'un projet. Le projet considéré doit être mis en contexte avant toute autre action. Un projet ne se limite pas à ses acteurs ou son but, sa portée est bien plus ample. Pour comprendre ce contexte et en estimer la portée, il est possible de se référer aux trois leviers respectifs d'action marketing « qualité de l'écosystème institutionnel » (cf. tableau 2), « qualité de l'écosystème produit » ou « écosystème projet » (cf. tableau 3), et « qualité de l'écosystème relationnel » (cf. tableau 4), proposés par Patrick Barthel [Barthel, 2013].

Il est à noter que dans ces tableaux la responsabilité sociétale et la valeur ajoutée en termes d'intégration d'attributs DD dans l'écosystème considéré est le troisième et dernier levier, ce qui semble indiquer un certain détachement de la volonté marketing à inclure une dimension responsable dans le modèle « classique ».

Tableau 2. Tableau des trois leviers d'action marketing sur l'écosystème institutionnel – Modèle P.R.I [Barthel, 2013]

ÉCOSYSTÈME INSTITUTIONNEL ET MODÈLE P.R.I		
Valeur ajoutée en termes économiques et d'utilités fonctionnelles (Innovations et économies organisationnelles)	Valeur ajoutée en termes de maîtrise des risques (Innovations et maîtrise des risques organisationnels)	Valeur ajoutée en termes d'intégration d'attributs DD dans l'écosystème « institutionnel » (Innovations relationnelles et respect des principes DD)
Orientation Qualité Totale du système de production-distribution de l'entreprise en vue d'un meilleur rapport qualité-prix-service à offrir aux clients	Gestion des risques organisationnels et évitement des externalités négatives liées aux activités de l'organisation (risques économiques, environnementaux, sanitaires et sociétaux)	Intégration des valeurs environnementales, sanitaires et sociétales (opérations de mécénat philanthropique...) dans les activités et le fonctionnement de l'organisation visée pour un meilleur service « clients » et une valorisation de l'image institutionnelle

Tableau 3. Tableau des trois leviers d'action marketing sur l'écosystème produit - Modèle P.R.I [Barthel, 2013]

ÉCOSYSTÈME PRODUIT ET MODÈLE P.R.I		
Valeur ajoutée en termes économiques et d'utilités fonctionnelles (Innovations et profitabilité)	Valeur ajoutée en termes de maîtrise des risques (Innovations et maîtrise des risques « Produit »)	Valeur ajoutée en termes d'intégration d'attributs DD dans l'écosystème « Produit » (Innovation et intégration d'attributs DD dans les produits commercialisés)
Identification et recherche d'une plus grande efficacité/rentabilité du portefeuille « Produits » Rapport qualité-prix-services satisfaisant à la fois pour le client et pour l'entreprise	Identification, gestion et prévention des risques/dangers liés à l'offre commerciale (Risques économiques, sanitaires, environnementaux ou sociétaux)	Intégration d'attributs DD dans les produits commercialisés (produits écoconçus, réduction des déchets, économie d'énergie, produit-partage...)

Tableau 4. Tableau des trois leviers d'action marketing sur l'écosystème relationnel - Modèle P.R.I [Barthel, 2013]

ÉCOSYSTÈME RELATIONNEL ET MODÈLE P.R.I		
Valeur ajoutée en termes économiques (Innovations et profitabilité)	Valeur ajoutée en termes de maîtrise des risques (Innovations et maîtrise des risques « Relationnels »)	Valeur ajoutée en termes d'intégration d'attributs DD dans l'écosystème « Relationnel » (Proximité Clients/PP et Innovations relationnelles)
Gestion du portefeuille « Clients et Parties Prenantes » efficace et rentable mais dans le respect de tous	Gestion des risques relationnels clients et respect de la souveraineté du consommateur et des PP impliquées dans l'élaboration des produits/services commercialisés (Services consommateurs, gestion pertinente des réclamations...)	Coopération et construction de relations de type gagnant-gagnant avec les clients et l'ensemble des PP impliquées dans l'élaboration des produits/services vendus Marketing de filière et échanges mutuels de connaissances ou de compétences

Les tableaux récapitulatifs 2, 3 et 4 permettent de situer le marketing responsable dans les trois types d'écosystèmes : « institutionnel », « produit », et « relationnel » en lien avec la démarche Qualité de l'organisation. L'auteur propose de gérer les trois leviers respectifs des différents écosystèmes et d'utiliser ce modèle P.R.I. dans une démarche en lien avec le management de la qualité.

Ce modèle montre que le marketing responsable joue un rôle important dans la prise de décision initiale d'un projet, puisque les enjeux économiques et utilités fonctionnelles sont en première ligne. Il est aussi possible de conclure quant à l'utilité du marketing responsable dans le management responsable, tant au niveau de l'innovation que des attributs DD, avec par exemple la proposition d'aller vers l'éco-conception de produits (cf. tableau 3) ou une approche par filière (cf. tableau 4).

3. APPORTS DU MARKETING AU MANAGEMENT RESPONSABLE DE L'INNOVATION

3.1. Leviers marketing pour le développement de projets collaboratifs

Les interactions entre la fonction marketing et les autres acteurs d'un projet est très importante puisqu'il s'agit ici de représenter, de façon théorique, ces échanges. Percevoir le marketing seulement comme un instrument de vente ou une sphère globale ayant pour simple but de borner le projet ou produit considéré serait réducteur et ne pourrait en aucun cas permettre d'y intégrer une vision responsable.

Le marketing perçu comme un ensemble, comprenant plusieurs leviers et plusieurs moyens d'actions semble en revanche adapté à notre propos [Van Laethem *et al.*, 2017].

Le marketing deviendra alors un appui, une entité fédératrice à part entière, en intégrant chaque partie prenante (de l'entreprise ou acteur d'un projet) dans sa démarche [Diviné, 2015], de façon à équilibrer les attentes de l'écosystème pris dans sa globalité.

La vision du marketing « classique », par opposition à un marketing responsable, repose sur le cycle de vie économique des produits mettant en relation le volume de vente des produits en fonction des différentes phases de ce cycle (développement, croissance, maturité, déclin) [Van Laethem *et al.*, 2017] (cf. Figure 2).

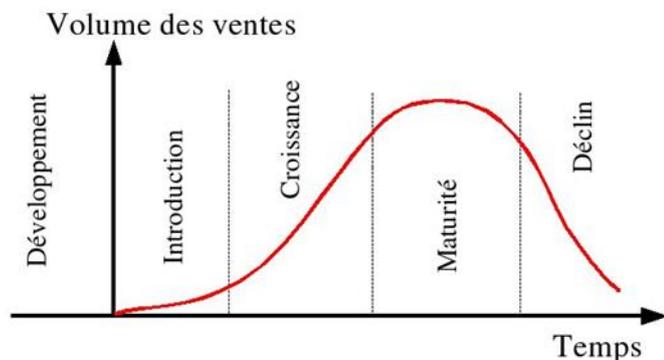


Figure 2. Cycle de vie classique d'un produit, sans démarche responsable - [Van Laethem *et al.*, 2017]

Ce cycle va à l'encontre d'une vision responsable et durable puisqu'il se base uniquement sur un début et une fin de vie d'un produit donné alors que l'approche par la responsabilité sociétale prendra en compte les différentes phases du cycle de vie environnemental qui suggère de créer de la valeur depuis l'extraction des matières premières jusqu'à la valorisation des déchets considérés alors comme matière première secondaire ou source de valorisation énergétique. La valeur n'est donc plus liée uniquement à la vente du produit mais son estimation nécessite de prendre en compte des externalités liées aux différentes phases du cycle de vie environnementale, et sociale, du produit. Seulement cette intégration très souvent n'est faite que dans le but de pousser à la consommation en appuyant les qualités durables ou responsables d'un produit. En effet l'intégration d'une boucle rétroactive dans le cycle de vie d'un produit n'est jamais sans frais et souvent l'effort d'investir du capital directement dans le retour d'expérience pour améliorer la pérennité des processus de conception responsable n'est pas la volonté première de l'industrie actuelle. En effet l'intégration de processus plus responsables pourrait diminuer le rendement premier d'une chaîne de production et ceci va à l'encontre du bénéfice direct.

Ainsi, le marketing responsable se doit d'intégrer des phases complémentaires comme le recyclage ou l'éco-conception et plus globalement les principes de l'économie circulaire, afin de participer à cette démarche durable. La représentation classique du cycle de vie économique définissant la mise sur le marché, la vente et le déclin des produits doit être modifiée pour prendre en compte cette vision holistique de la valeur d'un produit, nécessitant un reengineering du marketing classique [Lambin, 2012]. Comme suggéré par l'ADEME, il s'agit de transformer le

marketing pour passer d'un modèle visant à créer, à tout prix, « de la valeur pour une marque » à un modèle orienté vers la création de valeur(s) « pour la marque et l'ensemble des parties prenantes » [ADEME, 2013]. Cette approche visant l'intégration des enjeux du DD dans les écosystèmes « produits » et « relationnel » peut être à l'origine de l'émergence de projets innovants collaboratifs qui pourront être managés responsablement avec l'aide notamment d'une fonction marketing soucieuse de ces enjeux.

3.2. Utilisation de leviers marketing pour aller vers un management responsable

Il est important de questionner le sujet de la transition écologique qui opère dans la société depuis plusieurs années désormais, allant dans le sens d'une économie plus responsable et intégrant de ce fait le marketing responsable.

La vocation du marketing demeure intacte même suivant une approche responsable et durable, cependant il est intéressant de séparer quelques aspects du marketing pour les détailler et les adapter au mieux dans le modèle théorique. Tout d'abord si l'on suit la définition énoncée dans le livre « Marketing : remède ou poison ? » [Chauzal-Larguier 2013] : « Le marketing [classique] a des fins mercantiles et repose sur la persuasion et la manipulation ». Or cette définition ne concerne essentiellement qu'une partie du marketing, à savoir la communication. La finalité économique énoncée ici apparaît unique et va donc à l'encontre d'une approche par la responsabilité sociétale.

Il existe cependant une branche de la communication qui utilise l'approche « responsabilité sociétale » pour vendre, et donc servir le marketing. Il n'est pas extraordinaire qu'une entreprise, par exemple, vende ses processus et méthodes sous l'angle du respect de l'environnement et s'engage dans une démarche de développement durable pour améliorer son image de marque et ainsi rejoindre cette finalité « purement mercantile ». De ce fait « l'approche responsabilité sociétale » est un outil marketing et la communication est la clé de la transition d'un marketing orienté produit et profit vers un marketing responsable et qui utilise même toute son ampleur pour s'améliorer.

Ce levier marketing qui consiste donc à inclure dans son cycle de vie des aspects environnementaux et sociaux permet d'avoir une approche différente du modèle FIRMIN [Marcandella, 2015] puisqu'il existe désormais une rétroaction du marketing vers l'exploration, qui sera la première phase d'un projet ([Exploration/Ouverture vers la sphère d'influence] dans la Figure 1). Ainsi les processus de conception/fabrication sont eux-mêmes revus selon des critères socio-éco-responsables. De même l'exploration peut aussi influencer le marketing puisque, comme énoncé précédemment, les processus de fabrication, la politique d'entreprise ou même l'approche managériale d'un projet ou d'un produit peuvent être des arguments marketing à part entière.

La communication est donc un biais important dans le marketing - ce qui n'est pas une nouveauté - en revanche

lorsqu'il est question du marketing responsable, cette communication peut être l'élément déclencheur d'une transition de technique marketing puisqu'il révèle au client le processus même de fabrication. C'est un partage d'informations qui intègre directement le client puisque ce dernier est non seulement une partie prenante du marketing, mais aussi un acteur direct notamment par sa façon de consommer influencée autant par le produit finalisé que le processus de fabrication et de conception de ce produit. Ces interactions peuvent s'étendre à un cercle de parties prenantes allant au-delà du client final, c'est-à-dire la communauté de projet.

3.3. Écosystèmes P.R.I., marketing responsable et management responsable

Le modèle développé par Patrick Barthel [Barthel, 2013] regroupant les trois écosystèmes produit - relationnel - institutionnel a permis de déduire la place et l'ordre théorique des leviers marketing mais sans prendre en compte un lien possible entre eux. La partie précédente sur l'utilisation des leviers marketing dans l'approche responsable montre une certaine corrélation entre les trois leviers, valeur ajoutée en termes économiques, valeur ajoutée en termes de maîtrise des risques, valeur ajoutée en termes de développement durable. Par exemple la valeur ajoutée en termes de développement durable peut être un argument utilisable pour le levier de valeur ajoutée en termes économiques. De même, le levier de valeur ajoutée en termes de maîtrise des risques peut être influencé par le levier de valeur ajoutée en termes de développement durable puisque les processus même de fabrication et de conception suivent eux-mêmes une démarche responsable et durable. La partie de maîtrise des risques est encore améliorée d'un point de vue responsable car elle prend désormais en compte tous les facteurs autour du produit et les impacts sur l'ensemble des parties prenantes.

Il serait complexe d'introduire dans ce modèle cet aspect d'interactions directes et indirectes entre les trois leviers marketing de chaque système. Il serait donc préférable de sous-entendre ces interactions dans un modèle plus global, qui prend en compte la totalité des acteurs d'un projet puisque le marketing est un des acteurs parmi d'autres. Pour cela il est nécessaire de définir les liens entre marketing responsable et management responsable dans un premier temps, puis d'intégrer ces liens dans un modèle complet

3.4. Marketing responsable - Place et évolution de FIRMIN

Comme développé précédemment, les écosystèmes produit, relationnel et institutionnel peuvent être modifiés pour intégrer dans les valeurs ajoutées en termes économiques et maîtrise des risques une dimension durable et responsable qui sera alors un levier marketing. Le modèle FIRMIN (cf. Figure 1), permet l'intégration de la responsabilité sociétale dans le management, et de ce fait ce modèle englobe les points de gestion d'un projet innovant collaboratif et par extension le développement d'un produit. Nous pouvons donc introduire cette notion de marketing responsable, avec les leviers adaptés, pour obtenir un modèle plus complet, théorique, mais applicable à un projet concret dans le but de le perfectionner. La figure 3, modèle FIRMIN intégrant le marketing responsable,

permet la visualisation de la sphère marketing comme acteur, partie prenante et entité d'interaction avec les pôles d'un projet innovant.

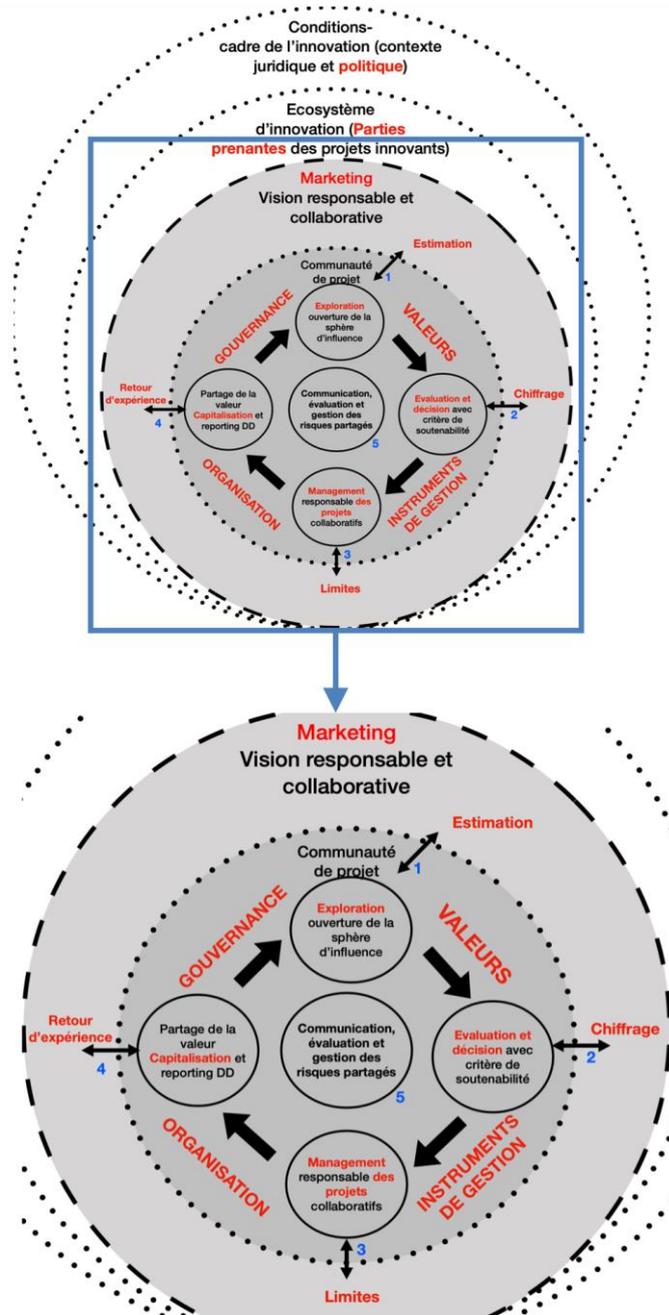


Figure 3. FIRMIN-PRI - Forme d'Intégration de la Responsabilité sociétale dans le Management de l'INnovation, en intégrant le marketing responsable et ses interactions [Auteurs].

Au cours de cette communication le marketing a été présenté comme une nécessité dans la gestion d'un projet. Il est effectivement essentiel lorsqu'il s'agit de projets type conception, fabrication et vente d'un produit ou service. Il serait dans ce cas envisageable de rajouter un pôle « marketing » (6^{ème} cercle au centre de la figure), puisque chacun des pôles interagit avec lui, de l'exploration du projet à la capitalisation. Mais intégrer le marketing de cette façon implique une présence systématique et surtout centrale de celui-ci. Or le management responsable développé dans la littérature ou plus particulièrement dans FIRMIN [Marcandella, 2015 ; 2017] ne concerne pas

obligatoirement un projet de ce type. L'intégration du marketing responsable doit donc se faire comme étant « optionnelle », pour rendre le modèle adaptable à des projets ne possédant pas de volet marketing pré-défini.

La sphère d'écosystème d'innovation des parties prenantes contient le marketing, puisque partie prenante et acteur de l'innovation. Il ne faut pas oublier le versant actionnable de FIRMIN et de ce fait ne pas placer le marketing comme incontournable et obligatoire pour chaque projet. Donc la création d'une sous-sphère *Marketing, vision responsable et collaborative*, à l'intérieur de l'écosystème d'innovation (parties prenantes des projets innovants) semble être la meilleure solution puisque cette sous-sphère est optionnelle, les limites qu'elle impose sont intrinsèquement liées au projet lui-même et si celui-ci ne comporte pas de fonction marketing, ces limites sont étendues à la sphère supérieure. Il aurait été envisageable de créer une deuxième sphère interne à celle de l'écosystème d'innovation, mettant le marketing directement au sein de la sphère « communauté de projet ». Cependant un tel modèle sous-entend une utilisation de la sphère marketing dans tout type de projet, transformant le modèle FIRMIN en un modèle applicable uniquement dans le cas de projets faisant intervenir un acteur marketing et ce n'est pas le cas de tous. Le tableau 4 définit chaque interaction de façon commune, pour n'exclure aucun projet de ce modèle.

Tableau 4. Les interactions entre marketing responsable et les autres pôles du modèle FIRMIN.

N°	Caractérisation des interactions
1	Le projet est à l'état de réflexion, la question marketing/financière n'est pas nécessaire mais toutefois à prendre en compte et à imaginer de façon qualitative De plus l'étude de marché apporte une valeur ajoutée dans les données du projet
2	Le projet devient plus concret, les acteurs de conception ont pu intégrer leur point de vue et la vision quantitative est nécessaire. L'étude de marché et les simulations marketing s'intègrent à cette vision
3	La base du chiffrage a permis une vision quantitative complète, celle-ci définit les bornes opérationnelles. La politique d'entreprise est respectée, le business plan fourni par le marketing a été adopté
4	Le retour d'expérience et le partage de la capitalisation permettent un regard critique sur les choix des parties prenantes et sont effectués sous forme pédagogique pour une mise en commun des différents aspects du projet avec du recul
5	La communication orientée vers les différentes parties prenantes aide chacune d'entre elles à mieux expliciter ses enjeux, à évaluer les impacts des décisions et à coopérer au service d'un projet innovant collaboratif prenant en compte les enjeux du développement durable.

Source : Auteurs

Le marketing est donc désormais au sein d'un modèle de management responsable, ses bornes dépendent bien entendu du projet et l'application de ce modèle pour la première fois mènera à un regard critique et fonctionnel. L'application de ce modèle à l'analyse de cas concrets devraient nous permettre d'identifier les projets innovants

collaboratifs qui gagneraient à être accompagnés en prenant en compte cette approche marketing responsable. Est aussi à prendre en compte le recul sur ce modèle, puisqu'évidemment il n'est que théorique et donc un retour d'expérience permettra sa validation opérationnelle et/ou son adaptation. Ce principe du retour d'expérience est très intéressant d'un point de vue collaboratif puisque ce seront des informations, des compétences et des connaissances qui seront partagées.

4. DISCUSSION ET CONCLUSION

4.1. Accompagner la transition sociétale et écologique

Il est de la volonté globale d'orienter la transition vers un mode de consommation plus responsable, plus transparent au niveau des processus de fabrication. Ainsi l'intégration du marketing et du management responsable se fera également via le biais de la consommation. Il s'agit ici d'influence des différentes sphères, car même si on intègre le marketing avec une vision responsable et collaborative, cette sphère est elle-même intégrée au sein de celle des parties prenantes. Ce qu'il faut comprendre ici, c'est que puisque la consommation change - c'est à dire la demande - l'offre doit s'adapter en conséquence pour pouvoir continuer à créer de la valeur. Ainsi au vu de la transition actuelle vers une écologie assumée et démontrée attendue par les consommateurs, l'offre n'a de choix pour survivre que de rendre l'écosystème produit plus durable.

En ce qui concerne la transition écologique et responsable « industrielle » en cours, reconsidérant la place de l'humain et de la nature au sein de l'industrie, le marketing a un rôle important à jouer par ses compétences en communication. Ces compétences peuvent être déployées au service de la stratégie « responsabilité sociétale » des industries en accompagnant les décideurs par une sensibilisation sur les impacts de leurs activités et les méthodes et outils permettant de les maîtriser (cf. Modèle P.R.I.).

Dans le modèle FIRMIN développé ci-dessus, la sphère globale juridique et politique peut elle-même être orientée suivant une politique plus écologique et responsable ou non. C'est le cas de beaucoup de pays durant la dernière décennie et cela influe évidemment sur le fonctionnement des industries qui doivent se plier à de nouvelles normes et lois et faire évoluer leurs processus en fonction des directives politique en vigueur.

Les leviers pour aller vers une transition sociétale et écologique sont donc de différentes natures : politico-juridiques, institutionnels, orientés produits mais également relationnels et communicationnels. En élargissant le cercle des parties prenantes à la sphère politique, le marketing responsable peut devenir un canal de sensibilisation de cette sphère politique aux enjeux du développement durable. En communiquant de façon professionnelle et éthique sur les impacts, sur les enjeux du développement durable, d'un management responsable des projets innovants collaboratifs, la fonction marketing responsable peut orienter les politiques vers des choix et des décisions plus responsables.

4.2. Critique du modèle, alternatives

Ce modèle final, FIRMIN modifié pour intégrer le marketing responsable et ses interactions, est le fruit de recherches et d'adaptations de modèles de marketing responsable et durable à la gestion d'un projet. Or la littérature actuelle sur le sujet est relativement pauvre. Le modèle de Patrick Barthel [Barthel, 2013] est un des rares modèles de marketing responsable précisant les leviers directs d'action marketing sur les écosystèmes environnants. Cette vision est en accord sur le principe avec le management responsable de l'innovation, car il prend en compte les interactions entre les différents acteurs et parties prenantes du projet.

De même, le modèle de management responsable de l'innovation proposé par Elise Marcandella [Marcandella, 2015 ; 2017] est une façon de penser le management responsable, qui correspond à une vision holistique de la responsabilité. Ce modèle de management, FIRMIN seul, a fait ses preuves opérationnelles dans le cadre de projets de recherche menés en collaboration avec un pôle de compétitivité [Marcandella, 2015] et un pôle territorial de coopération économique [Marcandella, 2017]. Ces projets nous ont permis de constater que les chefs de projets qui accompagnent les projets collaboratifs innovants au sein de ces structures ont besoin de compétences variées qui relèvent aussi bien de connaissances technologiques que managériales. Nos réflexions sur les apports du marketing au management responsable de l'innovation nous incitent à aller vers l'intégration d'un module marketing responsable dans la formation de ces futurs chefs de projets. Ces travaux sur FIRMIN, ont également amenés un des auteurs à concevoir un cours à destination d'élèves-ingénieurs et d'étudiants du DUT QLIO (Qualité Logistique Industrielle Organisation) intitulée « Ma petite entreprise responsable » [Marcandella, 2015b] auquel nous allons ajouter un module « marketing responsable ».

5. REMERCIEMENTS

Nous tenons à remercier ici Géraldine THEVENOT, enseignante-chercheuse en marketing, pour ses éclairages précieux concernant le marketing responsable.

6. RÉFÉRENCES

- ADEME (2013), Étude exploratoire sur le marketing responsable, Septembre 2013, p.25
- AFNOR (2010). ISO 26000, *Lignes directrices relatives à la responsabilité sociétale*, AFNOR Éditions, 2010.
- AFNOR (2014). FD X50-273, Management de l'innovation - Intégration du développement durable dans le management de l'innovation, AFNOR Éditions, Juin 2014.
- Barthel, P. (2013). Le marketing responsable, un outil de management de la qualité et de promotion de l'innovation responsable. *10^{ème} édition du Congrès international QUALITA, 19-22 mars 2013, Compiègne*.
- Benaceur, F. Z. (2016). Le marketing responsable : une contribution du marketing dans le développement durable. *Recherches économiques et managériales, 19*, 7902-1112.
- Benhallam, M. (2016). La résistance socialement responsable à la consommation : proposition de conceptualisation et implications pour les pratiques du marketing (Doctoral dissertation – Université Paris-Est Créteil).
- Brundtland, G. H. (1987). *Notre Avenir à Tous*, rapport de la commission mondiale sur l'Environnement et le Développement, Les Éditions du Fleuve, Paris (traduction française de Our Common Future).
- Chauzal-Larguier, C. (2013). Chapitre 12. Le marketing responsable : entre utopie et réalité. Dans : Patrick Bourgne éd., *Marketing : remède ou poison : Les effets du marketing dans une société en crise* (pp. 285-301). Caen, France : EMS Éditions.
- Corneloup, J. (2013). Chapitre 10. La transition culturelle, alternative au marketing. Dans : Patrick Bourgne éd., *Marketing : remède ou poison : Les effets du marketing dans une société en crise* (pp. 249-264). Caen, France : EMS Éditions.
- Dawson, L. M. (1969). The human concept: New philosophy for business: Marketing concept outmoded today. *Business Horizons, 12*(6), 29-38.
- Desjeux, D. (2013). Chapitre 3. Le marketing entre cadrage, consommateur acteur et nouvelles émergences sociétales. Dans : Patrick Bourgne éd., *Marketing : remède ou poison : Les effets du marketing dans une société en crise* (pp. 93-115). Caen, France : EMS Éditions
- Diviné, M., & Stal-Le Cardinal, J. (2015). Management des communautés externes et des équipes internes : vers des frontières flexibles ? *Décisions Marketing, 85-99*.
- Feldman, L. P. (1971). Societal adaptation: a new challenge for marketing. *Journal of Marketing, 35*(3), 54-60.
- Kotler, P., Keller, K.L., Dubois, B. & Manceau, D. (2006). *Marketing management*, Pearson Éducation France, 12^{ème} édition, 2006.
- Kotler, P., & Levy, S. (1971). Demarketing—Yes, Demarketing. *Harvard Business Review. On Management, London, Heinemann*.
- Lambin, J. J., & De Moerloose, C. (2012). *Le marketing Stratégique et Opérationnel. Du marketing à l'orientation-marché*. Dunod.
- Marcandella, E. (2015). Management responsable de l'innovation : Concept, méthodologie, perspectives. *11^{ème} édition du Congrès international QUALITA, 17-19 mars 2015, Nancy*.
- Marcandella, E. (2015b). Enseigner l'Innovation et le DD en IUT – Retour d'expérience sur « Ma petite entreprise responsable. *Congrès National de la Recherche en IUT sur « L'innovation dans les IUT », 12-13 mars 2015, Epinal*.
- Marcandella, E. (2017). Faciliter le management responsable de l'innovation dans un Pôle Territorial de Coopération Économique : apports de Firmin et de la recherche partenariale. *12^{ème} édition du Congrès international QUALITA, 31 août - 1^{er} septembre 2017, Bourges*.
- Pastore-Reiss, E. (2012). *Les 7 Clés du marketing durable*. Editions Eyrolles, Paris.
- Van Laethem, N., & Durand-Mégret, B. (2017). *La boîte à outils du responsable marketing omnicanal*. Dunod.

CIGI 2021

Use of Competency Management Methods and Tools in Project-based Organizations

KAOUTHAR MLAOUHI¹, LILIA GZARA², CELINE CHOLEZ³

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France
Kaouthar.Mlaouhi@grenoble-inp.fr

² Univ Lyon, INSA Lyon, UCBL, Univ Lumière Lyon 2, DISP, EA4570,69621 Villeurbanne, France
Lilia.Gzara@insa-lyon.fr

³ Univ. Grenoble Alpes, CNRS, Grenoble INP, Science Po Grenoble, PACTE, 38000 Grenoble, France
Celine.Cholez@grenoble-inp.fr

Résumé – La gestion des compétences relève traditionnellement de la gestion de ressources humaines (RH) notamment dans une perspective d'évaluation et de gestion de carrière. Avec la compression des délais de mise sur le marché de nouveaux produits, les besoins accrus d'agilité face aux aléas et l'augmentation du turnover, la gestion des compétences devient un élément fondamental dans la gestion de projets. Elle permet d'abord de gérer les compétences individuelles dans le but d'affecter les acteurs les plus adéquats aux missions d'un projet, ainsi que la prise en compte des compétences collectives afin d'améliorer les performances. Par ailleurs, elle garantit l'évolution du capital humain ainsi que le transfert de compétences en cas de turnover. L'objectif de ce travail est d'étudier l'usage des méthodes et outils de gestion de compétences dans le pilotage de projets. Outre l'examen de la littérature, des entretiens semi-directifs ont été menés avec un nombre de praticiens travaillant dans des entreprises organisées par projets. Ces entretiens ont été analysés à l'aide d'une méthode qualitative afin de comprendre comment les entreprises gèrent leurs employés dans les projets. Les résultats de notre étude ont montré que la gestion des compétences reste en grande partie la responsabilité de la Direction RH et que les dispositifs utilisés à cet effet ne sont pas mobilisés dans les projets.

Mots clés - Gestion des compétences, Organisations par projets, Etude exploratoire, Entretiens semi-directifs.

Abstract – Competency management has traditionally been the responsibility of human resources (HR) management particularly to evaluate employees and manage their careers. With the compression of time to market for new products, the increased need for agility in the face of hazards and the increase in turnover, competency management is becoming a fundamental element in project management. On the one hand, it enables managing individual competencies with the aim of assigning the most appropriate people to a project's missions, along with the consideration of collective competencies in order to improve performance. On the other hand, it guarantees the evolution of human capital as well as the transfer of skills in case of turnover. The objective of this work is to study the use of competency management methods and tools in project management. In addition to reviewing the literature, semi-directive interviews were conducted with a number of practitioners working in companies organized by projects. These interviews were analysed using a qualitative method in order to understand how companies manage employees' competencies in projects. The results of this study showed that competency management remains largely the responsibility of the HR department and that tools used for this purpose are not mobilized in projects.

Keywords – Competencies management, Project-based organizations, Exploratory study, Semi-structured interviews.

1 INTRODUCTION

About 40% of the global economy is project-based, using management by projects as main process for producing both products and services [Turner et al., 2010]. Numerous benefits have been associated with the adoption of a Project-Based Organization (PBOs). They refer to better processes, control and lead-time reduction, higher output quality, increasing ability to respond quickly and flexibly to customers' needs, more flexible application and integration of organizational skills, learning within the project boundary, coping with risks and uncertainties and increasing ability to innovate in collaboration with clients and suppliers [Bonjour et al., 2002; Di Vincenzo & Mascia, 2012]. PBOs are especially needed in significantly customized industries, such as filmmaking and media, complex products and systems, software development, construction, engineering design, and biotechnology. Many authors confirm that PBOs are the prevalent mode in a variety of industries [Di Vincenzo & Mascia, 2012; Loufrani-Fedida, 2006].

One of the most challenging issue that such a work environment could arise is the management of human resources [Boucher et al., 2007]; in particular the management of human operational competencies within projects. Indeed, projects are considered as a favorable environment to highlight the important role that individual as well as collective competencies play in the

progress of companies. So, *how can enterprises effectively allocate people to do their tasks within a project and how can they advance their human capital for the success of projects?*

At the beginning of a project, there is no overall perception of what will be needed, what risks are likely to be encountered, or what technical skills are necessary for the project's success, etc. As the project progresses, the more problems are identified and the more insight is gained to identify the required skills [Boucher et al., 2007]. However, it is difficult to quickly find the appropriate competencies and integrate them effectively into the project [Midler, 1995].

In order to overcome this weakness, firms have to update their professional practices and search for new methods to manage the human capital along with competencies.

Before doing that, a preliminary study, subject of this paper, has to be done with the purpose of exploring methods and tools suggested in literature or currently used in industries to manage different human competencies. Based on the two theoretical and practical studies, the definition of a new competency management method will be pragmatic and well argued.

This paper consists of 3 sections. In section 2, a theoretical exploratory study is conducted and cover three main concepts: competency management, means used to manage human capital and articulation between competencies and project management

(PM). Section 3 deals with the empirical exploratory study based on the interviews conducted with different representatives of companies from different sectors. Finally, in section 4, a conclusion summarizes the most important results and the limits of this work. Future related possible work is also discussed.

2 THEORETICAL STUDY

2.1 Competencies Definition and Classification

Going through papers dealing with the topic of “competency”, a good number of different definitions and taxonomies was found. The aim of this paper is not to exhibit these definitions; rather the focus is on tools used to manage competencies. These definitions, although expressed in numerous ways, converge in their meaning. Thus, a holistic definition is proposed here for the term “competency”. It refers to the *ability of an individual, a work group or an enterprise to mobilize and combine material and immaterial resources in order to implement a given activity or a process of actions in a certain situation with the achievement of a result* [Durand, 2000; Le Boterf, 1994].

In some papers, the term “competency” is confused with the word “skill”. However, skill could be defined as the facility or the dexterity that is acquired or developed through training or experience, while as it was defined above a competency is more than just skills. It involves the ability to meet complex demands by drawing on and mobilizing environmental end psycho-social resources (including skills and attitudes) in a particular context. Competencies, therefore, may include skills, but their meaning is larger than the skills’.

Based on literature, competencies could be classified according to two separate aspects: the competency’s owner point of view and their nature or content.

From the owner’s point of view, two large clusters of competencies can be distinguished through the managerial literature: individual competencies and collective competencies. A wealth of definitions exists in the literature and among practitioners to present the concept of individual competencies. The triptych of “know, know-how and know-being” is one of the simplest and most widespread definitions [Gilbert, 1990]. The “know” corresponds to the structured set of information assimilated and integrated into a frame of reference that allows the individual to conduct activities and operate in a specific context, mobilizing for this purpose, different interpretations. Whilst, the “know-how” refers to the ability to act in a concrete way according to a predefined process or objectives. For “know-being” or attitudes, they are neglected in the resource-based perspective, as in the still emerging theory of competency. Yet the question of “know-being”, and even more of identity and will, is an essential part of an individual’s or an organization’s ability to accomplish anything [Durand, 2000]. However, as in [Le Boterf, 1994], it is not enough to have a list of knowledge, know-how and skills to be recognized as competent, it is rather the way in which an individual puts all these elements together with environmental resources to face a new or unexpected situation at the workplace.

Although the topic of individual competencies is the subject of a large literature, collective competencies are a surprisingly little explored field compared to the former. These appear essentially within the framework of project teams. The context of projects unquestionably represents the most widespread and oldest example to express collective competencies. These competencies do not consist solely of individual competencies’ aggregation, they are rather the manifestation and the interaction of these individual competencies within a group of people who share a common objective and work in a cooperative and

collaborative environment. Thus, the literature provides a definition of the collective competency that appears at the intersection between the individual and organizational level in terms of team line-up. However, it remains difficult to measure and mobilize. Only its descriptive character, rather than perspective gives it a usable aspect [Nordhuag, 1996].

From a content point of view, the nature of the competency depends largely on the context in which it is used and, above all, on the enterprise field of activity. There are papers that distinguish technical competencies from personal competencies, interpersonal competencies and methodological competencies [Gronau et al., 2017; Hertle et al., 2017; Müller et al., 2017; Bohlouli et al., 2017; Lili, 2017; Arous, 2011]. Other papers use an action verb-based taxonomy of competencies: “Remember, understand and apply” to address technical and methodological competencies and “analyze, evaluate, create” to address conceptual competencies [Enke et al., 2015].

2.2 Competencies Management

When going through scientific papers related to the research subject, such as in [Bonjour et al., 2002; Belkadi et al., 2007], it has been noted that competencies management addresses different aspects, namely the *identification*, the *modeling*, the *evaluation*, the *development* and the *usage* of competencies.

The identification of competencies can be declarative or deductive. Indeed, this operation could be done through annual appraisal or intermediate interviews with the individual or with his/her supervisor, or through tests, observation of the person and traces that he/she leaves [Berio & Harzallah, 2005; Loufrani-Fedida, 2006].

In order to identify and model competencies, companies such as IBM, HP and TEMEX have developed individual competencies repositories according to these up-cited identification approaches. The major goal of these repositories is to facilitate the assignment of employees, by helping managers identify and select the actors having the required competencies for the job [Loufrani-Fedida, 2006].

The literature also shows the use of other modeling tools such as matrix models [Daniel et al., 2009; Decius & Schaper, 2017; Hlaottinun et al., 2008], competence trees [Bohlouli et al., 2017], Bloom’s taxonomy [Enke et al., 2015], a verb-based model that focuses on the knowledge of individuals and mathematical models [Müller et al., 2017; Bohlouli et al., 2017; Lili, 2017; Daniel et al., 2009; Shahhosseini & Sebt, 2011; Stadnicka et al., 2017; Guillaume et al., 2014; Harzallah, 2000, Hlaottinun et al., 2008] that are useful in the competencies usage phase. In [Belkadi et al., 2007], a situation-based approach was developed in order to model and characterize competencies through the traceability of design activities, and led to a repository of design competencies.

As for competency evaluation, this operation can take three different forms. The first form concerns the individual’s level of competency itself. The second form consists in the comparison of this level with the level required by the company, define the gap (positive or negative) and act accordingly. Here, the development either of competencies (negative gap) or of career (positive gap) is addressed. The third form is necessary when comparing individuals with each other according to their level of competency. In this case, the problem of human resource allocation is tackled [Skorková, 2016], in other words competencies mobilization.

As for competencies development, it is done by means of training: it is provided by the company or is the initiative of the employee. Companies nowadays rarely use classical training in specialized centers or through e-learning interfaces; rather a

migration towards a model that ensures adequacy with work situations is noticed (training inspired from work situations, applied in work situations or based on work situation) [Berio & Harzallah, 2005]. Moreover, concepts such as workshops and learning factories have emerged. Throughout literature a special interest is given to learning factories [Gronau et al., 2017; Müller et al., 2017; Mayer & Toubal, 2016]. For example, in [Gronau et al., 2017], a didactic concept was developed enabling competencies development and training related to the IoT context by using an IoT laboratory. Festo, which is a German mid-sized company specializing in automation systems, experimented new training methods to develop the adaptability and versatility of its employees by the creation of mini training centers within its factories. These structures developed training programs as close as possible to the real conditions of production [Mayer & Toubal, 2016]. Workshops are also used within industries. For instance, in [Hertle et al., 2017], a work-oriented approach was proposed. It included five workshops and a company-specific learning factory, which aim to develop relevant technical, methodological and social competencies. In addition, problem-solving circles are used to integrate competency development into daily work routines.

When it comes to competencies usage, their mobilization is closely related to the possibility to inquire the acquired and required competencies [Berio & Harzallah, 2005] with the aim of assigning the most competent person to work on a certain mission. For this, several different tools, between mathematical resolution and computer software, have been created. Some studies address mathematically the problem of task allocation under constraints of competencies [Guillaume et al., 2014; Nasibov, 2007]. There are some interesting papers where researchers have modeled the uncertainty on the competencies possessed by the person and the imprecision on those required for the task. Semantic methods based on keywords retrieval and research engines were also used for employees allocation. However, finding exact matches seems to be too much restricted. Therefore, a similarity measure between competencies or approximate search was added to these classical semantic methods [Colucci et al., 2011]. A solution that combines software technologies, mathematical and statistical methods and offers tools to experts in real enterprise environments to either evaluate employees for a job assignment and training or for recruitment processes, was as well developed [Bohlouli et al., 2017]. Another tool, quite simple, used by a good number of companies is the competencies matrix. It provides a visualization of the stock of competencies within the firm and serves as a decision-making tool regarding the staff allocation and the need for competencies development [Decius & Schaper, 2017].

In addition to the resolution models and computer solutions, software for competencies management were identified on the net. These latter are marketed in the form of solutions dedicated to the management of human resources, such as "Inser Compétences" which responds very concretely to the problem of proactive management of training tools. "Eurécia" is another human resources management solution with a module dedicated to the management of competencies by assessing the adequacy between the existing and the expected. "SkillsBoard" is also a solution that offers a collaborative competencies management and a wide range of functionalities such as the management of job/competencies repositories, the planning and management of evaluations and training, the search for competencies and their matching.

All of the methods and solutions mentioned above to mobilize competencies are just examples of what might be found online

or in the literature. The field explored through this work is too wide and cannot be limited to these examples. Furthermore, researchers and practitioners are constantly seeking for the best solution, scientific or commercial it is.

2.3 Articulation between Competencies and PM

The question that remains to be asked now is: *are such competencies management tools used to successfully manage projects? Are competencies addressed when dealing with project management?* Papers on this issue are rare. The majority of these works theoretically address the importance of projects and competencies articulation, but this importance has not really been quantified except in certain cases where the allocation of resources in projects was treated. Some of these cases are already drawn in [Boucher et al., 2007; Belkadi et al., 2007] and dealt with industrial engineering projects, such as design, production management, organization and supply chain projects.

In a context of project management, practitioners are more focused on the achievement of projects' objectives while maximally respecting budget, due dates and clients' requirements. However, in such entities of organization there is always risk of complications. The efficiency of the project is certainly based on the definition of tasks, their planning, the analysis of project risks, but also on the competencies of the actors involved, their ability to mobilize the relevant knowledge to design the product or service, or their ability to acquire them [Charue-Duboc & Midler, 2001; Loufrani-Fedida, 2006; Ben Mahmoud-Jouini, 1998]. The project is thus the result of the good combination of competencies within the enterprise.

While competencies require time to become more reliable and develop, project management increasingly responds to the need of reducing delays and developing new products and services. The combination of these two different logics raises difficulties and can be the cause of many discrepancies. However, researchers, such as [Loufrani-Fedida, 2006; Ben Mahmoud-Jouini, 1998; Charue-Duboc & Midler, 2001; Paraponaris, 2000], claim that this articulation is fundamental to ensure the sustainability of companies. However, their works were essentially exploratory and descriptive, even though it involved real PBOs studies. For these researchers, projects offer a validation environment, a basis for gaining experience, experimenting and memorizing competencies. It is therefore a matter of using existing competencies and building new ones.

Competencies were also linked to resource allocation within projects [Daniel et al., 2009; Stadnicka et al., 2017; Shahhosseini & Sebt, 2011; Hlaoittinun et al., 2008]. This is the case, for instance, of "the Best-Fitted-Resource methodology" helping decision makers allocate their human resources in software projects, by determining the suitability between the complete set of competencies available in candidates and required competencies for tasks. It is as well the case of aircraft maintenance projects where aircraft industries have to optimize the distribution of operators according to their competencies in order to minimize total service time and delays; or even more the case of construction projects where potential employees are traditionally interviewed and the most qualified one are subjectively selected. In [Hlaoittinun et al., 2008], a competency-modelling-based multidisciplinary team building method was developed. It uses jointly a competency matrix, a clustering algorithm and a task assignment mathematical model. Such a tool helps project managers, mainly design project managers, to balance between project costs and competency dynamics.

To summarize what was discussed in this section, literature

showed that competencies management tools and methods differ in terms of the purpose: identification, evaluation, development or mobilization. Each tool developed by researchers in the field is limited to a certain context of the industry. Some competencies management tools were created to try overcoming the limits of others. Commercialized IT solutions are more oriented towards the search for profiles and the assessment of the adequacy between required and possessed competencies. They are used rather in HR management (HRM) services and do not address projects.

As already mentioned, papers dealing with both competencies and projects management are scarce. Such papers only treated the importance of this articulation in terms of the role that competencies play in the success of projects and in terms of the role that projects have in the development of competencies. Little are the papers that were interested in their usage and are proving empirically the relevance of competency-based methods and approaches. However as it was discussed in literature, competencies management and PBO cannot be considered as two separate spheres. With this, the following questions could be asked: *are the tools identified in the literature used by real companies? If not, what tools and methods are currently employed in PBOs to manage competencies? Do such tools and methods mainly address competencies, as in the literature, from a human resources management point of view or do these organizations obey to the necessity of projects and competencies articulation?* Along this study, research papers dealing with real cases were found, especially in industrial engineering and IT fields [Daniel et al., 2009; Stadnicka et al., 2017; Shahhosseini & Sebt, 2011; Hlaioittinun et al., 2008, Belkadi et al., 2007; Boucher et al., 2007, Bonjour et al., 2002]. Still, it is not sure that such tools developed as part of research works are afterwards deployed within or supported by companies.

In order to answer the questions asked above, it is necessary to explore firms' world and involve people working within PBOs in this study. Real cases will therefore be treated in order to answer the research question in its practical aspect.

3 EMPIRICAL STUDY

3.1 Methodology

The study conducted through this research is exploratory. Indeed, the aim is to extract relevant information about how competencies are managed within real-cases of PBOs. The focus is particularly put on competencies management software and methods used to extract and mobilize competencies. To do so, a qualitative approach was adopted. Based on Blanchet and Gotman's book "Survey and its methods: the interview" (English translation of "L'enquête et ses méthodes: L'entretien"), the field of PBOs is explored through semi-structured interviews as a qualitative approach. These interviews allow collecting precise information on considered subject and giving the interviewees freedom to express themselves without deviating from the subject [Blanchet & Gotman, 2005]. To succeed in conducting these interviews, it was necessary to build an interview guide. In parallel, a sample for data collection was constructed. The study ends with a thematic analysis based on verbatim transcripts of the recorded interviews.

3.2 Data Collection

The phase of interview guide construction is considered critical since this guide is the basis of the empirical study. The development of the guide begins by defining what is

investigated, which is, in this case, the research question. It is then eligible to identify how the phenomenon under study takes shape in the practice of interviewees in order to finally define the points to be addressed to access it. Therefore, structuring factors of the guide were determined in the first place. As methods and tools mobilized by PBOs in competencies management together with the articulation between competencies management and project management are studied, the main elements structuring the questionnaire are the enterprise, the profile of the interviewee and his/her position, project characteristics, project dynamics as well as project management tools and competencies management tools. For each factor, indicators helping to solve the problematic are associated. Each indicator corresponds to one or more questions in the guide. Some questions serve as reminders to help the interviewees express their ideas. In fact, the guide is essentially based on situations' narration. The interviewees are asked to choose one example of the projects in which they participated. Hence, the interview is conducted around this project.

This exploratory study targets companies practicing management by projects. The objective is to interview a large number of people in different fields in order to answer the research question defined above. The data collection method is qualitative, thus, telephone interviews were conducted, due to the difficulty of visiting companies added to the possibility to record conversations. The key idea is to use social networks such as "LinkedIn". Indeed, it is a growing subject that attracts the attention of a large curious and ambitious community. Following this method, information was collected through seven interviews with individuals from different companies and domains. Table 1 summarizes the data for each of these companies. These interviews lasted on average one hour each. After they agreed to collaborate with us, interviewees were contacted by telephone to schedule an appointment. At the beginning of each interview, the context and objectives of the study are reminded. The interview procedures are then indicated. In addition, interviewees were asked permission to record the conversation and topics to be covered throughout the interview were presented. Two of the interviewees expressed their dissatisfaction with the recording by indicating that they will not answer anything that is confidential. The seven people who replied to us showed from the start their interest in the subject because it affects them directly and asked for feedback on the results.

3.3 Data Analysis

The interviews were recorded under the agreement of interviewees. This recording served to analyze the collected data. Indeed, these interviews were transcribed word by word and were studied thematically.

Thematic analysis is considered as the best method for conducting an exploratory study [Blanchet & Gotman, 2005]. The three most relevant interviews were used at the beginning to build a table of keywords related to the research topic: it is the coding phase. From these keywords, themes and sub-themes around which the analysis is articulated were defined. Then, the rest of interviews were examined and organized according to the identified themes. Succeeding this step, links with the literature and between themes and sub-themes were studied.

Examining the transcripts, a list of key words was organized according to four main themes, namely, project-based organization, competencies management, contextual framework and articulation between competencies and project management. The first two themes seem obvious as they explicitly reflect the research question. However, what was

Table 1. The studied sample and used tools

Interviewees	Positions	The enterprise's business sector	Project management tools		Competencies management tools & software				
			Tool name	Purpose	For Identification & modelling	For Evaluation	For Development	For Usage	Software
Interviewee 1	Strategy & innovation purchasing director	Digital transformation and automation of energy management	Unmentioned.	Unmentioned	Competencies repository. Individual annual interviews.	Individual competencies review. Procurement Excellence system to assess methodological and process competencies. SWOT to assess collective competencies.	Individual development plan. Training. Experience in projects.	Competencies assessment and identification are for developing and talent recruiting purposes. Mobilization is then done based on positions not competencies.	Talent Link
Interviewee 2	Co-founder and consultant	Information technologies and services	MS Project / Excel / Redmine	Manage schedule and budget.	Competencies matrix. Individual annual interviews.	Individual annual interviews. Self-assessment.	Individual development plan. Training. Self-training. External competencies recruitment. Experience in projects.	Talents chasing for project execution. Assignment is not purely based on competencies.	Competencies matrix on Excel
			Issue Log	Orders and organizes problems by type and severity in order to prioritize issues associated with the current milestone or iteration.					
Interviewee 3	Manager and electrical engineer	Building engineering (Energetic performance)	Excel	Manage schedule and budget.	Individual interviews every 2 years.	Individual interviews every 2 years.	Training. Experience in projects. External competencies recruitment.	Assignment is not based on competencies.	None
			Graphic interface	Model the processes of certain phases.					
Interviewee 4	Consultant and technical expert	Information management services	MS Project	Manage schedule and budget.	Competencies matrix on Excel.	Competencies matrix on Excel.	Training. Colleagues help.	Talents chasing for project execution. Assignment is not purely based on competencies.	Competencies matrix on Excel
			Internal software to assign HR	Assign project team members with respect to two conditions: the project capacity and the rest to be done, taking into account holidays as well.					
Interviewee 5	The head of the company's training academy and training consultant	Trade credit insurance	MS Project Excel	Manage schedule and budget.	Competencies repository. Individual annual interviews.	Competencies repository. Individual annual interviews.	Training. Colleagues help.	Management by profession.	None
Interviewee 6	Instrumentation engineer	Oil and energy (subsea, onshore/offshore and surface projects)	Tools developed by the company	Manage documents on internal databases. Manage the schedule.	Competencies repository. Individual annual interviews.	Competencies repository. Individual annual interviews.	Training. Experience in projects. External competencies recruitment.	Assignment is not based on competencies.	HRweb
Interviewee 7	Operation manager	Urban planning	GeoSas	Used by companies that practice urban planning. Allows to have financial balance sheets of the operations and to integrate all the project phases in terms of delays and budget distribution.	Individual annual interviews.	Individual annual interviews.	Training. Colleagues help.	Assignment is not based on competencies.	None
			The software "Gestion de projet"	Allows project follow up.					

Table 2. Data Analysis : Themes and Subthemes

Themes	Subthemes	Definition
Project-based organization	PBOs' characteristics	Organization and structuring of the company in the context of project-based management. The size of projects, their complexity and typology, as well as their place in the company are also considered.
	Actors	Participating entities and elements in all the activities related to project management. The constitution of project teams is also considered.
	Tools for project management	Tools mobilized to carry out the projects. They are computer-based solutions.
	Encountered problems	Difficulties and conflicts related to project management and to the complexity of PBOs context.
	Implemented solutions	The ways in which these problems were confronted.
Competencies Management	Required and acquired competencies	Competencies required to work on projects, as well as competencies that a person can acquire in such an environment. By this, one of the aspects of the articulation between competencies and projects is tackled. Collective competencies are also considered.
	Competencies management devices	Methods, approaches and tools used for competency management.
	Actors	Key actors responsible of competency management and the integration of employees in the enterprise and in projects.
	Encountered problems	Difficulties and conflicts related to competencies management, especially to the lack of competencies.
	Implemented solutions	Actions executed to solve the encountered problems.
Context	Internal	Impact of company internal elements on project and / or competencies management.
	External	Impact of external factors on project and / or competencies management.
Articulation between competencies and project management	Competencies importance in projects	Are enterprises aware of competencies role in the success of projects? → Determine competencies position inside the company.
	PBOs maturity in competencies management	The company's level of maturity in relation to competencies management and their mobilization in projects.

mainly studied through the analysis is the interaction between these two themes. Using the list of keywords, each mega theme was also broken down into sub-themes. Table 2 shows this decomposition and defines each subtopic. In what follows, details of the thematic analysis will be given by referring to excerpts from the interviews' transcripts.

Project-based organization. Among the companies dealt with in this empirical study, there are four large international companies. The rest are small businesses or startups. To measure the dependence level of the company's activities on projects, it was determined through the conducted interviews that all the enterprises use projects as a mode to operate on the market. For instance, interviewee 1 affirmed that they *"have a big part of activities organized by projects"*. Interviewee 6 said that *"the company is organized according to the sectors ... these are separate sectors ... and of course each sector is organized by projects"*. As for interviewee 3, they work in project mode they have no choice, all their activities are in project mode.

Interviewees have classified projects within their firms according to their size, complexity, activity or outputs. Large projects and small ones, complex projects and simple projects, business projects (B2B projects) and proposal projects (B2C projects), development or transformation projects are distinguished. Another classification has been proposed if the projects are addressed from a customer point of view: package projects and client-managed projects.

As for the actors involved in the conduct of projects, there are several appeals in the studied PBOs that are essentially related to the context. However, the project manager remains an essential element in such structures. The client is also a fundamental part because without its needs, we would have no project to lead. Internal and external resources to the company are as well solicited to execute projects.

In most of the studied cases, the project manager chooses project team members and the staffing. This project leader can be an outsider client or an employee within the company.

His/her choice is based on the availability, the skills, the knowledge, the experience and the reputation of individuals and is done in such a way that the members will have complementary competencies: the promoter will not pay twice the same thing. He arranges to have complementarities in all areas (interviewee 3). The goal is to vary competencies and to have a team where there is all the competencies that we seek for the project realization (interviewee 7). Moreover, it is admitted

by the interviewees that the project team composition changes according to the project phases and other factors. Staffing and members' selection are done according to the nature of activities during the project. Generally, there are phases in the project, and we do not find the same teams in each phase (interviewee 1). During one phase, the company tries to keep as much as possible the same team composition to avoid delays and leave a history, especially when this phase lasts a long time (interviewees 4 and 7). However, this composition can also change if individuals quit their jobs for either personal, professional or medical reasons (interviewees 3 and 5).

In terms of tools deployed to better manage projects, firms use either internal tools or the tools the client uses itself. The most commonly used software are Excel and MS Project (see Table 1). Nevertheless, some organizations prefer to use their own developed software such as the enterprise where work interviewees 4 and 6. In the interviewee 6 case, for example, they use their own-developed software. They have a large IT team who can develop software that exactly meets the enterprise needs. The nature of their work is such that they cannot work with standard software because they are not adapted to their way of doing.

Situational narratives were used to highlight the problems and conflicts that interviewees may encounter in a project management context. The examples mentioned by the interviewees were therefore various. These problems can be categorized into deliverable quality issues, dead-time issues and budget issues. Each class of problems can be the consequence of the other. Failure to meet deadlines, budget and quality may be related to organizational problems (interviewee 3), delays of suppliers or providers (interviewees 1, 5 and 6), complexity of tasks (interviewees 1 and 5), lack of communication and coordination (interviewee 2), lack of motivation (interviewee 1 and 4), poor expression of need (interviewees 1, 2 and 7) but especially lack of experience and competencies (all interviewees). Problems related to competencies will be addressed later in the competencies management part. Different solutions and action plans have been implemented to solve such problems. The resolution is done mainly through meetings where project team members meet to discuss ongoing issues and try to find answers (all interviewees). *"The frequencies of the meetings are adapted to the issues and problems. When everything is progressing well and there are no worries, the frequency is low and we increase the frequency when we have*

problems. We have topics that are treated, either problems or risks. Sometimes things do not work, or things may not work, or they may be questionable. In fact, this is just the problem solving of the project” (interviewee 4). For instance, the enterprise where works interviewee 6 uses a penalty policy to avoid problems with vendor delays. In addition, this company mobilizes a team of project engineers dedicated exclusively to resolving the difficulties encountered during the projects.

Competencies Management. Depending on the company intervention field, the competencies required in the projects and mentioned by the interviewees are different. However, they can be classified, like what was found in literature, into technical competencies related to the tasks and the used technologies, methodological competencies related to the experience and the expertise of individuals and personal competencies related to behaviors and attitudes, as well as interpersonal competencies of collaboration and mutual aid. Competencies building is done either outside projects or during projects. In fact, there are two methods deployed by the studied firms: either the training is used (all the interviewees), or the competencies are acquired and improved through the exploration and participation in projects (interviewees 1, 2 and 5). Some interviewees rely on self-training to develop their competencies (interviewees 1 and 2). As for the competencies possessed by the employees and the acquired competencies, they are assessed on an ad hoc basis through annual interviews and according to a declarative approach in which the evaluated person and / or his manager is involved. The actors involved in this evaluation are therefore the person whose competencies are assessed, his/her manager and HR managers. Before doing this assessment, a previous step is very important which consists in identifying the competencies for each position or task. The HRM team and a business referent create a repository where a generic list of key competencies needed for the execution of tasks exists. The individuals are then evaluated according to this two others use what is called a matrix of competencies on Excel as they claim that they find it quite simple and efficient. Table 1 summarizes the tools used by the under-study firms. All these are essentially deployed by the human resource department, in a logic of talent acquisition and for a main objective, which is the development of the company’s human capital.

The subject of collective competencies, despite the fact that it was mentioned many times by the interviewees, these competencies were always associated with the terms “project team”, “group work” and “company”. For interviewees, as well as their companies, collective competencies are not important as much as individual competencies. For them, as long the necessary individual competencies are available, a team can work all together to make the project a success.

Regarding competency-related problems, all interviewees mentioned these. They mentioned that there are always delays due to lack of competencies or poor staffing of the project. Sometimes it can happen that a conflict is misinterpreted and confused by a personal conflict when in reality the problem lies in a lack of competencies (interviewee 2). The solutions implemented to overcome this kind of problems is the replacement of incompetent people by others who are much more competent and we can even resort to external recruitment. Sometimes the client is dissatisfied with the competencies deployed by its provider, so it gets involved in the process of choosing members to critical tasks. The solutions are therefore instantaneous. Companies do not seek long-term solutions.

Context. Throughout the interviews conducted during this study, it was clarified that the context has an important influence on the way in which the company manages its projects, even the

competencies of its employees. Therefore, this mega theme has been divided into two parts. The first one concerns everything that is internal to the company and the second part is about the firm’s external environment. The internal context especially covers the nature of the company’s activity, its working strategy, its vision, its values and rules, the available means and resources, etc. As for the external context, it includes everything related to the market, customers, suppliers or service providers, regulations, technological evolution, etc.

The company’s activity sometimes forces it to use project management tools and work approaches among others. This is the case of interviewees 6 and 7 (see table 1).

Large companies have a clear policy in relation to the management of their projects but especially in relation to competency management. According to interviewee 3, they are few so they all know each other. They know their competencies and they improve them progressively. Yet they do not have that culture of saying, here are our competencies and let us make a list of them. Interviewee 1 shares the same opinion: “*very often in small structures as we do not have resources, when we created the four startups for example, we all did a little of different trades. It is good. However, at one point it has limits ... when a startup is created the first challenge is to build a team and then start assigning the roles, who does what. The more we are within matrix organizations where you have 10 leaders and you work around the world, the more complex it is and if you do not put everything in order it does not work*”.

The conversion in the PBO’s strategy has changed the required competencies in projects. The shift, for example, from production to services requires soft competencies around, according to interviewee 1, cultural cross, cooperation, partnership and networking. Further, according to all interviewees, the technological revolution has a major impact on management methods, especially on the required competencies within companies. Today’s trades are no longer those of tomorrow, so an adaptation of competencies is needed to deal with such changes (interviewees 1, 4 and 5).

In the “problems encountered in PBOs” section, it was already mentioned that the client and suppliers can delay the project: the client with its requirements, its poor expression of need, dissatisfaction with the competencies mobilized for the project and suppliers because of resources and competencies lack. As for the regulatory constraints, they are always taken into account before the start of projects. “*There were a lot of regulatory constraints. We have to anticipate everything to avoid problems of this kind*” (interviewees 3 and 6).

Articulation between competencies and project management.

The duality between PM and competencies management was mentioned by the interviewees only in terms of competencies development through projects. Even the tools used for competencies management are not deployed for the benefit of the projects; an independence between the two management systems is noticed. In fact, competencies management is done according to a job or a position and not to projects or tasks. Even if the project manager is the one who assigns people to projects, he bases himself on his intuition, his experience and/or on the perception of human resources managers.

In terms of maturity in competencies management, it has been noted that large international PBOs (case of the enterprises which interviewees 1 and 6 are part of) are more aware of such a concept importance in the progress of their businesses. Indeed, these companies have set up a whole system of competencies management, which remains highly depending on and managed by the HR department.

3.4 Discussion

Comparing what was found in literature with the results of the empirical study, it was concluded that both PBOs and researchers are aware of the role that competencies play in the success of projects. However, the policy used by these PBOs for competencies management is oriented for HRM purposes. A process of recruiting for positions (designer, production manager...) and not for projects is rather followed. In a logic of articulation, there is no tool used in the studied PBOs that deals simultaneously with projects and competencies management. This leads to inquiry the fundamental causes of such weak instrumentation. Is it linked to the uncertainties on the projects, to their more or less long duration? Is it linked to the high turnover of employees in such organizations? Is it related to know-being-problems difficult to anticipate? In all cases, such uncertainties could cause difficulties in identifying competencies required for projects beforehand. In addition, confusion around the project leader's authorities could prevent good competency management in a project context [Flamant, 2002; Midler, 1993], not forgetting to mention the confusion that business line actors in particular and actors in the company in general may have regarding to the definition of competency and its categorization [Coutant, 2014]. Indeed, these questions require further investigation, but what is as well certain is that a merger between the strategic and operational levels of PBOs must be made, the HR department and the project actors should work jointly within the company. A real paradox that needs to be resolved is the fact that human resources should not be directly related to HR department. Furthermore, companies should show more interest, invest in academic research and work together with researchers on the problematic of competencies and projects articulation. As far as possible, they could offer a validation or experimentation field to the already-developed methods and tools.

4 CONCLUSIONS & PERSPECTIVES

Although the tools for managing competencies vary in literature and there is a slight orientation towards what is deductive to identify and evaluate competencies, PBOs continue to use traditional declarative methods even if they were proved to be inefficient and do not necessarily give reliable results. This study allowed confronting French companies with competencies management together with project management. Despite the significant results obtained, this research has some limits: the sample size is very small and the empirical analysis was done manually. The themes identified for the analysis depend heavily on the person conducting the study and may therefore not be the same if another researcher led it. Thanks to this work, the design of an ideal tool for managing competencies in projects could be undertaken. Such a tool can cover all aspects from identifying human competencies to their evaluation, assignment and development. A questionnaire can be drawn up on the basis of the results here-presented to identify the selection criteria and the characteristics of an ideal tool for competencies management. Analysis could be done quantitatively for more relevance.

5 ACKNOWLEDGMENTS

The authors would like to thank all the interviewees who accepted to participate in the study and made this research possible. This work was conducted in 2019 and is part of the ANR project CaTCaP.

6 REFERENCES

- Arous, K. (2011). Les pratiques de la gestion des compétences : Cas des grandes entreprises tunisiennes. CREG – Université de Pau et des Pays de l'Adour.
- Belkadi, F., Bonjour, E., Dulmet, M. (2007) Competency characterization by means of work situation modelling. *Computers in Industry*, Elsevier, 2007, 58 (2), pp.164-178.
- Ben Mahmoud-Jouini, S. (1998), *Stratégies d'offres innovantes et dynamiques des processus de conception*, Thèse de Doctorat en Sciences de Gestion, Université Paris IX- Dauphine.
- Berio, G. & Harzallah, M. (2005). *Knowledge Management for Competence Management*, 0(1), 21–28.
- Blanchet, A., & Gotman, A. (2005). *L'enquête et ses méthodes : l'entretien*. Paris: Armand Colin.
- Bohlouli, M., Mittas, N., Kakarontzas, G., Theodosiou, T., Angelis, L., & Fathi, M. (2017). Competence assessment as an expert system for human resource management: A mathematical approach. *Expert Systems With Applications*, 70, 83–102.
- Bonjour, E., Dulmet, M., & Lhote, F. (2002, October). An internal modeling of competency, based on a systemic approach, with socio-technical systems management in view. In *IEEE International Conference on Systems, Man and Cybernetics* (Vol. 4, pp. 6-pp). IEEE.
- Boucher, X., Bonjour, E., & Grabot, B. (2007). Formalisation and use of competencies for industrial performance optimisation: A survey. *Computers in industry*, 58(2), 98-117.
- Charue-Duboc, F. & Midler, C. (2001), *Développer les projets et les compétences : le défi des hiérarchiques dans les métiers de conception*, *Gérer et Comprendre*, 63, mars, 12-22.
- Colucci, S., Tinelli, E., Di Sciascio, E. & Donini, F.M. (2011). Automating Competence Management through Non- standard Reasoning. *Engineering Applications of Artificial Intelligence* 24, 1368-1384.
- Coutant, H. (2014). La « technique » comme activité ou comme représentation partagée: Une ethnographie de la R&D dans une entreprise aéronautique. *Annales des Mines - Gérer et comprendre*, 117(3), 49-58.
- Daniel, L., Centeno, G., Ruiz-torres, A. J., & Otero, C. E. (2009). A systematic approach for resource allocation in software projects. *Computers & Industrial Engineering*, 56(4), 1333–1339.
- Decius, J. & Schaper, N. (2017). The Competence Management Tool (CMT) – A new instrument to manage competences in small and medium-sized manufacturing enterprises. *Procedia Manufacturing*, 9, 376–383.
- Di Vincenzo, F., & Mascia, D. (2012). Social capital in project-based organizations: Its role, structure, and impact on project performance, 30, 5–14.
- Flamant N. (2002). « On ne sait plus qui est le chef ». *Travailler et diriger dans l'industrie spatiale*, Terrain, n° 39, pp. 109-120.
- Durand, T. (2000), *L'alchimie de la compétence*, *Revue Française de Gestion*, 127, janvier-février, 84-102.
- Enke, J., Kraft, K., & Metternich, J. (2015). Competency-oriented design of learning modules. *Procedia CIRP*, 32(C1f), 7–12.
- Gilbert P. & Thionville R. (1990). *Gestion de l'emploi et évaluation des compétences*, Paris, ESF Ed.
- Gronau, N., Ulrich, A., & Teichmann, M. (2017). Development of the Industrial IoT Competences in the Areas of Organization, Process, and Interaction based on the Learning Factory Concept. *Procedia Manufacturing*, 9, 254–261.
- Guillaume, R., Houé, R., & Grabot, B. (2014). Robust competence assessment for job assignment. *European Journal of Operational Research*, 238(2), 630–644.
- Harzallah, M. (2000) *Modélisation des aspects organisationnels et des compétences pour la réorganisation d'entreprises industrielles*, Ph.D. Thesis, University of Metz, France.
- Hertle, C., Jokovic, B., Weber, C., Tisch, M., Meißner, A., Ardelt, T., Metternich, J. (2017). Innovative approaches for technical, methodological, and socio-communicative competency development in production areas. *Procedia Manufacturing*, 9, 299–306.
- Hlaottinun, O., Bonjour, E., Dulmet, M. (2008). A multidisciplinary team building method based on competency modelling in design project management. *International Journal of Management Science and Engineering Management*, 3, pp.163-175.
- Le Boterf, G. (1994), *De la compétence : essai sur un attracteur étrange*, Paris : Editions d'Organisation.
- Lili, Z. (2017). An Inverse Optimization Model for Human Resource Allocation Problem Considering. *Procedia Computer Science*, 112, 1611–1622.
- Loufrani-Fedida, S. (2006). *Management des compétences et organisation par projets : une mise en valeur de leur articulation*. Gestion et management. Université Nice Sophia Antipolis.
- Bidet-Mayer T., Toubal L., (2016). *Mutations industrielles et évolution des compétences*, La Fabrique de l'industrie (n° 5, Avril 2016) 12 p.
- Midler, C. (1993). Le responsable de projet, portrait d'un rôle d'influence. *GESTION* 2000, 9, 123-123.
- Midler, C. (1995). "Projectification" of the firm: The renault case, *Scandinavian Journal of Management*, 11(4), 363-375.
- Müller-frommeyer, L. C., Aymans, S. C., & Bargmann, C. (2017). Introducing competency models as a tool for holistic competency development in learning factories: Challenges, example and future application. *Procedia Manufacturing*, 9, 307–314.
- Nasibov, E. N. (2007). A problem of task allocation with fuzzy information and twostage solution algorithm. *Automatic Control Computer Sciences*, 41(4), 196–202.
- Nordhuag O. (1996), « Collective competences in organization », *Rethinking the Boundaries of Strategy*, Copenhagen Business School Press.
- Paraponaris, C. (2000), *Gestion des compétences et production des connaissances dans le management par projet*, *Revue de Gestion des Ressources Humaines*, 36, juin, 3-17.
- Shahhosseini, V., & Sebt, M. H. (2011). Sharif University of Technology Competency-based selection and assignment of human resources to construction projects, 18, 163–180.
- Skorková, Z. (2016). Competency models in public sector. *Procedia - Social and Behavioral Sciences*, 230(May), 226–234.
- Stadnicka, D., Arkhipov, D., Battaia, O. & Chandima Ratnayak, R.M. (2017). Skills management in the optimization optimization of Skills aircraft management, 6912–6917.
- Turner, J.R., Huemann, M., Anbari, F.T., Bredillet, C.N. (2010). *Perspectives on Projects*. Routledge, New York.

Supplying wood from private forests: Review of the barriers and means to increase mobilization

MINH LE, LUC LEBEL, NADIA LEHOUX

UNIVERSITÉ LAVAL
2325 rue de l'Université, Québec, G1V 0A6. Canada

minh.le.1@ulaval.ca, luc.lebel@sbf.ulaval.ca, nadia.lehoux@gmc.ulaval.ca

Résumé – L'augmentation de la consommation de ressources forestières met en lumière le besoin d'utiliser les forêts à leurs pleines capacités en respect leur durabilité. La mobilisation de bois, ou l'idée de récolter le bois disponible sans dépasser les possibilités forestières dans le but d'approvisionner les usines de première transformation, est une approche qui est prometteuse dans le contexte où un résultat rapide est souhaité. Une revue systématique de littérature professionnelle et scientifique a permis d'identifier les mesures et moyens qui permettraient une augmentation de la mobilisation des bois des forêts privées. Les résultats suggèrent (i) que des initiatives de mobilisation de bois, p. ex. le développement de capacités techniques ou la consolidation de moyens de production, peuvent réussir lorsqu'elles sont adaptées à un contexte régional spécifique, (ii) que ces dernières peuvent produire des résultats positifs quoique pas toujours quantifiés, (iii) que la littérature sur les initiatives de mobilisation réussies et quantifiées n'est pas étendue et (iv) que la mobilisation de bois par des moyens logistiques pourrait produire des résultats favorables. Un cadre de référence est proposé pour traiter cet enjeu au Québec (Canada) où la demande en fibre est croissante et où la forêt privée pourrait contribuer aux solutions.

Abstract – Increasing consumption of wood resources has put to light the need to use forests to their full sustainable potential. Wood mobilization, or the idea of harvesting the wood available without exceeding annual increments and making it available to processing plants, is an approach that shows potential particularly in situations where a rapid outcome is wished. A systematic literature review of professional reports and scientific articles on wood mobilization approaches and initiatives has identified ways and means that could lead to increased wood mobilization from private forests. The findings suggest (i) that wood mobilization initiatives, e.g. technical capacity building or production consolidation, can be successful if tailored to a specific regional context, (ii) that they may produce positive results, though not always quantified, (iii) that the literature on successful and quantified wood mobilization initiatives is not extensive and (iv) that wood mobilization through logistical means could yield favourable results. A wood mobilization framework is suggested to analyze the issue in Quebec (Canada) where fibre demand is increasing and where the private forests might contribute to the solutions.

Mots clés – Forêt privée, mobilisation du bois, collaboration, logistique, revue de littérature systématique.

Keywords – Private forest, wood mobilization, collaboration, logistics, systematic literature review.

1 INTRODUCTION

Wood consumption is increasing as a direct consequence of world population growth and its ecological merit as a material. The renewable nature of wood and its carbon sequestration capacity are two significant reasons explaining the increase in demand. According to the United Nations Economic Commission for Europe (UNECE) and the Food and Agriculture Organization of the United Nations (FAO), projected European and North American wood demand will keep growing until at least 2030 (UNECE 2010, 2012). In Quebec, round wood consumption by sawmills has increased from 21.6 million cubic metres (Mm³) in 2010 to 25.8 Mm³ in 2019 (Ministère des Forêts de la Faune et des Parcs, 2019). In Europe, wood transformation fulfills an important need for bioenergy (Hetsch *et al.*, 2007; Prins *et al.*, 2015). A similar trend is expected in North America, especially considering the biomass potential and the important role of bioenergy played in climate mitigation strategies (Durocher *et al.*, 2019; Röder *et al.*, 2019).

In a situation where annual felling is close to annual increments, i.e. a high wood utilization ratio scenario, increasing sustainable wood supply would require increasing wood production through more intensive forest management. In a situation where annual felling is significantly below annual increments, i.e., a low wood utilization ratio scenario, increasing sustainable fibre supply could be achieved through means that increases the rates of wood mobilization. Whereas a long time horizon is required to generate wood yield increase, wood mobilization can enhance wood supply on a shorter one. Schmithüsen and Hirsch (2010) showed that utilization ratio is generally lower in private forests in European countries dominated by public tenure. Smaller land areas and higher fragmentation levels in private forests are obstacles to wood mobilization. In contrast, in European countries with a large proportion of private ownership, such as Finland, utilization ratio in their private forests is often higher than in their public forest (*Ibid.*). In the province of Quebec, Canada, wood mobilization from private forests has the potential to satisfy a larger portion of wood demand. In 2019, 21.3

Mm³ of timber were harvested from Quebec's public forests, out of 32.7 Mm³ of annual increments, representing a utilization ratio of 65% (MFFP, 2019). For the same reference year, 6.4 Mm³ of timber were harvested from Quebec's private forests to supply primary processing plants out of 16.7 Mm³ of annual increments. Firewood production accounted for an additional 1.8 Mm³. The ratio of wood mobilization is 38% when excluding firewood production and 49% when considering it (MFFP, 2019). An estimation of the timber volume that can be realistically harvested considering historical supply to primary processing plants and demand forecasts suggests that an additional 2 Mm³ could be mobilized yearly from private forests (Fédération des producteurs forestiers du Québec, 2014). Wood mobilization should therefore be considered as a way to rapidly increase wood supply.

Lawrence (2018) provided a systematic literature review of wood mobilization interventions. The author reviewed how technical means, governmental measures and owner motivational factors affected harvesting levels. Several studies have also focused on forest logistics, especially transport operations (Audy et al., 2011, 2012; Audy & D'Amours, 2008; Dorval, 2015; Frisk et al., 2010). Logistic approaches to increase wood mobilization from private forests therefore seem to deserve further attention. The preliminary results from SIMWOOD pilot projects support that claim as close to 50% of the mobilization measures considered are logistical in nature.

The goal of this paper is to identify ways and means that could lead to increased wood mobilization from private forests. To achieve that goal, specific research questions were asked and a systematic literature review (SLR) was undertaken. Specific key words and research engines were targeted. The results were then organized to highlight the value of logistical solutions to wood mobilization and how they could be used in Quebec's context, where demand from sawmills has increased by 33% since 2009 and where the utilization ratio for public forests is already high (MFFP, 2019). It also allows us to build a wood mobilization framework that serves to analyze regional wood mobilization needs. The following sections are ordered as follows: first the detailed methodology for the SLR is presented, results are introduced next, a framework adapted to a specific regional context (Quebec) is then proposed and a conclusion highlights the key findings.

2 METHODOLOGY

A systematic literature review (SLR) was conducted to better understand how wood mobilization is addressed around the world. This SLR contributed to identify barriers to wood mobilization and means to increase it. A SLR is a process by which information on a subject is researched and collected in an organized and reproducible way (Tranfield *et al.*, 2003). The purpose of the SLR is to acquire an in-depth understanding of a

subject. To this end, the SLR was divided into three steps: planning, conducting and analyzing. First, the research orientation was determined. The strategy was to target specific information that answers a set of questions. In the case of this study, the results of the SLR answer specific research questions. The first three questions aim to establish a common understanding of key concepts related to wood mobilization whereas the last three questions aim to explore the potential of logistics strategies for wood mobilization from private forests, specifically in the province of Québec (Canada).

Q1: What is wood mobilization?

Q2: What are the barriers to wood mobilization from private forests?

Q3: What are the measures used to increase wood mobilization from private forests?

Q4: What is the place of logistics solutions in wood mobilization?

Q5: Could logistics measures increase wood mobilization from Quebec's private forests?

Q6: Are there existing wood mobilization frameworks suitable for Quebec's private forests?

As a second step, we determined what information would be researched and where. The research of the professional literature available for Quebec's private forests value chain was undertaken as a preliminary step. Organizations such as the ministère des Forêts, de la Faune et des Parcs (MFFP) and the Fédération des producteurs forestiers du Québec (FPFQ) have shared relevant knowledge and information pertaining to private forests and wood mobilization via professional reports and memoirs (Table 1). This information was obtained through the organizations' public websites or through personal inquiries.

Table 1: Organizations involved in Quebec's private forest that produced knowledge/information

Organization	Type
Ministère des Forêts, de la Faune et des Parcs (MFFP)	Government
Fédération des producteurs forestiers du Québec (FPFQ)	Forest owners' representative
Syndicats et offices de producteurs de bois	Forest owners syndicates/union
Agences de mise en valeur des forêts privées	Regional gouvernement body
Groupements forestiers Québec (formerly RESAM)	Forest owners association
Fédération québécoise des coopératives forestières	Forest cooperative network
Forêt compétence	Workforce association

Three academic databases (Web of Science, SpringerLink and OvidSP) and a search engine (Google Scholar) were used to survey the peer-reviewed literature. The keywords exploited are presented in Table 2. Search operators were also used to minimize search iterations.

A screening process was used to eliminate less relevant literature by analyzing the titles and abstracts first, then the introductions and the conclusions, and finally by evaluating the quality of the information and the research context. For example, literature specific to biomass mobilization was excluded as this research focuses on the mobilization of saw logs. However, literature pertaining to biomass mobilization was not excluded if saw logs and pulp log mobilization were both considered in the article. As a third step, the results of the SLR were analyzed. The contributions of each knowledge source were discussed and synthesized. The typology proposed by Orazio *et al.* (2017) was particularly useful to classify wood mobilization barriers and measures pointed out by the authors based on multiple studies.

Table 2: Keywords and operators used for the SLR

Keywords	Operators
(Private OR Small-scale OR Family) AND forestry	(Private OR Small-scale OR Family) AND forest*
Non-industrial forest owner	
(Wood OR Timber) AND Mobilization	(Wood OR Timber) AND Mobili*

3 RESULTS

Each of the following sections provide an answer to a research question stated in the methodology.

3.1 Q1: What is wood mobilization?

While the term “wood mobilization” started to become more frequent when Europe’s reflection on the future contribution of its forests was going on in the early 2000s, the concept of harvesting behaviour was already being discussed in the late 1980s and 1990s (Karppinen, 1998; Kuuluvainen *et al.*, 1996; Lönnstedt, 1989). de Galembert (2003) and Dengg (2003) suggested wood mobilization as a solution for European projected increase in wood demand.

Rauch and Gronalt (2005) defined wood mobilization as “*creating an accessible inventory for the mills from the growing forest asset*”. Orazio *et al.* (2017) defined it as “*initiatives and measures leading to harvesting and extraction of wood and woody biomass from forest*” with the purpose of enhancing “*the supply chain from raw timber and residues in the forest towards the industrial transformation into solid material, energy or biochemical uses of wood*”. Another definition for wood mobilization would be the increase of wood harvest for processing (MFFP 2017). In this study, because of Quebec sawmills’ increasing need for timber, we define wood mobilization as *increasing harvest levels and creating an accessible inventory for processing mills*. There is an emphasis on sawmills because of their increasing need for timber in Quebec (MFFP, 2019).

3.2 Q2: What are the barriers to wood mobilization from private forests?

Few authors have formally identified barriers for wood mobilization before de Galembert (2003). Rauch and Gronalt (2005) contended that wood mobilization is

curbed by the properties of small-scale ownership: weak market position, lack of forest management knowledge, low volumes supplied per forest owner, low machine utilization ratio and difficulties in promoting wood harvest and gaining support from public authorities. Hetsch *et al.* (2007) identified, during workshops, factors decreasing wood mobilization. They also proposed a framework to represent wood supply bottlenecks. Orazio *et al.* (2017) were the first to systematically identify wood mobilization barriers in parallel to wood mobilization means. They defined barriers as “*conditions that inhibit the development and uptake of effective solutions for wood mobilization*”. Lawrence (2018) has provided an extensive literature review on factors that affected harvesting behaviour.

In Quebec, Côté *et al.* (2015, 2017) identified that forest management activities, and thus wood mobilization, was affected by personal motivation and owners’ objectives. A shift in woodlot owners’ profile could explain why an increasing number of them are likely to seek to maintain the natural functions of their woodlots rather than to mobilize their wood for financial reasons, assuming that these objectives are not compatible. Table 3 is a synthesis of sources referring to specific barriers categorized in the typology from Orazio *et al.* (2017). In brackets is the number of independent studies that the source reports. While they provide an adequate typology, Orazio *et al.* (2017) do not give reported examples of barriers. Following the same typology, Table A1, in the appendix, synthesizes all the wood barriers identified by SIMWOOD’s regional pilot project on wood mobilization (Vuillermoz *et al.*, 2017).

Table 3: Sources reporting wood mobilization barriers following the typology of Orazio *et al.* (2017)

Barriers (category)	Source
Forest resource barriers	Hetsch <i>et al.</i> (2007) Lawrence (2018) [6] Rauch and Gronalt (2005) Stern <i>et al.</i> (2013) Vuillermoz <i>et al.</i> (2017) [5]
Regulatory and legal framework	de Galembert (2003) Hetsch <i>et al.</i> (2007) Vuillermoz <i>et al.</i> (2017) [1]
Finances and material inputs	Hetsch <i>et al.</i> (2007) de Galembert (2003) Lawrence (2018) [2] Vuillermoz <i>et al.</i> (2017) [2]
Organization and co-operation	de Galembert (2003) Hetsch <i>et al.</i> (2007) Lawrence (2018) [12] Rauch and Gronalt (2005) Vuillermoz <i>et al.</i> (2017) [10]
Knowledge and attitudes	Côté <i>et al.</i> (2015) Côté <i>et al.</i> (2017) de Galembert (2003) Hetsch <i>et al.</i> (2007) Lawrence (2018) [15] Rauch and Gronalt (2005) Vuillermoz <i>et al.</i> (2017) [9]

3.3 Q3: What are the measures used to increase wood mobilization from private forests?

Wood mobilization levels are enhanced by measures or means which are “a set of knowledge, techniques and/or a governance approach that is capable of lifting wood mobilization barriers” (Orazio *et al.* 2017). Rauch (2007) highlighted the potential of forest owner co-operation to increase wood mobilization from private forests. The author suggested that the most promising options are to optimize processes within the supply chain from the ownership to the industry, to use mechanized logging strategies and to develop new business models for forest owner co-operation. Hetsch *et al.* (2007) developed strategies to increase wood mobilization. A guide that identifies sustainable wood mobilization measures as well as examples from the field is also available by Ministerial Conference on the protection of Forests in Europe (MCPFE) *et al.* (2010). While a significant contribution because of more than 25 practice examples reported, outcomes in terms of additional volume mobilized was rarely available. Stern *et al.* (2013) emphasized the need to consider regional market conditions and forest owner types when implementing wood mobilization measures. Orazio *et al.* (2017) and the SIMWOOD project went further and produced a significant body of work regarding sustainable wood mobilization. The authors provide over 100 examples across each category of measures. However, the authors do not provide much detail concerning these examples. Lawrence (2018) reported that, from a final selection of 25 papers where wood mobilization outcome was evaluated, three were deemed unsuccessful and 22 reached their goals, “but often to a lesser degree than had been initially planned”. None of the peer-reviewed articles evaluated the net increase of wood mobilization, and only half of the professional reports did. Moreover, when the gains were evaluated, a direct indicator such as volume mobilized, was not always used. The impact of the intervention on harvested volume was also implied, but not demonstrated. Regional contextualization, technology use (current and new), stakeholders’ involvement and efficient organizational structures are among the key factors contributing to a successful wood mobilization intervention. Ultimately, the authors concluded that successful wood mobilization interventions “are tailored to local context, multifaceted and integrated”.

In Quebec, FPFQ (2014; 2013) identified ideas for wood mobilization that mostly fit well in the typology of Orazio *et al.* (2017). However, the importance of the synergy between public and private forests cannot be fitted into that framework because of the peculiar nature of the forestry regime in Quebec. Quebec’s government adopted, in collaboration with the stakeholders from the private forest, a triennial wood mobilization action plan (MFFP, 2017). The wood mobilization action plan is a strategy in the form of objectives, indicators, targets, and actions. Some of the actions in that plan fit with the framework typology. However, several actions do not.

Table 4 is a synthesis of studies referring to specific measures using the typology from Orazio *et al.* (2017). In brackets is the number of examples or case studies reported when the source was a SLR or a collaborative effort. Following the same typology, Table A1 in the appendix details all the wood measures set by SIMWOOD’s regional pilot project on wood mobilization by Vuillermoz *et al.* (2017).

Table 4: Sources reporting wood mobilization measures following the typology of Orazio *et al.* (2017)

Measures (category)	Study
Activation of forest management and enterprise	de Galembert (2003) FPFQ <i>et al.</i> (2013) FPFQ <i>et al.</i> (2014) Hetsch <i>et al.</i> (2007) Lawrence (2018) [19] MCPFE <i>et al.</i> (2010) Orazio <i>et al.</i> (2017) [8] Stern <i>et al.</i> (2013) Vuillermoz <i>et al.</i> (2017) [14]
Regulatory and legal framework	FPFQ <i>et al.</i> (2014) Hetsch <i>et al.</i> (2007) Lawrence (2018) [5] MCPFE <i>et al.</i> (2010) [2] MFFP (2017) Orazio <i>et al.</i> (2017) [12] Stern <i>et al.</i> (2013)
Finances and material inputs	FPFQ <i>et al.</i> (2013) FPFQ <i>et al.</i> (2014) Hetsch <i>et al.</i> (2007) Lawrence (2018) [4] MCPFE <i>et al.</i> (2010) [3] MFFP (2017) Orazio <i>et al.</i> (2017) [26] Stern <i>et al.</i> (2013) Vuillermoz <i>et al.</i> (2017) [4]
Organization and co-operation	de Galembert (2003) FPFQ <i>et al.</i> (2013) Hetsch <i>et al.</i> (2007) Lawrence (2018) [8] MCPFE <i>et al.</i> (2010) [2] MFFP (2017) Orazio <i>et al.</i> (2017) [32] Rauch (2007) Stern <i>et al.</i> (2013) Vuillermoz <i>et al.</i> (2017) [6]
Knowledge and attitudes	de Galembert (2003) FPFQ <i>et al.</i> (2013) FPFQ <i>et al.</i> (2014) Hetsch <i>et al.</i> (2007) Lawrence (2018) [11] MCPFE <i>et al.</i> (2010) [8] MFFP (2017) Orazio <i>et al.</i> (2017) [44] Stern <i>et al.</i> (2013) Vuillermoz <i>et al.</i> (2017) [21]

3.4 Q4: What is the place of logistics solutions in wood mobilization?

Measures relating to the *activation of forest management and enterprise, organization and co-operation and knowledge and attitudes* are the ones that are most closely related to logistics (Orazio *et al.*, 2017). The

analysis of SIMWOOD’s pilot projects highlighted several findings. First, addressing a single wood mobilization barrier often requires the application of multiple measures. Second, quantifiable mobilization volume gain has only been observed in a few cases. Third, the success of wood mobilization can, in the short term, be measured by other indicators than volume gain (e.g. knowledge gain, road construction and participation in capacity building). Finally, among the measures considered by the project leaders in each region, roughly 50% were of logistical nature (Table A1). This supports the idea that wood mobilization can be addressed with logistical solutions. Table 5 lists all the logistics solutions considered by these projects.

Table 5: Logistics measures used in SIMWOOD’s pilot projects

Capacity building	Efficient service delivery
Harvesting measures	Merger of ownership
Harvesting methods	Collaboration
Improved coordination between harvest and logistics operations	Forest cooperative consolidation
Landscape specific treatments	Joint timber marketing
Securing harvesting contractors	Homogenization of management
Securing harvesting costs	Forest owners database
Biomass contracts	Decision support systems
Biomass sales	Technical and scientific tools
Firewood production	Procurement simulations

3.5 Q5: Could logistics measures increase wood mobilization from Quebec’s private forests?

Malovrh *et al.* (2017) suggested that better logistics and infrastructure could help increase wood mobilization. Their findings point to the importance of considering forest owners’ profiles in order to implement the most efficient private forest owner’s organizational model for wood mobilization. Considering that there is an increase in timber demand in Quebec and considering that the utilization ratio is lower in private compared to public forests, wood mobilization from private forests is therefore an avenue that should be considered (FPFQ, 2019; MFFP, 2019).

Literature on forest logistics in private forests does exist. Meek (2013) proposed a cartography of Quebec’s private forest value chain and its processes which can be useful to research on logistics solutions. Dorval (2015) looked at transportation in the private forest sector in certain regions of Quebec. The author used an optimization tool to estimate the savings (2%) a reorganization of transport in a collaborative business model could yield. Elsewhere, research has shown that improved transport logistics, through collaboration, could generate significant profit gain or cost reductions in the forest sector (Audy *et al.*, 2012; Frisk *et al.*, 2010). Indeed, in Quebec transport is the second most important timber procurement cost, after harvesting, and it represents from 29% to 35% of total procurement costs (Groupe DDM, 2020). A logistics approach to wood mobilization from Quebec’s private forests then becomes

interesting. While the scientific literature in Quebec has explored factors relating to owners’ motivations in wood mobilization, it has yet to fully explore the possibility to increase wood mobilization through logistics (Côté *et al.*, 2015, 2017). Considering that logistics can be used to address wood mobilization issues in Europe and that there is a significant potential for cost reduction in forest transport, it would seem that logistics measures could potentially increase wood mobilization from Quebec’s private forests.

3.6 Q6: Are the existing wood mobilization frameworks suitable for Quebec’s private forests?

While our results suggest that a logistics approach could increase wood mobilization levels in Quebec, there is still the issue of applying a foreign framework to a local reality. Hetsch *et al.* (2007), Olschewski *et al.* (2019) and Lawrence (2018) argued on the importance of tailoring wood mobilization solutions to regional contexts. They highlighted the desirability of long-term business relations, planning security and supply channel development. Sjølie *et al.* (2019) suggested that while timber price shifts did not influence harvesting decisions for historically inactive forest owners, it did for historically active forest owners. The findings suggest that different mobilization measures can be successful in different environments.

Hetsch *et al.* (2007) proposed a representation of wood supply bottleneck in the form of a triad, where the first pole is the accessibility of forest, infrastructure and information. The second pole is the usability of resources, and the third pole is the availability of the forest based on the owners’ motivation and objectives. This conceptualization facilitates the understanding of the state of wood mobilization from one region to another. It is helpful in a wood mobilization process because it facilitates the identification of measures that would be promising in a specific region through the comparison of wood mobilization contexts. However, we believe that it requires some adaptations to fit Quebec’s private forest context as 2 out of 3 poles do not adequately represent regional stakes. FPFQ (2014) proposed its own framework to describe the conditions needed to mobilize wood from Quebec’s private forests. The framework is based on a wheel that is initiated with the owners’ personal motivation as a prerequisite and that ends with the decision to mobilize wood as the outcome. Financial opportunities, favourable regulations, fiscality, and technical support are the other factors. While more complete and more adapted to the context of Quebec’s private forests than the framework proposed by Hetsch *et al.* (2007), it suggests a more rigid conceptualization of wood mobilization. It assumes that owners’ personal motivation is a prerequisite. While it certainly is a crucial factor, owners’ personal motivation may also be affected by other factors. First, even though Côté *et al.* (2015) reported that most of Quebec’s private forest owners do not depend on their woodlot as a primary or as a secondary source of income, the authors also reported that 65% of owners would undertake more

forestry work if timber price rose. The findings of Sjølie *et al.* (2019) mirror that idea. Improved logistics has the potential to increase wood mobilization efficiency, to diminish its cost and therefore to increase the owners' benefit margin and their willingness to mobilize wood. Second, logistics measures could help simplify wood mobilization and thus positively affect owner motivation. For example, an effective communication approach and the use of an information diffusion platform that would accurately inform owners on contractors' availability, subsidy possibilities, transport schedules, prices and expected profits could influence the decision of inactive owners to mobilize wood (Orazio *et al.* 2017). Third, wood mobilization can help private owners in the non-harvesting use of their land. Indeed, having better access to their woodlot through road building, managing windthrows and limiting tree pathologies are reported benefits that could influence a forest owner's decision to mobilize wood. Having science-based information can help forest owners mobilize wood when most convenient for them. While the framework of FPFQ (2014) is suitable, one that explicitly accounts for logistics would be more adapted to our research goals.

4 AN ADAPTED FRAMEWORK

Two recent studies lead us to believe that logistics solutions can contribute to an increase in wood mobilization: Malovrh *et al.* (2017), which highlighted the potential of logistics, and Sjølie *et al.* (2019) which indicated that harvesting behaviours could be influenced by timber price. Considering what our SLR showed, wood mobilization of Quebec's private forests through logistics solutions is an avenue that warrants further attention. As Hetsch *et al.* (2007), Lawrence (2018) and Olschewski *et al.* (2019) suggested that successful wood mobilization initiatives should be tailored to regional particularities, we have to develop a framework specific to Quebec's private forests. We are using the conceptual representation of wood supply bottleneck by Hetsch *et al.* (2007) and adapting it to the context of Quebec. However, instead of representing wood supply bottlenecks, our framework highlights wood mobilization desirability (Figure 1).

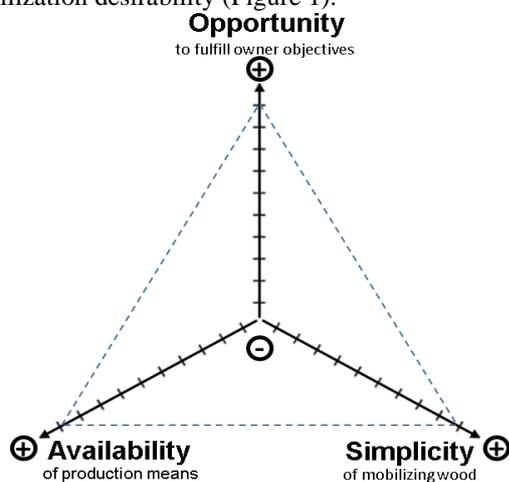


Figure 1: Wood mobilization desirability

By doing so, we consider owners' motivation without making it a prerequisite, but rather an outcome. The first pole of the triad is the opportunity for wood mobilization to fulfill the owner's objectives, e.g. generating incomes, land access, hunting, etc. The second pole is the simplicity of mobilizing wood, *i.e.* regulations, administration and silviculture, etc. The third pole is the availability of production means, *e.g.* mills, machinery, expertise, etc. Wood mobilization is most desirable when the opportunity to fulfill owner objectives exists, when wood mobilization is simple and when production means are available. On the other hand, wood mobilization becomes less desirable when one or more of the poles are affected by some barriers. Logistics solutions have the potential to affect all three poles.

In practice, wood mobilization desirability varies from one region to another because of specific barriers affecting them. In Figure 1, the origin of the axes indicates the lowest prevalence of one or more poles whereas the ends indicate their highest prevalence. The largest the area covered by the triangle defined by the dotted lines, better are the conditions favoring wood mobilization. Thus, the proposed framework can be used to describe wood mobilization contexts, to compare them between regions and to identify measures for wood mobilization based on regional similarities (Table 6).

Table 6: Classification of wood mobilization contexts derived from the proposed framework

Wood mobilization context	Pole(s) affected
Wood mobilization is simple	0
Wood mobilization is somewhat difficult and requires some effort	1
Wood mobilization is difficult and requires considerable effort	2
Wood mobilization is very difficult and requires tremendous effort	3

5 CONCLUSION

The goal of this paper was to identify ways and means that could lead to increased wood mobilization from private forests. The methodology chosen was a systematic literature review. Findings from the SLR suggest (i) that wood mobilization initiatives, mostly from the SIMWOOD program and other European endeavours, can be successful if tailored to the regional context, (ii) that they may produce positive results, though not always quantifiable, (iii) that the literature on successful and quantified wood mobilization initiatives is not extensive and (iv) that wood mobilization through logistics means could yield favourable results. A framework on wood mobilization desirability is proposed in order to accurately contextualize future wood mobilization initiatives in the province of Quebec, Canada.

Future research should aim to develop the applicability of the proposed framework. First, a multicriteria

approach for each of the poles of wood mobilization desirability would help quantify them. Second, this enhanced framework could be used to analyze wood mobilization requirements for specific regions. Afterwards, a comparison between the data from a SLR on wood mobilization and the data from interviews with stakeholders should allow to identify the most interesting wood mobilization measures for a particular wood mobilization contexts. Finally, Vuillermoz *et al.* (2017) put an emphasis on stakeholder engagement as a condition for the success of wood mobilization initiatives. It would therefore be interesting to look at the factors that affect stakeholders' engagement. One such

factor is financial incentives. It would be useful to know at what levels stakeholders engagement could be triggered. This could prove useful in contexts where stakeholders may have conflicting or competing objectives.

7 REFERENCES

6 ACKNOWLEDGMENTS

The authors would like to thank and acknowledge FORAC research consortium, Fonds de Recherche du Québec – Nature et technologies and Fonds de recherche et de développement en foresterie for their financial support in this project.

- Audy, J.-F., & D'Amours, S. (2008). Impact of Benefit Sharing Among Companies in the Implantation of a Collaborative Transportation System - An Application in the Furniture Industry. In L. M. Camarinha-Matos & W. Picard (Eds.), *Pervasive Collaborative Networks* (pp. 519–532).
- Audy, J.-F., D'Amours, S., & Rönnqvist, M. (2012). An empirical study on coalition formation and cost/savings allocation. *International Journal of Production Economics*, 136(1), 13–27.
- Audy, J.-F., D'Amours, S., & Rousseau, L.-M. (2011). Cost allocation in the establishment of a collaborative transportation agreement-an application in the furniture industry. *Journal of the Operational Research Society*, 62(6), 960–970.
- Côté, M.-A., Généreux-Tremblay, A., Gilbert, D., & Gélinas, N. (2017). Comparing the profiles, objectives and behaviours of new and longstanding non-industrial private forest owners in Quebec, Canada. *Forest Policy and Economics*, 78, 116–121.
- Côté, M.-A., Gilbert, D., & Nadeau, S. (2015). Forest Policy and Economics Characterizing the pro fi les , motivations and behaviour of Quebec ' s forest owners. *Forest Policy and Economics*, 59, 83–90.
- de Galember, B. (2003). Wood supply for the growing European pulp and paper industry. *Strategies for the Sound Use of Wood*, March, 437.
- Dengg, J. G. (2003). Wood promotion in Germany – a joint initiative of forestry and the wood based industry since 1990. *Strategies for the Sound Use of Wood*, 437.
- Dorval, J.-S. (2015). Réingénierie du transport forestier: Cas d'étude et modèles collaboratifs. Université Laval.
- Durocher, C., Thiffault, E., Achim, A., Auty, D., & Barrette, J. (2019). Untapped volume of surplus forest growth as feedstock for bioenergy. *Biomass and Bioenergy*, 120, 376–386.
- Fédération des producteurs forestiers du Québec. (2014). Optimiser la production et la mobilisation de bois dans les forêts privées québécoises. 32.
- Fédération des producteurs forestiers du Québec. (2019). La forêt privée chiffrée. 32.
- Fédération des producteurs forestiers du Québec, Regroupement des sociétés d'aménagement forestier du Québec, & Association des entrepreneurs en travaux sylvicoles du Québec. (2013). Enjeux liés à la production de bois dans les forêts privées québécoises. 12.
- Frisk, M., Göthe-lundgren, M., Jörnsten, K., & Rönnqvist, M. (2010). Cost allocation in collaborative forest transportation. *European Journal of Operational Research*, 205(2), 448–458.
- Groupe DDM. (2020). Enquête sur les couts d'opération forestière dans les forêts du domaine de l'État ainsi que sur les couts et revenus de l'industrie du sciage du Québec 2019. Presented to Ministère Des Forêts, de La Faune et Des Parcs, 23.
- Hetsch, S., United Nations Economic Commission for Europe, & Food and Agriculture Organization of the United Nations. (2007). Mobilizing Wood Resources: Can Europe's forests satisfy the increasing demand for raw material and energy under Sustainable Forest Management? 1, 124.
- Karppinen, H. (1998). Values and Objectives of Non-industrial Private Forest Owners in Finland. *Silva Fennica*, 32(1), 43–59.
- Kuuluvainen, J., Karppinen, H., & Ovaskainen, V. (1996). Landowner Objectives and Nonindustrial Private Timber Supply. *Forest Science*, 42(3), 300–309.
- Lawrence, A. (2018). Do interventions to mobilize wood lead to wood mobilization? A critical review of the links between policy aims and private forest owners' behaviour. *Forestry*, 91(4), 401–418.
- Lönnstedt, L. (1989). Goals and cutting decisions of private small forest owners. *Scandinavian Journal of Forest Research*, 4, 259–265.
- Malovrh, Š. P., Kumer, P., Glavonjić, P., Nonić, D., Nedeljković, J., Kisin, B., & Avdibegović, M. (2017). Different organizational models of private forest owners as a possibility to increase wood mobilization in Slovenia and Serbia. *Croatian Journal of Forest Engineering*, 38(1), 127–140.
- Meek, P. (2013). Cartographie de chaînes de valeur et Innovation en forêt privée. *FPInnovations*, 25.
- Ministère des Forêts de la Faunes et des Parcs. (2017). Plan d'action national sur la mobilisation des bois en forêt privée – 2016-2019. 17.
- Ministère des Forêts de la Faunes et des Parcs. (2019). Ressources et industries forestières du Québec, portrait statistique 2019. 160.
- Ministerial Conference on the protection of Forests in Europe, European Commission, United Nations Economic Commission for Europe, & Food and Agriculture Organization of the United Nations. (2010). Good Practice Guidance on Sustainable Mobilisation of Wood

in Europe. 75.

Olschewski, R., Tzanova, P., Thees, O., & Polosek, P. (2019). How does wood mobilization depend on marketing decisions? A country comparison based on choice experiments. *Annals of Forest Science*, 76(4).

Orazio, C., Kies, U., Edwards, D., Cordero Montoya, R., Lovrić, N., Hayes, S., Green, T., Schreiber, R., Koch, M., Dhubain, A., Kerr, G., Schelhaas, M.-J., Nabuurs, G.-J., Lawrence, A., Tidey, P., Olivar, J., Sabin, P., & Kitching, A. (2017). Handbook for wood mobilisation in Europe: Measures for increasing wood supply from sustainably managed forests.

Prins, K., United Nations Economic Commission for Europe, & Food and Agriculture Organization of the United Nations. (2015). Forests in the ECE region : trends and challenges in achieving the global objectives on forests. 212.

Rauch, P., & Gronalt, M. (2005). Evaluating organisational designs in the forestry wood supply chain to support forest owners' cooperations. *Small-Scale Forest Economics, Management and Policy*, 4(1), 53–68.

Röder, M., Thiffault, E., Martínez-Alonso, C., Senez-Gagnon, F., Paradis, L., & Thornley, P. (2019). Understanding the timing and variation of greenhouse gas emissions of forest bioenergy systems. *Biomass and Bioenergy*, 121, 99–114.

Schmithüsen, F., Hirsch, F., & United Nations. Economic Commission for Europe/Food and Agriculture Organization of the United Nations. Forestry and Timber Section. (2010). Private forest ownership in Europe. Geneva: United Nations. 120.

Sjølie, H. K., Wangen, K. R., Lindstad, B. H., & Solberg, B. (2019). The importance of timber prices and other factors for harvest increase among non-industrial private forest owners. *Canadian Journal of Forest Research*, 49(5), 543–552.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14, 207–222.

United Nations Economic Commission for Europe, & Food and Agriculture Organization of the United Nations. (2010). The European forest sector outlook study II: 2010-2030. 107.

United Nations Economic Commission for Europe, & Food and Agriculture Organization of the United Nations. (2012). The North American forest sector outlook study : 2006-2030. 65.

Vuillermoz, M., Edwards, D., Ambrose-Oji, B., Kerr, G., Ni Dhubhain, A., Aurenhammer, P., Azevedo, J., Kelly, A., Ruch, Philippe Koch, Marc Calvert, A., Kies, U., Kitching, A., Richardson, W., Little, D., Pérez-rodríguez, F., Ginet, C., Peuch, D., Magaud, P., Perinot, C., Vayreda, J., & Carbonell, X. (2017). D 4.3. – Regional Pilot Projects's Added Value for Wood Mobilisation - Synthesis Report. SIMWOOD. 70.

8 APPENDIX

Table A1: Listing of wood mobilization barriers and barriers and measures identified by the pilot projects in Vuillermoz et al. (2017)

Barriers		
Soil Sensitivity	Land fragmentation	Lack of information
Terrain steepness	Land size	Lack of knowledge of owners
Cost of management	Lack of communication between stakeholders	Lack of skills of owners
Operation security	Lack of coordination between owners	Importance of non-productive forest functions
Stand unproductiveness	Resistance to co-operation	Lack of interest
Administration complexity	Scarcity of contracting parameters	Lack of owner awareness
Lack of infrastructure	Lack of market	Low forest owner activity
Lack of human resources	Lack of demand for local products	Unmanaged private forests
Land accessibility	Low timber price	Lack of management tools
Measures		
Capacity building	Wood storage area construction and maintenance	Field demonstration events
Harvesting measures	Added value to products	Field training events
Harvesting methods	Demand creation	Focus study
Improved coordination between harvest and logistics operations	Merger of ownership	Contact between stakeholders
Organizational efficiency of harvest operations	Readjustment of land property	Information diffusion
Landscape specific treatments	Collaboration	Regional learning labs
Management intensity	Forest cooperative consolidation	Stakeholder consultation
Securing harvesting contractors	Joint timber marketing	Third party events
Securing harvesting costs	Homogenization of management	Workshop and training sessions
Silviculture	Forest owner mobilization	Forest owners database
Biomass contracts	Counselling	Decision support system
Biomass sale	Information campaign	Technical and scientific tools
Firewood production	Capacity building: individuals	Procurement simulations
Efficient service delivery	Contact with owners	Self-formation documentation
Road construction and maintenance	Publication diffusion	

Automation of competences extraction

HumaNeTrust: an experimental computer system used in an academic laboratory

OLIVIER POUYDEBAT¹, CHRISTOPHE COURTIN¹, KAMEL-EDDIN BEMMAMI¹,
ELODIE GARDET², GEORGES HABCHI¹¹ Univ. Savoie Mont Blanc, SYMME
5 Chemin de Bellevue - 74940 Annecy-le-Vieux
olivier.pouydebat@univ-smb.fr, christophe.courtin@univ-smb.fr,
kamel-eddin.bemmami@univ-smb.fr, georges.habchi@univ-smb.fr² Univ. Savoie Mont Blanc, IREG
4, Chemin de Bellevue - BP 80439 - 74944 Annecy-le-Vieux Cedex
elodie.gardet@univ-smb.fr

Abstract

In an organisation, competences form human capital, an intangible asset that is decisive for overall performance and strategic management. The aim of this paper is to understand, in an objective and undeclared way, how competences in an organisation evolve. Then this research proposes to fill this gap by setting up a computer system to provide an automated overview of competences (individual and collective). After defining the competences, we show that traces can be used to identify them. Based on the PROXYMA model, we switch to an actual dimension for the first time through the implementation of an application (HNT: HumaNeTrust) that allows to validate the different functions of a trace-based system. Then, the HNT application is experimented in an organisation, based on of the traces produced during publication of scientific articles. The results show that the graphs produced can be used by the organisation's managers to obtain an objective overview of the professional interpersonal relationships and areas of expertise held within the organisation (which are some of the facets of the complex concept of competence). We notice that the areas of expertise found with HumaNeTrust software are not exactly the same as those highlighted on the organisation's website. This research marks a step in the quest for an automatic competence mapping solution, to provide the structure's managers with an objective view of reality.

Keywords – Competences mapping, Collective and Individual competences, Human information, Knowledge management, Skills, Competencies, Digital interaction, Trace-based system, TBS, HAL.

1 INTRODUCTION

One of the major components of a company's value is the human capital that is a part of its immaterial capital [Boucher et al., 2007]. The strategic management of a company implies the management of this human capital that comprises the individuals' information. A clear and undistorted view of this information is crucial for the leaders of a company.

The use of software tools for strategic management could facilitate its extraction [Balog et al., 2012] especially in a turbulent environment forcing the different actors to adapt and develop their competences. For managers, the challenge is to be able to identify quickly and objectively the internal competences that can be mobilized for a project development. In addition, the competences mapping can help managers to reveal the competences that are missing or those that need to be strengthened (for example, through training or hiring). Beyond their usefulness for managers, individual and collective competences are essential to establish the strategic diagnosis of a company and thus establish the main orientations.

Despite the various definitions of competences, due to the polysemic nature of the term, most researchers agree that they are closely related to the activities [Aubret et al., 2005; Bernabé-Moreno et al., 2019; Gluga et al., 2011; Lingappa et

al., 2016; Ughetto and Paris-est, 2002] and include multiple facets in terms of knowledge, know-how and skills [Durand, 2015].

Like the mainstream [Kandampully, 2002], we consider that competences being human dependent and temporally situated imply trust level and regular update issues, respectively. The difficulty then lies in successfully operationalizing this vision of competences. One of the proposals often used, is to base the analyses on the competences declared by the individuals themselves. However, some researchers believe that a distortion can exist [De Siqueira and Gough, 2016], because there is a gap between declared and real competences.

To fill this gap, we propose to get an automated inventory of elements of competences by means of the publications activities' analyses. Most studies based on publications analysis are conducted to estimate the impact factor of publications [van Leeuwen and Wouters, 2017], or to specify parameters influencing this impact factor [Sahoo et al., 2018]. As our objective is concerned, it consists in getting this automated inventory through publishing activities' analyses mainly from the HAL (Hyper Articles on Line: <https://api.archives-ouvertes.fr>) platform. These analyses are carried out in a Trace-Based System (TBS) and involve an evaluation of the level of reliability of the used instrumentation technique [Courtin, 2008].

The two main concepts of this study are interactive digital traces and competences. This work is based on the experience acquired from previous research work on a TBS called BEATCORP [Chebil et al., 2011], where we observed the relevance of low level traces for identifying potential individual competences [Chebil et al., 2015]. In this study, we collect traces at the task level to underline the link between tasks and competences as presented later. Therefore, we have implemented a software, based on the PROXYMA model [Chebil et al., 2015], to collect and organize higher level traces in order to identify competences.

The underlying question is to know whether it is possible to extract individual or collective competences from a trace-based system. This work thus offers both academic and practical value to an automated vision of competences. First, the concepts associate the definition of competences with the main dimensions of characterization. Second, the notion of trace is explained and how it can be useful to track competences. Then on these theoretical bases, our project is described (the HumNeTrust software). The main objective is to collect, organize, identify and map competences from traces of activities. Thanks to the first acquired results, we present a first mapping of competences. After the results presentation, we conclude with an in-depth discussion.

2 COMPETENCE

A huge volume of research works were dealing with the individual competences until the last decade, producing an impressive number of competence definitions. Due to the polysemy of this term, there is no consensual and universally accepted definition [Aubret et al., 2005; Lingappa et al., 2016; Ughetto and Paris-est, 2002]. Before suggesting a definition, we present a picture of the main common features that emerge from this abundance of definitions.

The scientific terminology highlights different meanings for the following concepts: skill, competency and competence [Fazel-Zarandi and Fox, 2010; Le Deist and Winterton, 2005]. The people competencies designate the cause and effect relationship between the individual characteristics, observed during a high yield activity, and their successful and effective job performance. The people competences highlight the link between a complex set of knowledge, skills, motivation, attitude, etc. and their job fits within a specific work context. In short, "competence" describes what people can do, while "competency" focuses on how they do it [Le Deist and Winterton, 2005; Sanghi, 2016].

In accordance with the above-mentioned definition of competence, we identify three dimensions in the set of elements: "Knowledge", "Know how to do" and "Know how to be" [Durand, 2015; Le Deist and Winterton, 2005].

This paper deals with the concept of competence which involves the above-mentioned complex items to carry out specified actions in a specific context to achieve a S.M.A.R.T. (Specific, Measurable, Achievable, Relevant and Time bound) objective [Menon, Saraswathi and Karl, Judith and Wignaraja, 2009]. According to this definition, this section tries to specify the meaning of the following terms:

Knowledge: an information acquired during learning activities or/and a set of facts, principles, etc. acquired through human experience and thought. It may be explicit or tacit [Nonaka, 1995]. Individual's expertise fields are those areas in which the person has in-depth knowledge and for which the person is socially recognized (reputation, recommendation, publications)

as a reference person.

Skills: a specific aptitude for a task, sport, etc., the ability acquired through training and/or practical activities [Durand, 2015; Le Boterf, 1994; Le Deist and Winterton, 2005].

Attitudes: a personal set of features that lead to a particular behavior in various work situations [Durand, 2015].

Action: concrete manifestation of someone's activity [Durand, 2015].

Context: a set of elements related to human aspects, financial aspects, time, material resources, etc., which define a specific situation [Brownell, 2008].

The competences are revealed through action and recognized by the others. Many studies indicate that competences can only be approved if they are socially recognized. The action and the social recognition are thus both essential aspects to address the three dimensions of the competences. People competences emerge thanks to a good implementation of their own knowledge in the action [Durand, 2015; Le Boterf, 1994].

The number of skill classifications being high, we present below an excerpt reduced to the most common types found in the literature [Durand, 2015]:

- Individual competences: an individual's own skills. Individuals generally include these competences in his/her job description, his/her curriculum vitae or his/her LinkedIn© profile for example. The subtypes listed below are used commonly to distinguish individual's competences. Studies [Belkadi et al., 2007] strives to synthesize the publications on individual competences as following :
 - Hard skills: demonstrable skills attested by the diplomas, certificates, degrees, titles ...
 - Soft skills: knowing how to collaborate, how to adapt to a situation, how to negotiate ...
 - Mad skills: atypical skills, acquired through the practice of a sporting or creative activity, or an original passion ...
 - Collective competences: they are acquired by the quality of the interactions that take place between the individuals in a company. Collective competences are made up of two complementary parts, one referring to the individual and another one to the organization. Any modification of one of these parts modifies the collective competence itself [Boreham, 2004; Le Boterf, 2000; Macke et al., 2013]. Valérie Michaud [Michaux, 2005] provides a synthesis of the types of collective competences dynamics : i) Collective competences in the perspective of pooling, synergy effect. (project team, operational team), ii) Collective competences in the perspective of change (collective learning, adoption of a new technology, creation of new knowledge), iii) Collective competences in the perspective of sharing (free or guided) experiences. (knowledge sharing, reflection group, progress group, etc.), iv) Collective competences in the perspective of daily management of hazards: cooperation, transversal communication, inter-individual exchanges within a network.
- Organizational competences: a set of competences that allow performing actions whose objective is to

successfully accomplish the organized actions, but also to influence the way in which these actions are organized. These competences allow to organize the framework of action, and allow the evolution and adaptation of it to new constraints [Grant, 1991; Michaux, 2005].

Business owners, and more specifically in the human resource department, collect and organize this information according to an internal skills repository. Such repositories have inventory values because they contain the description of the positions and the jobs. These repositories become referentials in the companies and they are now the subject of many studies [Jarnias and Oiry, 2013]. The ESCO survey accessible online provides a multilingual repository of individual skills in the form of an updated and effective taxonomy for the European labor market (ESCO is the multilingual classification of European Skills, Competences, Qualifications and Occupations: <https://ec.europa.eu/esco/>). The use of referentials is constantly increasing and a common frame of references with a shared representation can be defined by means of consensus or even negotiations. Such a representation would enable to identify both the activities to be carried out and thus the corresponding actions to do and the procedures to be applied to reach the objectives. This information can be input into software tools to assist the human resource managers during staff appraisal or the leaders to establish a strategic management.

As noted above, the holding of competences is evidenced through actions. Research and analysis of the traces of these actions will serve as evidence of possession of competences.

Competence having been defined, it now appears as a complex and composite concept, with a very varied typology. In light of this complexity, it seems clearer that managers in organizations need tools to help them identify the elements that make up the competences present in their structures. A competence extraction tool will be able to detect each of the facets that make up competences. It will provide a limpid, clear vision that is as exhaustive as possible.

3 TRACE

As it was mentioned in the previous section, competences may be identified through activities containing functional and business tasks. When executed in a Digital Working Environment (DWE), the corresponding actions leave digital interaction traces that can be analyzed. During the last two decades, there has been an increasing interest in analyzing activities especially in the learning management systems (LMS) in the form of observable objects. In the scientific literature, many research works deal with the notion of trace, but there is no standard definition to this day [Settouti et al., 2009]. As defined by [Settouti et al., 2009], traces are considered as a collection of temporally located observables which are any structured information stemming from the observation of human-machine interaction (Observable: structured information derived from observation of an interaction). These interaction traces are collected in a trace-based system (TBS) to be transformed into higher abstraction levels traces to identify the corresponding tasks listed in referentials. The functional and business tasks imply soft and hard competences related to behavior and application fields, respectively. These competences are listed in referentials similarly and may be thus identified through the actions of individuals or groups of users. Through the ESCO European referential, it is thus possible to identify the competences and

attitudes that are linked to the specific tasks. In conclusion, the interaction traces provide information on two of the three dimensions of competences: know how to do and know how to be. In terms of knowledge identification, [Barazzutti et al., 2015] presents an interactive tool, named "Transmut", to support knowledge discovery (KD). This tool enables the extraction of knowledge by means of traces processing following a cycle composed of three main phases. Firstly, traces are collected, selected, and transformed. Secondly, an algorithm is used to detect frequent sequential episodes considering the constraints specified by the analyst. Thirdly, the results of the operation are transmitted to the analyst who chooses the most relevant models according to his knowledge of the domain. In terms of competences and attitudes identification, there is no significant proposal of studies focusing on the exploitation of interactive traces. Nevertheless, some attitudes (soft competences) can be measured by the traces collected from different actions of user activity. H.Tadger's work on PBL (Problem Based Learning) environments for the improvement of students' soft competences on the basis of activity traces, shows that it is indeed possible to measure some soft competences (e.g. in communication, in time management, initiative, curiosity) [Tadger et al., 2020]. The system developed in H.Tadger's work under the name "SYDIS" (SYstem of Development and Improvement of Skills) allows to measure through traces (traces of direct actions; traces of interaction with the system) the level of these soft competences and their evolution, as demonstrated in the experimental study of the SYDIS.

As shown above, traces stemming from actions may be used to identify the three dimensions of competences. Through the ESCO standard, it will be possible to identify the competences and attitudes that are linked to the identified tasks. The basis of our approach lies on the exploitation of these traces.

The UML model presented figure 1 enables to describe interaction traces with various abstraction levels based on various granularity levels [Chebil et al., 2015]. A signal corresponds to an elementary action identified by a single event, such as save a page in a word processing software. A sequence has some duration and is composed of a chronological succession of signals and other sequences. These traces are stored in the database of a TBS based on the CARTE model (Collection, activity Analysis and Regulation based on Traces Enriched) [Courtin, 2008] and are exploited with several application programming interfaces (API) to be collected, transformed, queried, notified and displayed. The analysis of these traces needs the description of the context of the activities where these actions take place. These contextualized traces are organized in corpora containing context metadata and other resources of the corresponding activities, such as provided or produced documents. In this study, these corpora are managed with the BEATCORP system (Benchmarking platform for Analysis of Trace CORPora) based on the PROXYMA approach [Chebil et al., 2015] whose originality is to exploit traces regardless of their original format. Their description is not the subject of this paper, but the TBS presented below is based on the BEATCORP architecture.

Thus, with the information mined behind traces, we can use a competence model, to evaluate users' competences on some subjects. N.Wang [Wang et al., 2016] proposes "Action-Knowledge model" of competence permitting the evaluation of users competences (eg. Java programming; PHP; C#). This evaluation is based on an indicator of competence ($C\{i,j\}$)

denoting the competence of user j on subject i , calculated according to a statistical method used for the research of information in textual documents.

We can also use a competence model to extract company competences in a context of decision-aid to identify collaborative networks. A.Harb presented an approach of ontology enrichment for competence extraction implemented through the developed system "UNICOMP" [Harb et al., 2011] to identify collaborative networks. This approach focuses on extracting essential facets of competences within the company by analyzing traces of competences on its website.

The following section proposes a first experimentation based on the concept of traces and competences. Through this experimentation, we develop a trace-based system to identify the researchers competences in a laboratory.

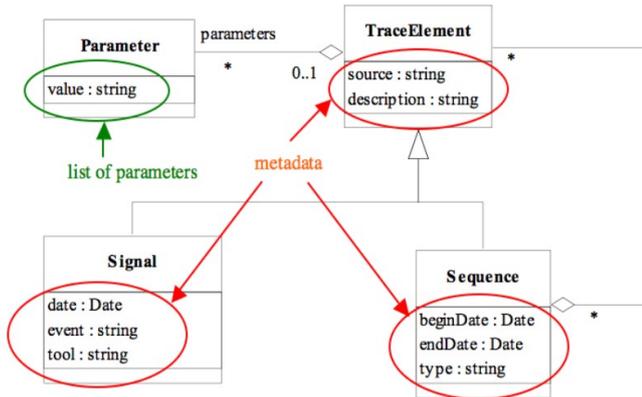


Figure 1 : Interaction Trace Model

4 FIRST EXPERIMENT : HUMANETRUST

Our objective was dual: (i) to equip an organization with a competence mapping tool (ii) to implement a solution according to the concepts specified in the TBS BEATCORP. The HumaNeTrust (HNT) prototype was implemented in PHP and javascript with a MariaDB relational database. We chose to study the actions of publication of articles for the researchers of a laboratory. It was a very instructive phase because it allowed us identifying the effective difficulties, and especially the limits of this first proposal. This is what it will be detailed below.

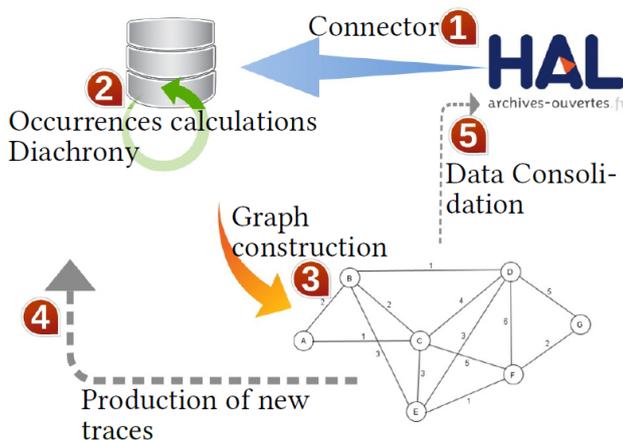


Figure 2 : HumaNeTrust process principle

The five main functions of HNT (Figure 2) are the following:

- Step 1: data recovery from HAL (a French repository

dedicated, among others, to research articles). A connector is in charge of retrieving the state of the data in HAL, at regular intervals, and storing this raw data in the local database.

- Step 2: the raw data are organized in such a way as to calculate the occurrences. (eg. Number of times a researcher has published using a given keyword. Number of times two researchers A and B have co-published.)
- Step 3: production of a graph illustrating these organized data. Users can interact with the graph by clicking on a node or edge. The user also has a search engine to find hidden terms in the graph.
- Step 4: the user can complete the displayed data with additional information that he can input. This information can be attached to a node or an edge of the graph.
- Step 5: this first experiment had the effect of highlighting some of the inconsistencies present in the HAL data source. The researchers themselves went to check the source and made corrections. It is a process of data consolidation.

The main functions of HNT being described; here are the details of the implementation.

4.1 Data acquisition

The HAL Restfull API (<https://api.archives-ouvertes.fr/>) is used to set up the functions that are specific to the connection to the data source (Mark N°1 in Figure 3). At this stage, data are stored as a time-stamped local buffer, in the original unmodified JSON format.

4.2 Data processing

The second program (HAL_DataCompilation.php) loads a raw JSON file, as retrieved from HAL and stored in the database by the first program. Then the program fills a set of relational database tables so that we can make powerful SQL queries (Mark N°2 in Figure 3). Three tables store organized source data: the first for all authors, the second for all publications and the third for all keywords. The calculated occurrences are stored in three specific tables: the first contains the co-publications occurrences, the second contains keywords occurrences used by authors, and the third contains the occurrences of co-use between two keywords.

4.3 Graphs processing

In a third program, data sets are built with these pre-calculated data to be displayed in the form of graphs. (Mark N°3 in Figure 3). There are two types of graphs in JSON format:

Graphs whose nodes are authors: a node represents a person. The diameter of the node carries the information of the number of times the author has co-published. The edges of this graph indicate co-publication between two authors. The thickness of the edge is proportional to the number of co-publications between two authors.

Graphs whose nodes are keywords: a node represents a keyword. The diameter of the node carries information on the number of times the keyword has been used in the laboratory's publications. The edges of this graph indicate the use of two keywords in co-publication. The thickness of the edge is proportional to the number of times a keyword has been used at the same time as another keyword in a publication. The graphs can contain a large number of nodes and edges that can degrade the fluidity of the consultation interface, the

readability of the entities and the general ergonomy of the tool. Some functionalities enable thus to compute graphs with all the nodes representing data stemming from HAL queries or only a part of them.

4.4 Data viewing

Several graph sets are stored in JSON format and the users may load one of them on their Web browser by a forth program (Mark N°4 in Figure 3). They can explore the interactive displayed graph by selecting nodes or edges to obtain their properties. A search engine allows the user to look for an information in the graph, and the resulting graphic entity can be emphasized by a zoom functionality. We use a library to spatialize the nodes according to the occurrences carried by the edges. Higher is the occurrence, higher is the attraction force between two nodes. Initially, we choose arbitrary to distribute evenly the nodes on a circle. Then, the spatialization program calculates the relative positions of the nodes according to the values of occurrences carried by the edges between these nodes. There are various methods of spatialization [Cogis and Schwartz, 2018; Thomas, 2015], and their limits are shown in the table below:

Table 1: Algorithm of spatialization.

Algorithm	Year	Limit
Fruchterman Reingold	1991	maximum 1000 nodes
Yifan Hu	2005	maximum 100000 nodes
Force Atlas 2	2010	maximum 1000000 nodes

The SIGMA-JS (<http://sigmaj.s.org/>) library is compatible with a graph in JSON input format, and enables to implement Force Atlas 2 in javascript on the client side.

4.5 Inputting Data in Networks

In a future development, we will complete the graph interface by functionalities allowing collecting user input, associated with nodes and edges. The objective is to collect three types of new data:

Feedback: This type of input will allow us to create new data, or categorize existing data based on nodes and edges. We will be able to build input forms specific to an information gathering campaign, and record the new data in our database. (Mark N°7 in Figure 3).

Consolidation: When users see an error (eg. in the spelling of their name), they can immediately notify it or even correct it themselves. Providing them with the tools to perform these operations could well be possible from HNT. (Mark N°6 in Figure 3).

Recommendation: It is also planned to offer to the HNT users some tools allowing them to evaluate (eg. from 1 to 5 stars) a level on a graphical entity (node or edge). For example, their level (self-evaluated) of expertise on a specific domain. In addition, they could evaluate the elements of their N+1 level network, if the graph represents their social network, centered on themselves (ego-centric network). (Mark N°5 in Figure 3).

4.6 Overall vision

The solution combines four functionalities: i) Data acquisition, ii) Data processing, iii) Graphs processing, iv) Data watching. Figure 3 shows these different levels of. It is planned to add a JSON data management module to HNT in order to deal with the buffering of raw data. We investigated the Elasticsearch stack (Elastic, Beats, Logstash: <https://www.elastic.co/>) solution which would be appropriate due to:

Its scalability: the possibility to easily increase the storage capacity.

The tools availability and compatibility: in particular Beats and Logstash are very good solutions to fill the Elasticsearch base. It will be able to load a local buffer from various and heterogeneous data sources.

Finally, arrows start from the visualization tool (HNT) or from the future recommendation system (Mark N°5 in Figure 3) and go until the HNT base (Mark N°7 in Figure 3) and the HAL base (Mark N°6 in Figure 3). These arrows represent flows that do not yet exist in an operational way, and whose purpose will be to: i) Consolidate HAL data from inconsistencies highlighted in graphs, ii) Produce new data from the interactions carried out in the graphs (data entry on nodes or edges, suggestion, etc.).

Since July 2019, two graph consultation terminals are in place in the laboratory and can be used by researchers to explore keywords used in publications. For ethical and confidentiality reasons, author's graphs are not displayed publicly. This technical platform will be used as a basis for the first experiments aiming to automate the extraction of competences from traces present in HAL. The graphs produced will be able to be used by the laboratory management to know in an objective way the interpersonal professional relationships and the competences held within the organization.

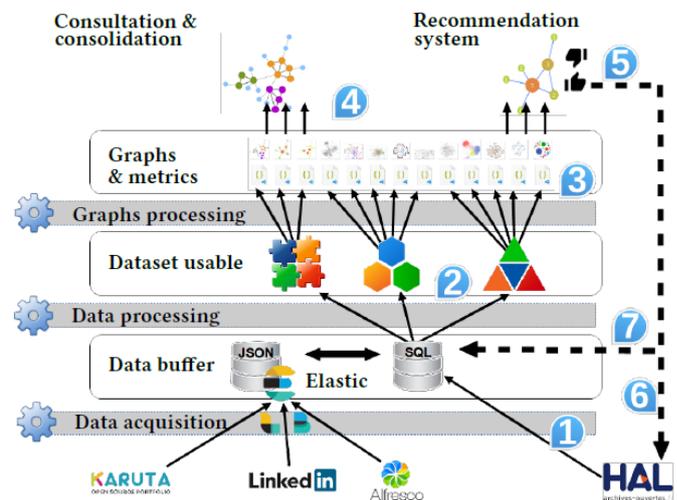


Figure 3: The HumaNeTrust architecture

5 RESULTS: HUMANE TRUST SOFTWARE

Finally, the HumaNeTrust software combines the two main concepts presented in sections 2 and 3 of this work: competences and traces. The results of these concepts are present in HumaNeTrust: figure 4 shows the software, and the architecture has been described in section 4. The used graphs are based on an export of data on 2020/02/14, processing 1209 publications. Table 2 shows the graphs volumetry.

Table 2 : Graph volumetry at 2020/02/14.

	Total publications	Total nodes	Total Edges
keywords graph	1209	1914	9029
authors graph	1209	3902	7556

This implementation is structured according to the Proxyma models.

This application is composed of the four zones:

The main background screen: The graph is displayed with an adjusted zoom factor. The node size changes according to the

value to be represented (number of publications for an author, or number of occurrences for a keyword). The links between nodes (edges) are of constant thickness. However, we can associate to this thickness either the number of co-publications between two authors, or the number of co-uses of two keywords in the publications. It is possible to click on two types of entities in the graph: i) Click on a node: highlights the subgraph of the nearest direct neighbors. This action also updates the contextual zone (left side), ii) Clicking outside the graph, in the background: cancels the current selection and returns to the total selection of the graph. This action also updates the contextual area (on the left side).

A search tool: It is possible to enter directly a term (a keyword or an author's name) and if the item is present in the graph, it will be automatically centered and selected.

A contextual zone: three subsections are present

- **Graph caption:** not finalized, this section will display the graph metadata.
- **Node information:** not finalized, this section will display the properties of the clicked node. Later, it will also be possible to enter data to be associated with the currently selected node.
- **Neighbors of keyword:** this section displays a cloud of keywords. Terms present in the cloud are the keywords of the neighbors connected to the currently selected node.

Navigation tools: global view, zoom in and zoom out. These tools complement the actions available directly in the graph, such as zooming in and out with the mouse wheel, and pan.

5.1 Uses

Two large touchscreens have been installed on two of the laboratory's sites: "point A" and "point B". In both cases, these spaces are strategic places in the laboratory where many researchers pass by, and can take a break to exchange ideas. The produced graphs answer the following questions that researchers ask themselves when discovering the software: What are the most relevant keywords? To which keywords are they related? Which authors use them in the laboratory?

At first glance, researchers are surprised to see the keyword "Simulation" being the most used in the publications. This term largely dominates the others. However, this field of expertise is not highlighted by the laboratory in its communication (website, social networks, etc.). It is nevertheless the field of expertise of the laboratory most shared by researchers and the most cited in their publications.

These three questions are the main ones that generate discussion. Unfortunately, for reasons of respect for individual rights, the graph for answering the third question is not available on public touch screens. However, during demonstrations on private workstations, we were able to observe researchers using the authors' graph. Their first reaction is to look for themselves in the graph. Some authors analyze their personal network, and mentally reconstruct the reasons for the different links. Finally, and this was a desired goal from the beginning of the experiment, some researchers get that they have multiple identifiers in the graph. The beneficial effect of these findings is that the researchers correct the data source (Mark N°5 in Figure 3).

It is planned to anonymize the graph of the authors and they will be represented with a unique identifier. The user can then click on this identifier and send a message to the author. This message states that a user has wished to identify him, and ask

him in return to reveal his identity. Answers to other questions have been requested. For example, which authors publish articles with the most popular keywords? Which authors introduce new keywords?

This last question introduces an important planned evolution: the management of the chronology of appearance of keywords in publications. For this reason, we will have to compute new graphs according to the following sequence:

1. Sort publications by increasing publication date.
2. Identify the emergence of new keywords.
3. Identify the authors of these publications.

The keywords graph can then be coupled to a time slider, which will allow visualizing the growth of the graph progressively as the new keywords appear.

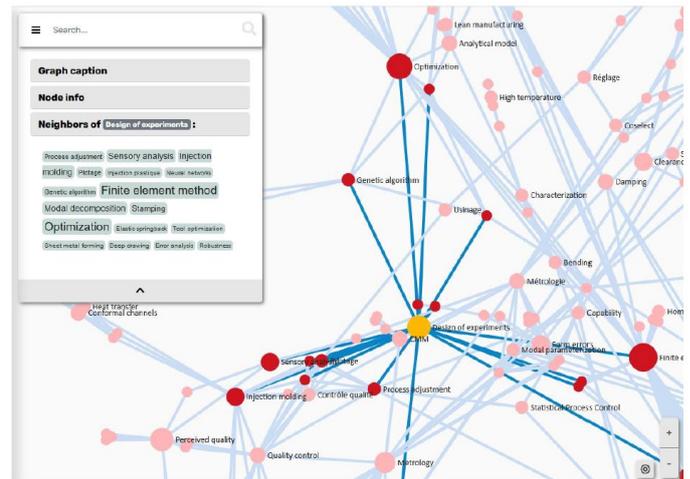


Figure 4: HumaNeTrust user interface. Graph of keywords used in the laboratory's publications, and their co-use links

5.2 Current limitations and prospects

Beyond the storage and processing mechanisms presented, it is important to keep in mind that the goal is to be able to extract a competences mapping from the system.

A critical analysis of the existing system is thus necessary:

Area of expertise and competence: what is the difference between publishing in an area of expertise and having recognized competence in the subject? As we saw in section 2, the field of expertise is one of the facets of competence, included in knowledge, one of the three dimensions that makes up competence (knowledge, expertise, interpersonal skills). Our solution for extracting competences identifies some of the information relating to the structure, but does not offer an exhaustive view of the individual and collective competences actually present. However, this information is valuable because it allows us to deduce all the activities of the researchers, complementary to the publication work around these fields of expertise: teaching, search for partnerships, participation in professional networks, bibliographic research, modelling, English writing, collaborative work, etc. For example, some of these activities address indirectly the expertise facets of competences that are linked to the writing of scientific articles. There is therefore additional work to be done to enrich the competence profile of individuals with other sources of information that it will be possible to cross-reference. On the other hand, it will certainly be useful to calculate and quantify the level of expertise. It would certainly be relevant to compare

these areas of expertise with a notoriety calculation (i.e. H index: <https://arxiv.org/abs/physics/0508025>) on these keywords in order to add the "social recognition" facet of competence.

Genericity: the built system is highly oriented on a type of data (scientific publications or studies). However, it will be easy to reconfigure the system to consider data from an e-mail box, LinkedIn, etc.

Diachony: build graphs provide a static view of the status of the publications. The basic functions are implemented for the dynamic construction of a network on user actions. The literature shows the established interest of the process-oriented competence vision [Bayad et al., 2006]. Chronological graphs will be completed with metrics calculations [Dakiche et al., 2019], and differential vision between two dates.

Going further in graph use: the analysis of the graphs was not carried out as thoroughly as desired. It would be very useful to calculate metrics based on graph theory and more particularly on the theory of strong links and weak links [Saeedi et al., 2018]. Highlighting structural holes and structural bridges, calculating "clustering coefficients" and "connection coefficients", are so many indicators whose relevance is not in doubt. The navigation in the graph, i.e. the process of exploring the graph, proposed by the application is quite basic. A development has been initiated to propose a billiard type navigation [Ventocilla et al., 2017].

Missing references: lot of noise is generated by the missing references.

- People repository: the use of an LDAP directory, coupled with text analysis, shows that duplicates were reduced to more than 98%. This experimental development has not been completed, and has not been integrated into the HNT source code.
- Keywords repository: the back up of HNT with an ontology system [Ilahi et al., 2014] would make it possible to simplify the graphs by making thematic groupings.

A problem of scale and granularity: there are many other competences in a laboratory than those reflected in the publications of the researchers. These competences are not proved, neither in action nor socially. Yet these two elements are inseparable from the definition of competence, as presented in section 2. A suitable system should enable to: i) put actions back at the center of the analyses, ii) increase the variety of observed actions, iii) recover a form of social validation relative to the quality of the implementation of actions.

Through this self-criticism, HumaNeTrust appeared to be an interesting step towards the use of traces to manage the competences of the laboratory. This step was necessary to partially validate the feasibility of the implementation of the PROXYMA models [Chebil et al., 2015].

As we have just underlined in this section, these first results make us want to go further and perfect our base to add functionalities to it. The initial reactions are encouraging, and show us that the needs of structure managers, in terms of competences mapping, are considerable.

6 CONCLUSION AND PERSPECTIVES

In this document, we validate an approach based on traceability to identify automatically the human capital of organizations and partially individual and collective

competences. Based on the academic work that led to the Proxyma model, the HumaNeTrust application was set up and implemented in two public places. Initial feedback indicates that HumaNeTrust offers a new vision of an organization's human capital: competences appear in an objective and factual manner. The work of data collection and organization is well automated, and snapshots can be taken at regular intervals in order to analyze the processes of building or losing competencies.

We have moderated the interpretation of our results, which we certainly consider encouraging, but which we would like to supplement with a number of improvements, which we propose to you in the following descriptions. It is interesting to note that the areas of expertise found with the HumaNeTrust software are not exactly the same as those highlighted on the organization's website. This is a form of evidence that there is a gap between the competences known and displayed by an organization and the competences actually held by its human capital.

We also note that the HumaNeTrust software provides us with an objective view of the interpersonal professional relationships present in the organization's activities. This display opens up ethical questions that we will have to answer with adapted technical solutions.

The work of defining competences proposed in this article shows that knowledge and expertise (extracted from the fields of expertise) are only part of competences. These elements are already rich in information. In the proposed approach, the tracing of actions as well as social recognition are lacking to prove that an individual has acquired competences. Further research will have to be carried out to identify the sources of real competences present in a group of individuals, capable of relying on a wider variety of actions.

This first experience has allowed us to measure the limits of our approach regarding the studied data and the implemented functions in the TBS. These limitations allow us to identify some research questions and the associated technical problems. One of these issues concerns the link between the concepts of trace and competence, both related to activities. In other words, can competence be considered as a high level of abstraction of the trace or can a trace be considered as evidence of competence? Secondly, how can these traces be better used for intelligent competence management in organizations?

7 REFERENCES

- Aubret, J., Gilbert, P., and Pigeure, F. (2005). *Management des compétences: réalisations, concepts, analyses*. Dunod.
- Balog, K., Fang, Y., De Rijke, M., Serdyukov, P., and Si, L. (2012). Expertise retrieval. *Foundations and Trends in Information Retrieval*, 6(2–3), 127–256.
- Barazzutti, P., Cordier, A., Fuchs, B., Barazzutti, P., Cordier, A., Fuchs, B., and Tool, I. (2015). *Transmute: an Interactive Tool for Assisting Knowledge Discovery in Interaction Traces*.
- Bayad, M., Boughattas, Y., and Schmitt, C. (2006). Le métier de l'entrepreneur : le processus d'acquisition de compétences. *8ème Congrès International Francophone En Entrepreneuriat et PME (CIFPME)*, 16.
- Belkadi, F., Bonjour, E., and Dulmet, M. (2007). Competency characterisation by means of work situation modelling. *Computers in Industry*, 58(2), 164–178.
- Bernabé-Moreno, J., Tejada-Lorente, Á., Herce-Zelaya, J.,

- Porcel, C., and Herrera-Viedma, E. (2019). An automatic skills standardization method based on subject expert knowledge extraction and semantic matching. *Procedia Computer Science*, 162, 857–864.
- Boreham, N. (2004). A theory of collective competence: Challenging the neo-liberal individualisation of performance at work. *British Journal of Educational Studies*, 52(1), 5–17.
- Boucher, X., Bonjour, E., and Grabot, B. (2007). Formalisation and use of competencies for industrial performance optimisation: A survey. *Computers in Industry*, 58(2), 98–117.
- Brownell, J. (2008). Leading on land and sea: Competencies and context. *International Journal of Hospitality Management*, 27(2), 137–150.
- Chebil, H., Courtin, C., and Girardot, J.-J. (2015). A case study to validate the PROXYMA approach to share and analyse contextualised interaction trace corpora in a TEL environment. *International Journal of Learning Technology*, 10(4), p.291-312.
- Chebil, H., Courtin, C., and Girardot, J.-J. (2011). BEATCORP: une plate-forme de benchmarking pour l'analyse de corpus de traces. *Conférence Environnements Informatiques Pour l'Apprentissage Humain (EIAH)*, 10.
- Cogis, O., and Schwartz, C. (2018). *Théorie des graphes* (Cassini (ed.); Poche). Collection L.
- Courtin, C. (2008). *CARTE : AN OBSERVATION STATION TO REGULATE ACTIVITY IN A LEARNING CONTEXT*. *†. Celda*, 191–197.
- Dakiche, N., Benbouzid-Si Tayeb, F., Slimani, Y., and Benatchba, K. (2019). Tracking community evolution in social networks: A survey. *Information Processing and Management*, 56(3), 1084–1102.
- De Siqueira, J. R., and Gough, M. J. (2016). Correlation between experience targets and competence for general surgery certification. *British Journal of Surgery*, 103(7), 921–927.
- Durand, T. (2015). L'alchimie de la compétence. *Revue Française de Gestion*, 41(253), 267–295.
- Fazel-Zarandi, M., and Fox, M. S. (2010). Reasoning about skills and competencies. *IFIP Advances in Information and Communication Technology*, 336 AICT, 372–379.
- Gluga, R., Kay, J., Lever, T., and Lister, R. (2011). *An architecture for systematic tracking of skill and competence level progression in Computer Science*.
- Grant, R. M. (1991). The resource-based theory of competitive advantage: implications for strategy formulation. *California Management Review*, 33(3), 114–135.
- Harb, A., Hajlaoui, K., and Boucher, X. (2011). Competence mining for collaborative virtual enterprise. *IFIP Advances in Information and Communication Technology*, 362 AICT, 351–358.
- Ilahi, M., Cheniti-Belcadhi, L., and Braham, R. (2014). Formal competence-based assessment: On closing the gap between academia and industry. *ACM International Conference Proceeding Series*, 581–587.
- Jarnias, S., and Oiry, E. (2013). Vers un repérage des types de référentiels de compétences. *@Grh*, 8(3), 11.
- Kandampully, J. (2002). Innovation as the core competency of a service organisation: The role of technology, knowledge and networks. *European Journal of Innovation Management*, 5(1), 18–26.
- Le Boterf, G. (1994). *De la compétence : essai sur un attracteur étrange*. Editions d'Organisation.
- Le Boterf, G. (2000). *Construire les compétences individuelles et collectives*. Ed. d'Organisation.
- Le Deist, F. D., and Winterton, J. (2005). What Is Competence? *Human Resource Development International*, 8(1), 27–46.
- Lingappa, D. M. S., De Fransisci Morales, G., and Gionis, A. (2016). Extracting skill endorsements from personal communication data. *International Conference on Information and Knowledge Management, Proceedings, 24-28-October, 1961–1964*.
- Macke, J., Sarate, J. A. R., and Vallejos, R. V. (2013). Collective Competence and Social Capital Analysis in Collaborative Networks. *Journal of Systemics, Cybernetics and Informatics*, 8(3), 18–23.
- Menon, Saraswathi and Karl, Judith and Wignaraja, K. (2009). Handbook on planning, monitoring and evaluating for development results. *UNDP Evaluation Office, New York, NY*, 68(3), 10.
- Michaux, V. (2005). Performance collective et compétences individuelle, collective et organisationnelle : construction d'une grille d'analyse unifiée. *Agrh*, 1–26.
- Nonaka, I. (1995). *Knowledge and Management: Chapter 2 of Ikujiro Nonaka and Hirotaka Takeuchi, The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press.
- Saeedi, A., Peukert, E., and Rahm, E. (2018). Using Link Features for Entity Clustering in Knowledge Graphs. In *Springer International Publishing AG, part of Springer Nature 2018* (Vol. 1, Issue March). Springer International Publishing.
- Sahoo, J., Mohanty, B., and Dash, I. (2018). An analytical study on the publication pattern and impact of top research papers: A case study of information processing and management. *Library Philosophy and Practice*, 2018(October), 1–22.
- Sanghi, S. (2016). *The handbook of competency mapping: understanding, designing and implementing competency models in organizations* (P. Enclave (ed.); Second Edi). SAGE publications India.
- Settouti, L. S., Prie, Y., Marty, J.-C. C., Mille, A., Prié, Y., Marty, J.-C. C., and Mille, A. (2009). A trace-based system for Technology-Enhanced Learning Systems personalisation. *Proceedings - 2009 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009*, 93–97.
- Tadger, H., Laffi, Y., Seridi-bouchelaghem, H., and Gülseçen, S. (2020). Improving soft skills based on students' traces in problem-based learning environments. *Interactive Learning Environments*, 0(0), 1–18.
- Thomas, S. A. (2015). *Data Visualization with JavaScript* (No Starch).
- Ughetto, P., and Paris-est, U. (2002). *Service Competence: an Overview*.
- van Leeuwen, T. N., and Wouters, P. F. (2017). Analysis of Publications on Journal Impact Factor Over Time. *Frontiers in Research Metrics and Analytics*, 2(April).
- Ventocilla, E., Bae, J., Riveiro, M., and Said, A. (2017). A billiard metaphor for exploring complex graphs. *CEUR Workshop Proceedings*, 1798, 37–40.
- Wang, N., Abel, M., Barthès, J., Negre, E., Wang, N., Abel, M., Barthès, J., Negre, E., Wang, N., Abel, M., Barthès, J., and Negre, E. (2016). *Mining user competency from semantic trace*. May 2015.

Session 4.1 - Session spéciale 6 : Vérification et Validation en Ingénierie Système

Chairman : Vincent Chapurlat - Pierre David

Synthesis of software architecture for the control of embedded electrical generation and distribution system for aircraft under safety constraints: The case of simple failures

LAURENT ZIMMER¹, PIERRE-ALAIN YVARS²

¹ DASSAULT AVIATION
78, Quai Marcel Dassault, 92252 Saint Cloud, France
laurent.zimmer@dassault-aviation.com

² Institut Supérieur de Mécanique de Paris (SupMeca)
Laboratoire QUARTZ
3 rue Fernand Hainaut, 93407 Saint Ouen Cedex, France
pierre-alain.yvars@supmeca.fr

Résumé – Dans l'aéronautique, le respect des exigences de sûreté de fonctionnement est le facteur déterminant pour la définition des architectures de systèmes. A l'avenir, les architectes systèmes devront s'aider d'outils d'assistance à la génération d'architectures correctes par construction. Dans ce contexte, nous nous sommes intéressés à la modélisation formelle des exigences et à la résolution informatique de ces modèles formels pour trouver des architectures admissibles. Dans ce papier, nous proposons une approche et un ensemble d'outils pour au moins vérifier puis synthétiser une architecture à sûreté de fonctionnement intégrée d'un système embarqué avion de production et de distribution de puissance électrique. Nous avons développé une approche à base de modèle qui utilise le langage DEPS (DEsign Problem Specification), un langage dédié à la modélisation et à la résolution des problèmes de conception. Les résultats obtenus montrent qu'il est possible de modéliser des exigences complexes de sûreté de fonctionnement au niveau requis par l'architecte système et que leur prise en compte pendant la résolution génère des solutions correctes vis-à-vis de celles-ci.

Abstract – In aeronautics, compliance with safety requirements is the main driver in the definition of system architectures. In the future, system architects will have to use tools to assist in the generation of correct architectures by construction. In this context, we are interested in formal requirements modeling and in the computer resolution of these formal models to find eligible architectures. In this paper, we propose an approach and a set of tools to at least verify and then synthesize a fail-safe architecture of an on-board aircraft electric power generation and distribution system. We have developed a model-based approach using the DEsign Problem Specification (DEPS) language, a language dedicated to modeling and solving design problems. The results obtained show that it is possible to model complex dependability requirements at the level required by the system architect and that taking them into account during the resolution process generates solutions that are correct with respect to them.

Mots clés - système embarqué avion de génération et de distribution électriques, sûreté de fonctionnement, synthèse à base de modèle, programmation par contrainte, vérification.

Keywords – on board aircraft electric power generation and distribution system, safety, model based system synthesis, constraint programming, verification.

1 INTRODUCTION

In this period of ecological and energy transition, several industrial sectors are being led to rethink their concepts to ensure a reduction in greenhouse gases. This is notably the case in the aeronautics sector, where we are talking about a

more electric, hybrid electric or all electric aircraft. These concepts, concerning the systems architecture domain, aims to replace hydraulic and/or pneumatic systems with electrical systems performing the same functions. On the one side, the expected benefits are numerous: better power rationalization, weight savings, improved aircraft availability and reduced

maintenance costs. All three will increase the profitability of commercial aircraft and meet current and future environmental requirements. On the other side as explained in [Menu et al, 2018] the increasing use of electrical components and electrical power systems gives rise to new constraints, particularly on the electrical power generation and distribution system architecture embedded in the aircraft. The replacement of hydraulic and pneumatic networks will result in a shift in reliability and safety requirements to the electrical network.

Engineers will have to design and verify new and more complex architectural solutions in order to respond to these requirements while limiting the impact in terms of mass [Giraud, 2014].

In this paper, we do not consider the sizing of system but we present through a simplified use-case inspired from an industrial system architecture how we can take into account the safety requirements during the design process:

In this paper, firstly we will present the issue, secondly our approach, thirdly some related work, fourthly we will present the use-case that is to say the embedded power generation and distribution system that we are studying as well as the associated safety requirements.

Fifthly we will present the DEPS problem modeling language [Yvars and Zimmer, 2014], dedicated to the formalization of system synthesis problems. The modeling of the problem in DEPS will then be detailed. The models will be compiled and solved under the DEPS Studio modeling and solving environment associated with the DEPS language.

Finally, some perspectives of evolution will be evoked.

2 ISSUE

In aeronautics, compliance with operational safety requirements is the main driving force behind the definition of system architectures.

These system-level requirements are generally the result of preliminary reliability analyses that determine the failure rate of functions critical to operational safety.

To ensure that failure rates are above the desired thresholds, system architects must implement architectures with a number of hardware and functional redundancies.

In particular, these dependability requirements will lead architects to develop back-up functions that will replace flight-critical functions when they fail. This is known as fail-safe architecture.

These architectures are difficult to develop and verify because redundancy clashes with other design criteria such as weight and cost, which favor the sharing of components. Failure of components common to normal and backup functions causes the fail-safe nature of the architecture to be lost. In order to avoid the problem it is mandatory to ensure that the realizations of normal and backup functions are segregated (i.e. they have no material elements in common).

In the following, we propose an approach and a set of tools to at least verify and later synthesize a fail-safe architecture of a power generation and distribution system.

More precisely, we will simultaneously address the following problems:

- to check that the electrical hardware architecture meets the the operational safety requirements with regard to the various failures that may occur during a flight;
- to synthesise the software architecture needed to control the reconfiguration of the electrical system when a failure occurs.

3 APPROACH

We adopt a model-based systems engineering approach with two levels of modeling: one level for the functional architecture of the system and one level for the physical architecture.

At the functional level, the power supply system services are organized in safety patterns that group together the normal and backup power distribution operating channels. It is at this level that the properties of physical segregation between channels will be expressed.

At the hardware level, we find the components that make up the physical architecture of the system, including: AC generators, bus bars, contactors, AC/DC converters and calculators.

We formalize the problem of verifying the conformity of the architecture to safety requirements as a generalized deployment problem (or resource allocation problem).

Generation, Distribution, Conversion, Command and Supervision functions, which are part of safety patterns, must be deployed on generators, distribution paths, convertors and calculators in a way that respects the segregation constraints.

It is important to point that these segregation constraints have to be expressed at the functional level even if they act at the hardware level. On the one hand, these segregation constraints are expressed on functions and channels and, on the other hand, we would like to have some form of genericity and reusability of the models.

In the following, we detail how we use DEPS a modelling language with constraint to set and solve the checking and/or synthesis problems of fail-safe architecture.

4 RELATED WORK

The majority of the published work in the field of embedded electrical power system design focuses on solving problems by satisfying some functional requirements rather than formalizing the problem and requirements. Thus, [Giraud, 2014] proposes an approach designed to solve the problem of load allocation on an aircraft electrical network. It involves sizing the energy sources and generating the connection paths between the energy sources and the loads in such a way that the energy demand is satisfied. The load allocation calculation is done using the implementation of a genetic algorithm. No formalization of the load allocation problem using an adapted description language is proposed.

Other researchers practice design by simulation of embedded electrical architecture as in [Yang et al, 2015]. Here again there is no abstract modeling of the problem. The performance of a given system is evaluated by simulation. If the result is suitable, the architecture is accepted; if not, it is necessary to manually modify some parameters of the system and then simulate it until a satisfactory solution is found. It can take a long time, especially when there are multiple requirements to be met. On the other hand, it is possible to achieve a very fine level of granularity for the physical models used.

From the point of view of taking into account safety requirements, [Menu et al, 2018] propose a two-step design process: the use of a graphical formalism intended to represent the architecture of the electrical system and a low level of variability reduced essentially to the range of possible values for the cardinality of the system components. A generation of the cartesian product of the candidate architectures is then carried out and fault trees are used to evaluate the failure probabilities of the different generated systems, with the

designer having to choose the satisfactory solution. The limits of this work concern on the one hand the weakness of the expressible variability and on the other hand the evaluation of the solutions.

However, as early as 2010, [Becz et al, 2010] and [Pinto et al, 2010] emphasize the need for both problem modeling formalism at a sufficient level of abstraction and synthesis tools to represent and solve complex system design problems in embedded aeronautics.

In our case, we will see that we have a typical case of formally representing and solving a problem of synthesis of software architecture for the control of the embedded electrical generation and distribution network in compliance with the safety requirements of this network.

[Creff et al, 2020] point out the difficulties of using formalisms such as SysML (System Modeling Language), initially designed to represent totally defined systems to model problems. They propose in their work to use the Clafer [Bak et al, 2014] formalism associated with the Choco [Lorca et al, 2014] constraint programming library to model and solve a problem of allocating calculators to embedded tasks. Clafer is a feature oriented modeling language [Kang et al, 1990] with very limited reusability capabilities. As a result, it remains a language dedicated to the configuration of software product lines. On their side, [Leserf et al, 2015] propose to add a first level of variability to the SysML language intended to be more universal. The approach is coupled with the Choco library to solve simple configuration problems. This work has so far remained at the research stage. The current limitations of these two approaches are the low level of variability that can be taken into account, the use of a solver handling essentially discrete constraints and a weak coupling between the formalism and the solver, which means that the development of models is carried out in case of a problem in the solver language and not in the Clafer or SysML language.

All these limitations have already been pointed out by Shah [Shah, 2010] and [Shah et al, 2012]. We will see in this paper that the DEPS language (DEsign Problem Specific language) [Yvars and Zimmer, 2014] and its recent integrated modeling and solving environment DEPS Studio [Yvars and Zimmer, 2019] are an attempt to address these limitations.

5 PROBLEM DESCRIPTION

5.1 System description

The case study is an aircraft electrical Power Distribution System (PDS). A PDS is responsible for the distribution of energy from the various generators to the various on-board loads (Figure 1).

From an electrical point of view, the system is composed of the following hardware elements:

- Three generators (GLC1, GLC2, RAT) which provide the electrical power;
- Eight bus bars (LH MAIN AC, LH ALT AC, LH MAIN DC, LH ESS DC, RH MAIN AC, RH ALT AC, RH MAIN DC, RH ESS DC) that provide electrical distribution services to the loads connected to them;
- Three converters to transform alternating current into direct current (T1, T2, SBT);
- Seventeen power contactors to connect and disconnect certain parts of the network in order to conduct or interrupt the transfer of power between these parts.

The network topology of the power generation and distribution system can evolve dynamically under the effect of calculators that control the contactors by means of their control ports. The calculators and the connection between the ports and the contactors are not shown on Figure 1.

Note that the system has been sized by electrical engineering specialists so that GLC1 or GLC2 alone can provide the electrical power required by all the loads.

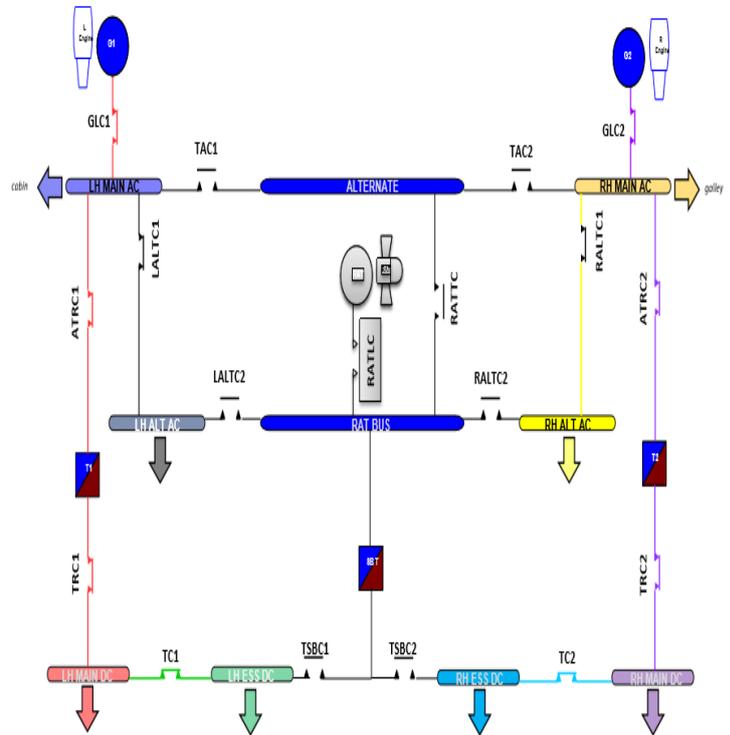


Figure 1. Aircraft electrical generation and distribution architecture

5.2 Control system description

To ensure the connection between the generators and the bus bars, the contactors must be connected or disconnected. To do this, they are controlled by means of processing programs running on calculators.

A calculator includes:

- a power supply unit (PSU),
- a microcontroller (μC).

A μC is a kind of processing unit on which are implemented the programs controlling the contactors. It interfaces with the environment via control ports.

It should be noted that for technological reasons the routing of contactor commands must respect the following property: a contactor can only be controlled by one and only one control port (P1).

It should be noted too that a PSU may fail frequently. Therefore a calculator can fail just like any other hardware equipment.

The role of the control system is twofold:

- It has to distribute electrical power to the bus bars by establishing distribution channels;
- It has to change the distribution channel in case of failure.

So we will have two kinds of μC programs; those controlling the distribution channels and those monitoring the switch from one channel to another.

Distribution channels are therefore made up not only of contactors but also of their treatments. In addition to this we have to consider transition channels made up of switching treatments.

5.3 Addressing safety requirements

Each bus bar can be seen as a system distributing power services to loads.

In normal operation, each bus bar of the PDS functions normally and distributes power via a so-called *normal distribution channel*.

If a power service is critical for the survival of the aircraft, then additional channels are added to the related system in order to guarantee service in the event of a failure. In this case a bar bus system will be composed of a normal distribution channel, a safety distribution channel and the related transition channel.

It is important to note that these safety channels are the result of preliminary safety studies. Therefore they are functional inputs of the problem we address.

In the following we will only consider the safety channels defined for single failures. Other safety channels exist in the case of multiple faults but there are out of the scope of the present paper.

In the case of single failures, the overarching requirement is as follows:

(R) In the case of a single failure occurring on equipment all bus bars must continue to be powered after reconfiguration of the system.

In this respect, it should be remembered that calculators are part of this equipment. Therefore calculator failures have to be considered.

In our case study, *R* means that the height bus bar system will have the same safety pattern made of: a normal distribution channel, a safety distribution channel and a transition channel.

To this pattern we must add the segregation constraints in order to avoid the hardware common points.

In an obvious way, the normal and safety distribution channels have to be segregated including the programs controlling the distribution.

Less obviously the normal and transition channels have to be segregated too. Indeed, if the program controlling the normal distribution and the transition program are on the same calculator then a calculator failure prevents switching to the safety distribution channel. The Figure 2 resumes the resulting pattern.

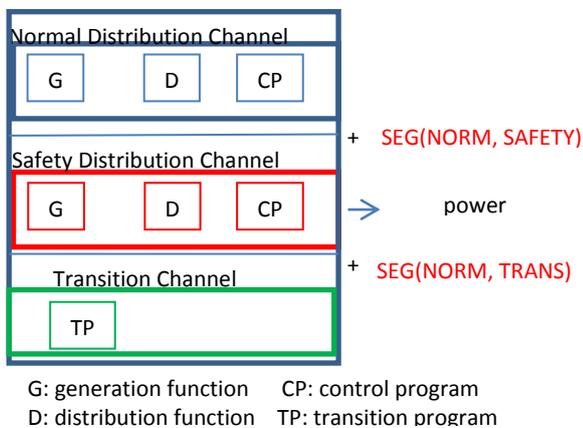


Figure 2. Example of safety pattern

Let's take as an example the LH MAIN AC system.

In the topology of Figure 3, the "LH MAIN AC" bus bar is powered by generator G1 via contactor GLC1. In the event of a single failure on generator G1, contactors TAC1 and TAC2 must be activated to allow continuous power supply to the bus bar via generator G2.

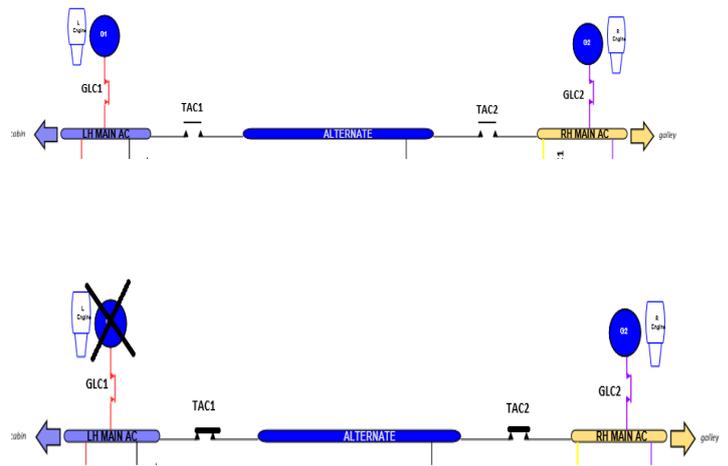


Figure 3. Example of normal mode and simple failure

Functionally, the LH MAIN AC system (see Figure 4) is composed of:

- A channel norm: G1, GLC1 and the control program,
- A channel safety: G2, GLC2, TAC2, TAC1 and the control program,
- A transition channel: the transition program.

Figure 4 shows a material deployment of LH MAIN AC.

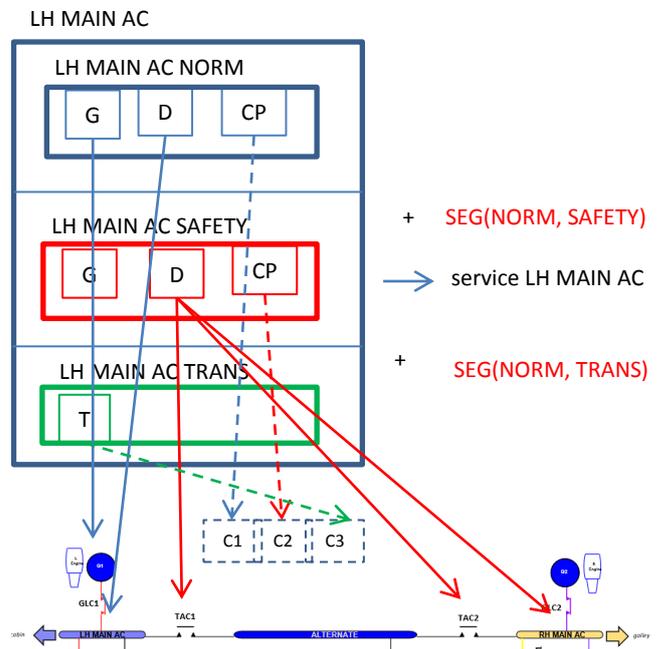


Figure 4. LH MAIN AC generation and distribution functions deployment

It should be noted that:

- For the generation (G) and distribution (D) functions of LH MAIN AC the deployment is known (NORM on G1 and GLC1, SAFETY on G2 and GLC2, TAC2, TAC1);

- For the control and transition programs the calculators and the ports used for the deployment are not known.

This is why in the following we put the emphasis on the deployment of the control and transition programs on calculators and on the routing between calculator control ports and contactors.

5.4 Software architecture synthesis

Thus, with a fixed electrical architecture, it remains the problem of determining:

- the necessary and sufficient number of calculators,
- how to allocate calculators to the control and transition programs,
- how to allocate the contactor commands to the calculator ports.

This last point requires specifying the relationships existing between contactor commands and contactors:

- R1: a contactor is controlled by one and only one contactor command of a calculator
- R2: a contactor command of a calculator controls one and only one contactor

All this has to be done in such a way that all safety requirements are met.

Finally it is not a question here of verifying that an existing control architecture verifies a posteriori the dependability requirements imposed on the electrical architecture, but rather of producing a control architecture that verifies by construction the dependability requirements. In other words, it is about to use the dependability requirements to build a solution. We are thus faced with a synthesis problem, requiring a formal model of the problem to be solved as well as a solving tool adapted to use this formal description to generate a correct architecture by construction [Pinto et al, 2010].

6 DEPS LANGUAGE

6.1 Paradigm

The Design Problem Specification (DEPS) language is commonly referred to as a Domain Specific Language (DSL). The target application domain is the specification and resolution of engineering problems, particularly those encountered in product or system design: sizing, configuration, allocation and architecture generation. The industrial interest is to use a unique formalism and tooling to model and solve all these categories of synthesis problems [Yvars, and Zimmer, 2021].

DEPS is an external (as opposed to internal or embedded) dedicated language. The source code is therefore independent of any host generalist language.

The DEPS language can be seen as a combination of a software or system modeling language and a mathematical programming language. From the former have been borrowed the features of structuring and abstraction which allow to represent the elements and the system under study. From the latter were borrowed the mathematical concepts necessary to solve the engineer's problems: unknowns, equations and inequalities.

This combination makes it possible both to represent design problems and to pose and then solve or optimize the systems of equations and inequalities that govern them [Yvars and Zimmer, 2019]. DEPS is supported by the non-profit organization DEPS Link (www.depslink.com).

DEPS has been used on problems of robot design [Yvars and Zimmer, 2014], battery synthesis [Diampovesa et al, 2020] and embedded avionics system synthesis [Zimmer et al, 2020].

6.2 Main characteristics

6.2.1 The Model

The fundamental feature of the language is the Model. Any Model is defined using the keyword *Model* followed by its name and its (possibly empty) list of arguments. It contains in order: a declaration-definition area for *Constants*, a declaration area for *Variables*, a declaration-creation area for *Elements* and a definition area for *Properties*. The definition of a DEPS Model ends with the keyword *End* (Figure 5).

The properties of a Model are the equations and inequalities that relate to the constants and variables of that Model. All the algebraic operators of the IEEE754 standard are available to build the properties. A Model therefore contains all the ingredients necessary to set the constraints that govern an instance of this *Model*. Some specialized constraints can also be set as constraints on data catalogs. Any instance of a Model will necessarily contain the set of constants, variables and Elements expressed in the Model and will necessarily have to verify the set of properties of the Model..

```

Model B(arg1, arg2) extends A
Constants
arg1 : Integer ;
Variables
CPU : CpuIndex;
Elements
arg2 : C[Integer, CpuIndex];
Elt1 : C(1, 2);
Properties

End

```

Figure 5. DEPS model example

As in an object language there is inheritance and composition: a Model can be extended and inherits constants, variables, elements and properties from another Model. Elements are instances of Models. They are either built inside the Model and in this case you have to call the constructor of the reference model with values given to its arguments or are passed as arguments to a Model creating an aggregation link with it. In this last case, the argument elements are named in the list of model arguments and declared in the Elements zone, specifying the signature of the model to which they refer. *Constants* can also be passed as arguments to a *Model*, allowing the creation of Parametric Models. DEPS also supports polymorphism.

Thus in Figure 5, we have a Model B that inherits (*extends*) from a Model A. Model B has two arguments: argument *arg1* which is an integer (*arg1: Integer;*) and argument *arg2* which is an instance of Model C which must necessarily have two arguments: the first being an integer, the second being a *CpuIndex*. All this is specified by the signature of the Model C (*arg2: C [Integer, CpuIndex] ;*). Finally, Model B is composed of an instance of Model C called *Elt1* and created thanks to the call *C(1,2)*.

Modeling a problem to be solved is therefore like as building DEPS Models and assembling elements in a main model without arguments expressed using the keyword *Problem*.

6.2.2 The Quantity

In DEPS, both constants and variables are associated with types of physical or technological magnitudes called quantities (Quantity). They are necessary in Systems Engineering (Figure 6).

A Quantity has:

- A basic type of quantity (called QuantityKind). For example, Real, Integer, Length;
- A min (resp. max) bound that represents the minimum (resp. maximum) value that can be taken by any constant or variable having the defined quantity as its type;
- A unit of the quantity. For example the meter *m* for a length and *u* for quantities without unit.

A QuantityKind has:

- a basic type (integer or real),
- a min and max bound,
- a dimension in the sense of the dimensional analysis of the quantity. For example L for a length or U for a dimensionless quantity;

QuantityKind Integer	Quantity CpuIndex
Type : integer ;	Kind : Integer ;
Min : -maxint;	Min : 1 ;
Max : +maxint;	Max : 4 ;
Dim : U ;	Unit : u ;
End	End

Figure 6. QuantityKind and Quantity

6.2.3 Implementation

The DEPS Studio integrated modeling and solving environment associated with the DEPS language includes model editing functions, project management functions based on a package mechanism, a compiler and a solver.

An integration approach rather than a model transformation approach has been chosen. Indeed, in the case of solving a sub-defined system synthesis problem, it will be necessary in case of unsatisfactory computation results to perform a model tuning in the DEPS language. By deliberately opting for an integration approach, we choose this fine-tuning process.

The computational methods we use are taken from the work on the resolution of CSP [Tsang, 1993]. The solver implements the well-known revised HC4 (HC4Rev) propagation method [Benhamou et al, 1993] on equations and inequalities. We have extended the HC4Rev algorithm, initially designed to handle open real intervals, to the following three types of domains: integer intervals, enumerated sets of floating values and enumerated sets of signed integer values.

The object-oriented architecture of the solver has been designed so that it can be extended to other propagation and/or resolution methods.

7 MODELING THE PROBLEM IN DEPS

We briefly present here some models extracted from the synthesis problem explained in the chapter 5. They concern the model of the LH MAIN AC system and they have been chosen to highlight some technical points.

7.1 Modeling the subdefinite systems

The LH MAIN AC system follows this model (Figure 7). We recognize the channels NORM, SAFETY and TRANS and the segregation constraints between NORM and SAFETY and NORM and TRANS.

The channels NORM and SAFETY have the set of contactors as parameters since they act on them.

Model S1 () extends ThreeChannelSystem[ContactorSet]
Constants
Variables
Elements
ch1: ChS1Norm(ContSet);
ch2: ChS1Safety(ContSet);
ch3: TRANSChannel();
seg1 : SEG(ch1, ch2);
seg2 : SEG(ch3, ch1);
Properties
End

Figure 7. DEPS LH MAIN AC model

7.2 Modeling a channel

This model of channel (Figure 8) is composed of a generation function, a distribution function and a control program for commanding contactors belonging to the set of contactors.

Model ChS1Norm () extends GDChannel[ContactorSet]
Constants
Variables
Elements
GenF: S1NormGenFunction();
DistF: S1NormDistFunction (ContSet);
ProcF: S1NormProcFunction(ContSet);
Properties
End

Figure 8. DEPS channel model

7.3 Modeling a function

This control function (Figure 9) commands one contactor. It extends a processing function with a constraint in the property zone which expresses that the control program and the contactor control are on the same calculator.

Model S1NormProcFunction() extends OneContactorProcFunction[ContactorSet]
Constants
Variables
Elements
Properties
L1.ProcIndex = ContSet.GLC1.ProcIndex;
End

Figure 9. DEPS function model

7.4 Modeling a segregation

This model segregates two control functions (Figure 10). The first one commands one contactor and the second three contactors. The second function is composed of three logic subprograms each of them controlling a single contactor.

Model SEG(ProcF1, ProcF2) extends SEG[ProcFunction[ContactorSet], ProcFunction[ContactorSet]]
Constants
Variables
Elements
ProcF1 : OneContactorProcFunction[ContactorSet];
ProcF2 : ThreeContactorProcFunction[ContactorSet];
Properties
ProcF1.L1.ProcIndex <> ProcF2.L1.ProcIndex;
ProcF1.L1.ProcIndex <> ProcF2.L2.ProcIndex;
ProcF1.L1.ProcIndex <> ProcF2.L3.ProcIndex;
End

Figure 10. DEPS segregation of two control functions

All the segregation requirements are expressed in the bar bus systems (Fig 7).

7.5 Modeling the problem

Finally, we create the full problem containing all the contactors of the electrical architecture (TheContactors), the eight bus bars (F1, F2, ..., F8) as well as complementary segregations (seg1, seg2, seg3, seg4) necessary for the complete expression of the problem and allowing to express the segregation of the NORM channels of the bus bars between the left and side parts of the aircraft (Fig 11).

Problem FailSafeVerification
Constants
Variables
Elements
TheContactors: ContactorSet();
F1: S1(TheContactors); F2: S1(TheContactors);
F3: S1(TheContactors); F4: S1(TheContactors);
F5: S1(TheContactors); F6: S1(TheContactors);
F7: S1(TheContactors); F8: S1(TheContactors);
seg1: SEG(F1.ch1, F3.ch1);
seg2: SEG(F2.ch1, F4.ch1);
seg3: SEG(F5.ch1, F7.ch1);
seg4: SEG(F6.ch1, F8.ch1);
Properties
End

Figure 11. DEPS model of the Problem

7.6 Solving the problem

After a compilation step, the DEPS solver produces fail-safe architectures with a control system requiring at the best four calculators.

The results of allocating the calculators to the contactors are shown in Table 1.

The results of the allocation of the calculators to the different contactor control programs for the normal and safety modes of the bar bus systems as well as their transition management programs are shown in Table 2.

Table 1. Routing between contactors and calculators

Contactors	Calculator Index
GLC1	1
ATRC1	1
TRC1	1
LALTC1	4
GLC2	2
ATRC2	2
TRC2	2
RALTC1	3
RALTC	1
RATTC	1
TAC1	3
TAC2	4
LALTC2	3
RALTC2	4
TC1	2
TSBC1	3
TSBC2	3
TC2	1

Thus, if we remember that for the LHMAIN AC bus bar, the NORM channel has to control contactor GLC1 (Fig. 3), we read in table 2 that the control program for this contactor is deployed on calculator 1. In the same way, the contactors to be actuated for the SAFETY channel are contactors TAC1, TAC2 and GLC2 which control programs are deployed on calculators 3, 4 and 2. The transition program TRANS is deployed on computer 3.

The whole is coherent since we can see in table 1 that contactor GLC1 is routed to calculator 1 and that TAC1, TAC2 and GLC2 are routed respectively to calculators 3, 4 and 2.

Table 2. Deployment of control and transitions programs on calculators

Bus bar system	NORM calculator Index	SAFETY calculator Index	TRANS calculator Index
LHMAIN AC	1	3, 4, 2	3
LHMAIN DC	1, 1	4, 3, 3, 2	2
RHMAIN AC	2	4,3,1	3
RHMAIN DC	2, 2	3, 4, 3, 1	1
LHALTAC	4	3, 1, 3	1
LHESSDC	2	4, 3, 3	1
RHALTAC	3	4, 1, 4	1
RHESSDC	1	3, 4, 3	2

8 CONCLUSION

In this article we have shown that it is possible to capture safety requirements that are difficult to formalize directly without a suitable formal language. We used the DEPS language which allowed us to model the right abstractions at the right level to describe a sub-defined model of architecture software architecture for the control of embedded electrical generation and distribution system for aircraft on the one hand, and safety requirement models on the other hand. The deployment problem was solved by applying constraint satisfaction methods to the properties of the problem with DEPS Studio environment. The structuring features offered by the language facilitate the reuse of the models developed.

Future work will focus on taking into account additional safety requirements as double faults and ultimate rescue mode and additional components as fault sensors. The evolution of the models will follow those of the DEPS language in particular on the possibilities of manipulation of collections of objects in the future versions of DEPS and DEPS studio.

9 REFERENCES

- Giraud, X. (2014). *Méthodes et outils pour la conception optimale des réseaux de distribution d'électricité dans les avions*. Phd Thesis INSA Toulouse.
- Yang Z, Qu J & Shi X (2015). Modeling and Simulation of Power Distribution System in More Electric Aircraft. *Journal of Electrical and Computer Engineering*.
- Menu, J., Nicolai, M., & Zeller, M. (2018.). Designing Fail-Safe Architectures for Aircraft Electrical Power Systems. *Proc of AIAA/IEEE Electric Aircraft Technologies Symposium*.
- Becz, S., Pinto, A., Zeidner, L. E., Banaszuk, A., Khire, R., & Reeve, H. M (2010). Design System for Managing Complexity in Aerospace Systems. *13th*

- AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.
- Pinto A., Becz S. & Reeve H.M (2010). Correct-by-construction design of aircraft electric power systems. *Proc of AIAA Aviation Technology, Integration, and Operations Conference*.
- Creff S, Le Noir J, Lenormand E & Madelénat S. (2020). Towards Facilities for Modeling and Synthesis of Architectures for Resource Allocation Problem in Systems Engineering. *Proc of 24th Systems and Software Product Line Conference*. Montreal.
- Bąk K., Diskin Z., Antkiewicz M., Czarnecki K. & Wąsowski A. Clafer: Unifying class and feature modeling. *Software and Systems Modeling, 2014*.
- Lorca X., Prud'homme C. & Fages J-G (2014). *Choco3 Documentation*. TASC, INRIA Rennes, LINA CNRS UMR 6241, COSLING S.A.S.
- Kang, K.C., Cohen S.G., Hess J.A., Novak, W.E. & Peterson, A.S. (1990). *Feature-oriented domain analysis (FODA) feasibility study*. Technical Report CMU/SEI-90-TR-021, SEI, Carnegie Mellon University.
- Leserf P., Saqui-Sannes P., Hugues J. & Chaaban K. (2015). SysML Modeling For Embedded Systems Design Optimization: A Case Study. *Proc of 3rd International Conference on Model-Driven Engineering and Software Development (MODELSWARD)*.
- Shah A.A. (2010). *Combining mathematical programming and SysML for component sizing as applied to hydraulic systems*. Master Thesis, Georgia Institute Of Technology.
- Shah A.A., Paredis C.J.J., Burkhart R. & Schaefer D. (2012). Combining Mathematical Programming and SysML for automated Component Sizing of Hydraulic Systems. *Journal of Computing and Information Science in Engineering (JCISE)*.
- Yvars, P.-A. & Zimmer, L. (2014). DEPS Un langage pour la spécification de problèmes de conception de Systèmes. *Proc of the 10th International Conference on Modeling, Optimization & SIMulation (MOSIM)*. Nancy.
- Yvars, P.-A. & Zimmer, L. (2019). DEPS Studio : Un environnement intégré de modélisation et de résolution de problèmes de conception de systèmes. *Proc of 8^{ème} Conférence en ingénierie du logiciel (CIEL)* Toulouse.
- Yvars, P-A & Zimmer, L (2021). Integration of constraint programming and model based approach for design synthesis. *Proc of IEEE Systems Conference SYSCON 2021*, Vancouver.
- Diampovesa S, Hubert A & Yvars P-A (2020). Modélisation d'un problème de conception en vue de réutilisabilité. Exemple d'une batterie Li-ion. *Proc of Symposium de Génie Electrique (SGE)*, Nantes.
- Zimmer L, Yvars P-A & Lafaye M (2020). Models of requirements for avionics architecture synthesis: safety, capacity and security. *Proc of the 11th Complex System Design and Management (CSDM) conference*.
- Tsang E (1993). *Foundations of Constraint Satisfaction*. London and San Diego: Academic Press.
- Benhamou F., Goualard F., Granvilliers L. & Puget J.F. (1993). Revising Hull and Box consistency. *16th International Conference on Logic Programming*.

Implementing Quality Function Deployment in Model Based Systems Engineering

IMEN AOUF¹, PIERRE DAVID¹ & ERIC BLANCO^{1,2}

¹ Univ. Grenoble Alpes, CNRS, Grenoble INP*, G-SCOP, 38000 Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes
prenom.nom@grenoble-inp.fr

² Ecole de l'Air
BA 701 13 Salon Air
prenom.nom@ecole-air.fr

Résumé – Afin de déployer l'Ingénierie Système, diverses méthodes et outils ont été développés pour gérer la transition entre les différentes étapes du cycle de développement. L'approche QFD (Quality Functional Deployment) permet d'intégrer la voix du client dans le processus de développement du produit, de la conception à la fabrication. Malgré la diffusion du QFD son intégration aux approches MBSE est rare [Benner, et al. 2003].

Le QFD est conçu pour améliorer la transition entre les différentes activités de l'ingénierie des systèmes, comme l'identification des besoins des clients, la spécification des exigences système et la définition de l'architecture système. Plusieurs difficultés existent pour son utilisation dans un contexte MBSE. Ce travail se focalise sur le développement d'outils et de méthodes pour faciliter la mise en œuvre du QFD au sein de MBSE en utilisant le langage SysML. Ce travail de recherche se concentre sur la première phase de l'approche QFD liée à l'analyse et la définition des exigences systèmes. Nous proposons une structure de données sous la forme d'un méta-modèle pour gérer les QFD en SysML. Une méthode est proposée pour assurer la mise à jour dynamique des données MBSE par l'échange automatique des données entre l'analyse QFD et les diagrammes SysML.

Abstract – Quality Functional Deployment (QFD) is developed to bring the voice of customer into product development process from conceptual design to manufacturing. QFD approach might be beneficial to apply MBSE by efficiently ensuring the transition between the different steps of systems engineer development cycle. However, until now, the implementation of QFD approach is still difficult due to complexity of data management and update.

The purpose of this research is to develop tools and methods to facilitate the implementation of QFD within MBSE using SysML language. This research work focuses on the first phase of the QFD related to the analysis and the definition of system requirements. This phase is also known in the literature as the house of quality (HoQ). A meta-model is proposed to describe the shared data between MBSE activities and HoQ analysis. A method is proposed to dynamically update MBSE data by automatically exchanging data between HoQ analysis and SysML diagrams.

Mots clés – Ingénierie Système Basée sur les modèles, QFD, Analyse des besoins.

Keywords – MBSE, QFD, needs analysis.

1 INTRODUCTION

The increasing number of customer needs leads to modify numerous systems and develop new products. New systems are massively concerned with an increasing number of interconnections, operational scenarios and potential interactors or stakeholders. Therefore, products become more and more complex to meet these needs. To manage this complexity, engineering processes has been developed as systems engineering (SE) standards (e.g. IEC 15288).

SE processes have to handle considerable amounts of data regarding, needs, requirements, interfaces, verification plans and system description through the system development life cycle (INCOSE, 2015). To deal with this characteristic, Model Based approaches have been developed to capture such information needed to conduct system-engineering activities. This approach is known as Model based systems engineering (MBSE) (Friedenthal et al., 2012).

On the other hand, in order to enhance SE processes, various methods and tools have been developed to manage the

transition between the different steps of development cycle. As examples, we may cite QFD (Quality Function Deployment) (Singh & Kumar, 2015) or FMEA (Failure Modes and Effects Analysis) (David et al. 2010) that use very similar approaches. The quality function deployment (QFD) approach permits to bring the voice of the customer into the product development process from conceptual design to manufacturing. Despite, the important role ensured by the QFD in the success of multiple projects, it can be noticed that it is rarely seen in MBSE (Benner et al., 2003).

The QFD is helpful to apply SE by efficiently ensuring the transition between the different activities in systems engineering development cycle, such as identifying customer needs, specifying system requirements and defining the system architecture. However, several difficulties still exist for QFD as the issues related to its implementation and the management of the associated data to assist SE activities.

This paper focuses on development of tools and methods to facilitate the implementation of the QFD within MBSE processes using SysML.

In the remaining of this paper, QFD and its limits is first introduced. Then, a meta-model of the data organization is then defined along with a framework for QFD deployment in MBSE. The paper is finally concluded by giving perspectives to this work.

2 QUALITY FUNCTION DEPLOYMENT

The Quality Function Deployment (QFD) is a product development approach devoted to translate customer's requirements into design specification in order to progress and develop new products (Carnevali & Miguel, 2008).

QFD has been developed since its inception in 1960. New methods and tools have been established to meet the changing business conditions of QFD users, their management, their customers and their products. These methods and tools are presented in the standard ISO 16355 which presents decades of improvements to QFD (ISO 16355).

According to its promoters, the application of QFD approach permits to improve: customer satisfaction with products at time of launch, cross functional communication, systematic and traceable design decisions and efficient use of resources. Likewise, it is claimed in its dedicated standard, that this method reduce rework, time to market and lower life cycle cost (ISO 16355). QFD is used in many companies that implement Design for Six Sigma methodology. It is recognized as a powerful representation to support group work on requirements. It allows interaction and coordination between Marketing, Engineering, manufacturing and supply chain specialists. It allows to concentrate in a single chart many information related to design specification.

2.1 The House of Quality

The majority of QFD studies cited in the literature are restricted to the house of quality phase because the method's potential benefits are based on the success of this phase (Verma et al.,1998).

As well, this works is concentrated to this phase. The construction of the house of quality (HoQ) is divided to six regions which are illustrated in figure 1

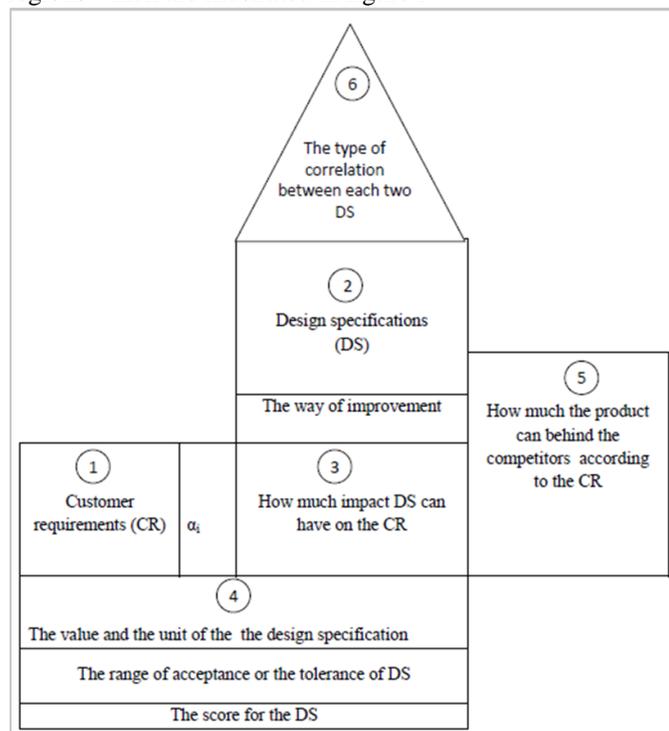


Figure 1. The house of quality (Moubachir & Bouami, 2015)

- In first region, the customer requirements are gathered through various methods and recorded in this region with a coefficient α_i that reflects the importance of each customer requirements.

- The second region: the how region is concerned to the design specification or system requirements. It is the crucial part of the HoQ as the customer requirements is translated to design specification to be implemented in the solution. The design specifications have to be controllable and measurable characteristics of the product. Each design specification is defined with a coefficient that reflects the way of improvement to meet the customer requirements: increasing, decreasing.

- The third region: "the What vs the How" region determines the relationship between customer requirements and design specification. In fact, the QFD team must assign a weighting based on the question "How important is the design specification A is in satisfying customer requirement B". The relationship can be: weak, moderate, or strong and carry a numeric value = {1, 3 or 9} to estimate the impact.

- The fourth region: The How Much region, this region is divided into three lines. In the first one concerns the target values for the design specification. In the second the range of acceptance or the tolerance for Design specification is defined. In the last line, the score for the Design specification is given.

- The fifth region: The Benchmarking region where the users' products are compared to the competitors in regards to the customer requirement.

- The sixth region: the How vs How region, defines the type of correlation between each two Design specifications whether it is positive, negative. It represents the impact of improving one specification on the other one. The QFD team should document strong negative relationships between design specifications and work to eliminate physical contradictions.

2.2 QFD deployment hinders

Many difficulties presented in the application of the QFD process may reduce its efficiency. Many of these difficulties concern methodological aspects. It is related to the elaboration of quality matrices stage: interpreting the voice of customer (Ginn & Zairi, 2005), identifying the most important customer demands (Yan, Khoo, & Chen, 2005) and project decision making since correlation among the demands are not clear.

Reducing the methodological difficulties in developing the HoQ is the fundamental point to enhance and expand the implementation of QFD.

In this direction, one hundred ninety-four recommendations about ways to aid application of QFD were identified which consists of submitting methods and tools to decrease difficulties relates to its use (Carnevali & Miguel, 2008).

The main recommendations presented in the literature were about the development of the quality matrix. In fact, to reduce the size of quality matrix in order to identify easily the demanded qualities, (Lowe & Ridgway, 2000) suggest a maximum number of 8 items for quality characteristics. This suggestion is not applicable in complex project. Then, (Chou, 2004) proposes another method to exceed matrix analyzing difficulties. The author proposes that instead of reducing the number of items in the quality matrix, the QFD users could divide the projects into groups of subprojects. This permits to analyze data separately in several smaller matrices. Nonetheless, new difficulties can be generated with this strategy because it does not easy to consider the correlation between demanded qualities and quality characteristics.

For the difficulties related to the interpreting of customer requirements, several solutions are presented in the literature

including the use of fuzzy logic (Shipley, de Korvin, & Yoon, 2004) the use of focus groups with brainstorming and cause and effect diagrams and using the Kano model.

The proposed methods and tools to resolve QFD application permits to increase the QFD performance but they raise its complexity of application. For example, the fuzzy logic cares to define the required degree of quality importance and help to do the correlation of matrix data. Nevertheless, these methods make the application of QFD even more complex.

Further work encompasses research in this direction to reduce the difficulties in the use of QFD. The work presented here proposes to facilitate the implementation of the QFD process by integrating it to MBSE processes. The approach tends to keep the simple formalism of the HoQ to facilitate interdisciplinary interaction and to organized data in order to feed the system model. Being integrated in a computerized environment as in MBSE will permit to automatically fill HoQ parts and to provide algorithms to trace and check inconsistencies between the project and HoQ results. In the remaining of the paper, technical development to create this integration are presented.

support the specification of system requirements in order to satisfy customer needs (Liu, Wang, & Fu, 2017). The HoQ approach is made to support this phase. When deploying MBSE, methods are needed to structure the use of languages, diagrams and tools. QFD can be such a structuring method for the customer needs analysis phase.

To establish this link and simplify the data exchange between HoQ and MBSE activities, it is necessary to define a unified data structure. This is why, a meta-model is developed to define the structure of input and output data from HoQ that are integrated in systems engineering activities. It also serves to present the data from HoQ to supportive SysML diagrams.

The meta-model of house of quality is presented with SysML via block definition diagram using PAPYRUS modeling tool. Figure 2 presents this meta-model.

In the proposed meta-model, “house of quality” block is composed of “VOC ”block and “VOE” block. The ”VOC ”: voice of customer is raw unprocessed information from the customer or stakeholder. The “VOE ”: voice of engineer is the strategy or technology service planned, communicated and in tune with related departments within organization. (ISO 16355).

3 A METAMODEL FOR QFD IN SysML

Nowadays, many systems engineers use model-based approach. Some are based on the use of SysML diagrams, to

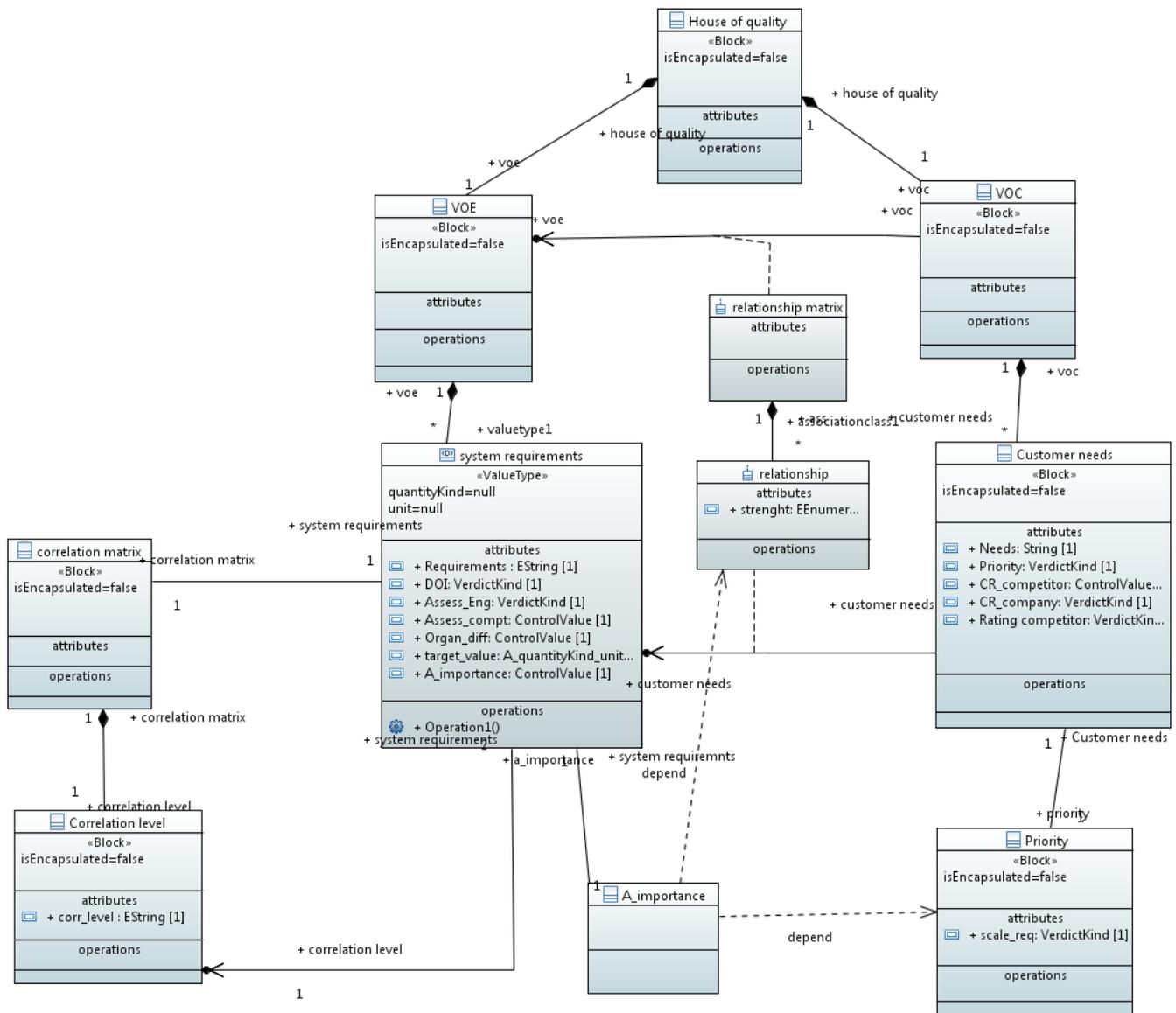


Figure 2. The meta-model of the house of quality

The relation between “VOC” and “VOE” is illustrated via the intermediate block “relationship matrix”. The “relationship_matrix” presents a part of house of quality matrices. It is composed of an intermediate block between “system requirements” and “customer needs” which is “relationships” block. The “relationships” is characterized by the strength attribute which defines the possible relationship can be done: it can be equal to 9, 3 or 1.

The “VOC” presents the matrix part of the house of quality. It is composed of “customer needs” block. Indeed, the “customer needs” is defined through the following attributes: needs, Priority, CR_competitor, CR_company, and rating competitor. These attributes respectively identify the description of customer needs, the importance rating of each need according to the costumer, customer rating company to identify the customer rating of each need in the company, customer rating competitor to determine the customer rating of each need against competitor.

On the other hand, the “VOE” is composed of the value type “system requirements”. A value types is used because it describes the type of values and may have an associate unit and dimension which is required to define the target value.

As well, the “system requirements” is detailed with the following attributes: requirements, DOI, Assess_Eng, Assess_comp, Organ_diff, target value and A_importance. These attributes respectively identify: a presentation of system requirements, direction of improvement of each system requirement which can be increase or decrease of each parameter, the engineering assessment of each requirement, the organizational difficulties which describes the assessment of the technical realization of system requirement in the company, target value which presents the desired value of each requirement to meet customer needs, and absolute importance indicates the importance of the realization of each requirement to satisfy the needs. As the absolute importance is a result of calculating the product of the relationship value and the priority value to customer needs, it is mentioned in the meta-model a dependency between the priority block and the relationship block.

In addition, for “system requirements”, an association relation is created to identify the correlation between each system requirements the “correlation matrix” block. “Correlation matrix” is composed of “correlation level” block which defined by the correlation attribute. A relationship between two system requirements can have the following levels: strong positive, medium positive, medium negative or strong negative.

4 A FRAMEWORK FOR QFD IN A MBSE PROCESS USING SysML

Currently, many systems engineers use model-based approach relying on SysML to capture data as identifying the customer needs, requirement, and comparing multiple architectures. In other hand, many of them have experience in using QFD analysis through dedicated tools, Excel sheets or paper documents. In this work a framework is proposed to organize the exchange of data between QFD (house of quality) and SysML models. The framework is sketched in Figure 3.

As a tool selection, it is proposed to work with Papyrus tool for SysML edition and Excel for QFD edition. Other tools may be chosen just modifying the technical part of the implementation. Firstly, the meta-model is created via papyrus. After that, an UML profile adapted to the meta-model is developed. This profile permits to easily define the necessary data of the HoQ and to present them in the SysML models. In another hand, an

XML structure is developed to describe the information embedded in QFD in order to support the transition between tools. Then, a conversion from XML files to Excel sheets is required with dynamic update between files.

Finally, the XML file is reloaded in papyrus to ensure a data update in SysML models.

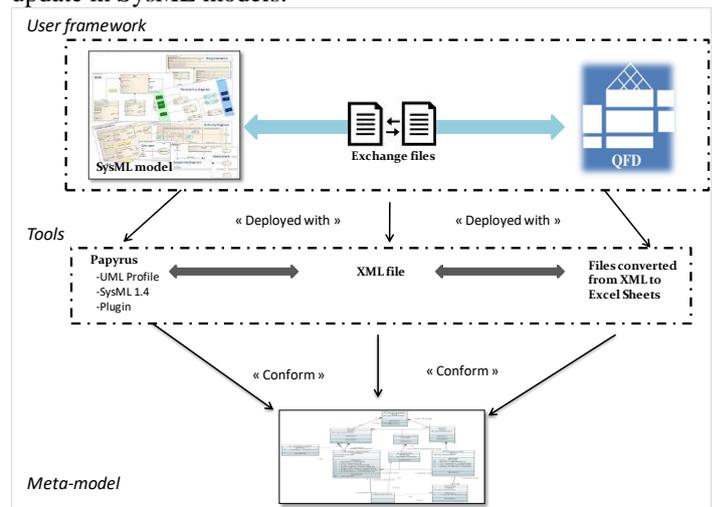


Figure 3. Global architecture of the framework

4.1 UML profile & Plug-in

After the development of the HoQ meta-model, a set of tools are exploited to facilitate the realization of the user framework. Firstly, the development of a UML profile is proposed. The basic premise of profiles is that all domain-specific concepts are derived as extensions or refinements of existing UML concepts, called UML meta-classes. These extensions are called stereotypes. A stereotype definition must be consistent with the abstract syntax and semantics of standard UML meta-classes it extends. Consequently, a profile-based model can be created and manipulated by any tool that supports UML standard.

In this paragraph, the parameters names are mentioned in *Italic*. The UML profile corresponds to the structure of the HoQ as shown in figure 3 and makes the link to UML meta-classes to permit their representation in UML tools. The profile is presented in figure 4. The proposal has been verified by tracing that all needed concepts of HoQ are covered in the profile. Academic examples have been run to verify the correct execution.

In fact, a stereotype of “house of quality” extended from the UML meta-class (class) is composed of two stereotypes extended from the same type of UML meta-class: the “VOC” (voice of customer) and “VOE” (voice of engineer).

The “VOC” is composed of “customer needs” stereotype and “VOE” is composed of “system requirements” stereotype. Likewise, the customer needs and the system requirement stereotype are presented with the same attributes cited in the meta-model.

The relation between “VOC” and “VOE” is presented via the stereotype “relationship matrix” extended from the UML meta-class “dependency”. The “relationship matrix” presents a part of house of quality matrices. It is composed of the stereotype “relationships” which presents an extension of both: “systems requirements” and “customer needs”. It permits to identify “the relationship strength” which can have the following value 9, 3 or 1.

For the “system requirements” stereotype, an association

relation is created to identify the correlation matrix between two system requirements. Thus, a “correlation matrix” stereotype is created which is extended from the dependency UML meta-class. This stereotype is composed of “correlation level” stereotype where the correlation attribute defines the type of correlation between two system requirements which can be strong positive, medium positive, medium negative or strong negative.

The corresponding level of correlation is mentioned in an Enumeration.

4.2 Exchange data: XML and Excel

XML technology is designed to manage and share structured data in a user-readable text file. This technology complies with industry standards and can be processed by various databases and applications.

To enable the data exchange between the SysML modeling tool and the QFD edition tool (for example Excel sheets), XML file is developed which presents the different parts of HoQ in tags forms. See an excerpt of the XML structure filled with the data of an example system on Figure 5.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <houseofquality>
3   <VOC>
4     <Customerneeds ID="cus_01" needs="quick selection of lighting mode" priority
5     <Customerneeds ID="cus_02" needs=" good vision of proximity " priority="10"
6     <Customerneeds ID="cus_03" needs="ease opening to replace batteries" priorit
7     <Customerneeds ID="cus_04" needs=" enlighten by keeping your hand free " pri
8     <Customerneeds ID="cus_05" needs="operate in the rain" priority="5" cus_rati
9     <Customerneeds ID="cus_06" needs="torche distance during displacement" prior
10    <Customerneeds ID="cus_07" needs="to be comfortable " priority="10" cus_rati
11    <Customerneeds ID="cus_08" needs="use standard batteries" priority="10" cus_
12    <Customerneeds ID="cus_09" needs="not too expensive" priority="10" cus_ratin
13  </VOC>
14  <VOE>
15    <Systemrequirements ID="req_01" requirement=" beam width" DOI="increase" As
16    <Systemrequirements ID="req_02" requirement="battery type" DOI="" Assess_E
17    <Systemrequirements ID="req_03" requirement="autonomy" DOI="increase"ASSE
18    <Systemrequirements ID="req_04" requirement="sealing" DOI="increase"ASSE

```

Figure 5 Excerpt of the XML structure

4.3 Implementation example

The framework principle has been tested on a headlamp system design. Various scenarios of utilization of QFD in a MBSE have been tested. First the data stored in a excel driven QFD have been propagated to Papyrus SysML diagrams. For example, System requirements raised in the HoQ have been created in the SysML tool repository with their associated parameters has their priority. On the other hand, elements of SysML diagrams are used to populated the QFD tables. For example, customer needs expressed through use cases are used to populate the Voice of Customer part of the QFD.

Then the contribution is to provide a dynamic update between the SysML model, and the house of quality. When a requirement attribute is modified in the HoQ it is propagated to its representation in the SysML Diagrams. The link is made through the shared XML files gathering the HoQ data has expressed in the meta-model defined in figure 3.

Figure 6 illustrates the proposal on the analysis of a customer needs. For example, in the use case diagram a set of activities of the headlamp is presented such as “illuminate without using hand”. This function can be detailed in the HoQ through a set of customer needs such: “easy selection of different modes”. With the usage of HoQ this use case of the headlamp can be enhanced by providing data about the priority of this activity regarding the customer point of view.

On other hand, the HoQ implemented via excel is converted into a SysML diagram with the application of the UML profile described in Figure 4.

The “easy selection of mode” requirement is presented with different attributes provided by the HoQ and its corresponding system requirements. These elements are later on reused in other SysML diagrams used by the project.

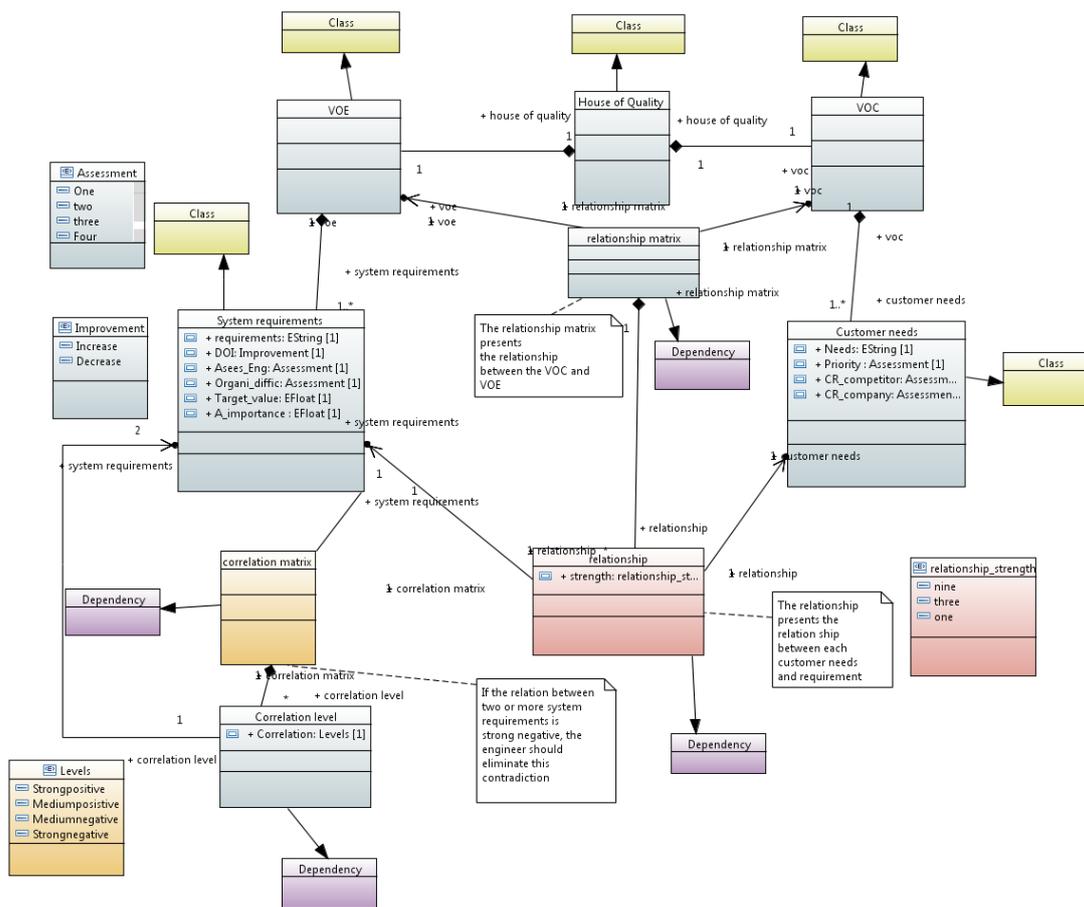


Figure 4. The House of Quality profile

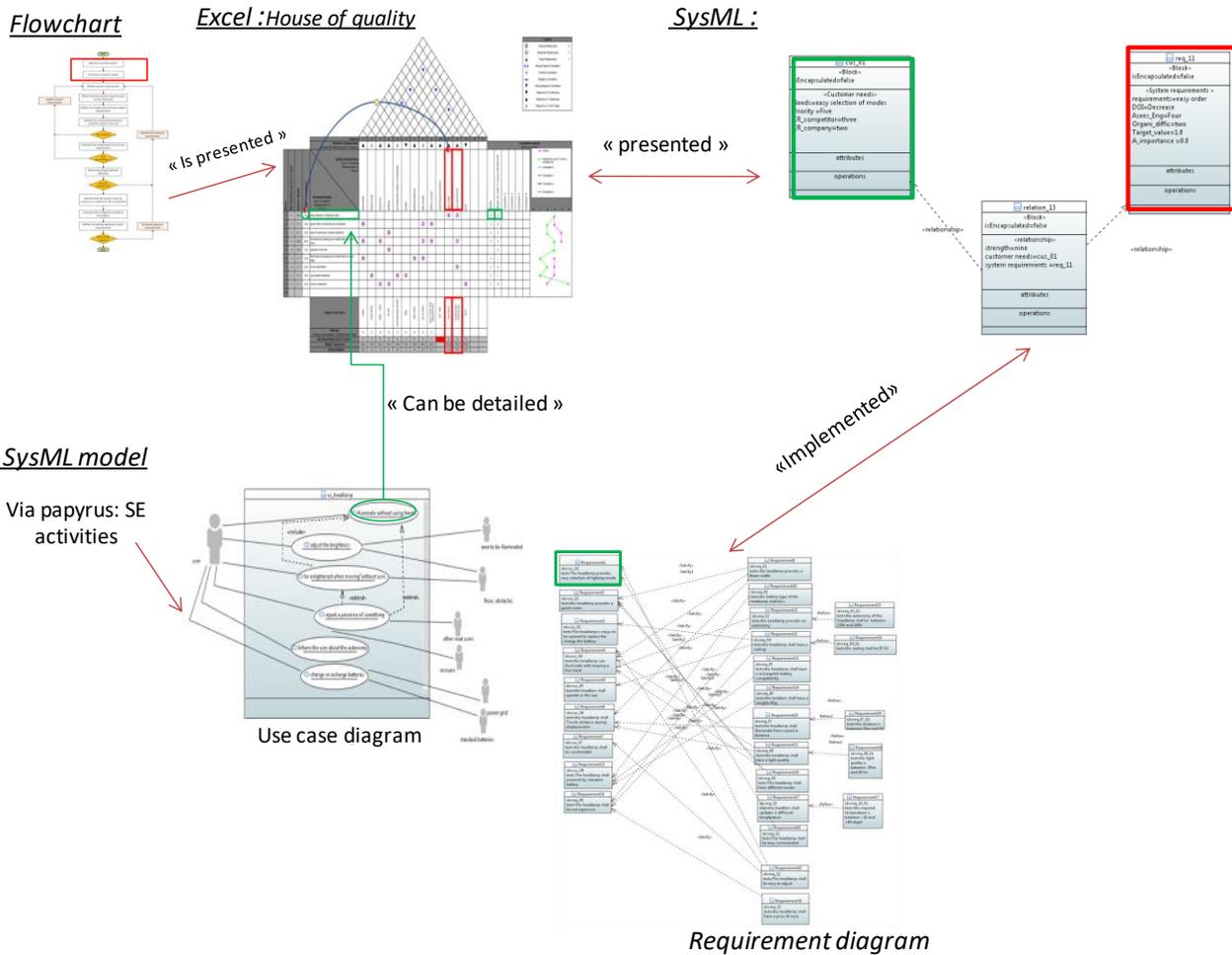


Figure 6. The implementation of a customer needs of house of quality in SysML model

5 CONCLUSION

The goal of this study is to make a dynamic link between QFD analysis approach and MBSE activities workflow. At present, the implementation of QFD in MBSE remains difficult because of data management and updating complexity. Whereas QFD is a complementary approach for MBSE deployment. It is well known by many design teams, it supports multi-actor design sessions and provides a combined view of complementary project features (VoC, technical specification, competitors positioning ...).

This paper focuses on developing tools and methods to facilitate the implementation of QFD within MBSE using SysML language through dynamic exchange of data. It focuses on house of quality phase (HoQ) of the QFD approach that tackles the systems engineering activities related to the analysis and the definition of system requirements. To reach this goal, a meta-model is created to define a unified structure of shared data between MBSE activities and HoQ analysis. In this manner, the data exchange between HoQ analysis MBSE models can be performed. Then, a framework to dynamically update data between HoQ analysis and SysML Diagrams is presented. Examples of the framework utilization are finally given to show the various interaction that can be supported by such a connection between QFD and MBSE practices. This approach allows to keep the QFD representation for the interdisciplinary work within the design team while maintaining MBSE advantages with the rigor of data management and tracking.

In future work, it is of interest to extend the proposed method to the remaining phase of QFD approach such as process design and process control. Moreover, QFD analysis is an approach to justify the traceability and compliance between customer needs and already defined system requirements considering external constraints. Therefore, it is interesting to develop extended method including rationale analysis to define system requirements and QFD to check the compliance. Additionally, validation works should be conducted to validate the proposed metamodel. Running case studies and ensuring that the data to be expressed is captured would allow this validation.

6 REFERENCES

- Benner, M., Linnemann, A. R., Jongen, W. M. F., & Folstar, P. (2003). Quality Function Deployment (QFD)—can it be used to develop food products? *Food Quality and Preference*, 14(4), 327–339.
- Carnevali, J. A., & Miguel, P. C. (2008). Review, analysis and classification of the literature on QFD-Types of research, difficulties and benefits. *International Journal of Production Economics*, 114(2), 737–754.
- Chou, S. (2004). Evaluating the service quality of undergraduate nursing education in Taiwan – using quality function deployment. *Nurse Education Today*, 24(4), 310–318.
- David, P., Idasiak, V., & Kratz, F. (2010). Reliability study of complex physical systems using

- SysML. *Reliability Engineering & System Safety*, 95(4), 431-450.
- Friedenthal, S., Moore, A., & Steiner, R. (2012). *A practical guide to SysML: the systems modeling language*. Morgan Kaufmann.
- Ginn, D., & Zairi, M. (2005). Best practice QFD application: an internal/external benchmarking approach based on Ford Motors' experience. *International Journal of Quality & Reliability Management*, 22(1), 38–58.
- International Council on Systems Engineering (INCOSE). (2015). *Systems Engineering Handbook*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- IEC 15288 (2002). ISO/IEC 15288:2002 Systems engineering -- System life cycle processes,
- ISO 16355. (2015). Application of statistical and related methods to new technology and product development process — Part 1: General principles and perspectives of Quality Function Deployment (QFD), 83.
- Liu, J., Wang, S., & Fu, C. (2017). Design Analysis Method for Multidisciplinary Complex Product using SysML. *MATEC Web of Conferences*, 139, 14.
- Lowe, A., & Ridgway, K. (2000). UK user's guide to quality function deployment. *Engineering Management Journal*, 10(3), 147.
- Moubachir, Y., & Bouami, D. (2015). A new approach for the transition between QFD phases. *Procedia CIRP*, 26, 82–86.
- Shiple, M. F., de Korvin, A., & Yoon, J.-M. (2004). Fuzzy quality function deployment: determining the distributions of effort dedicated to technical change. *International Transactions in Operational Research*, 11(3), 293–307.
- Singh, S., & Kumar, M. (2015). Product development through QFD analysis using Analytical Network Process. *International Journal of Advanced Engineering Research and Applications (IJAERA)*, 1(3), 112–126.
- Verma, D., Chilakapati, R., Fabrycky, W. J., & Professor Emeritus, L. (1998). Analyzing a Quality Function Deployment (QFD) Matrix: An Expert System Based Approach to Identify Inconsistencies and Opportunities. *Journal of Engineering Design*, 9(3), 252-262.
- Yan, W., Khoo, L. P., & Chen, C.-H. (2005). A QFD-enabled product conceptualisation approach via design knowledge hierarchy and RCE neural network. *Knowledge-Based Systems*, 18(6), 279-293.

CIGI 2021

A Computer-Aided Verification Process for Engineered Systems

VICTOR ROMERO¹, ROMAIN PINQUIE¹, FREDERIC NOEL¹

¹ UNIV. GRENOBLE ALPES, CNRS, GRENOBLE INP*, G-SCOP, 38000 GRENOBLE, FRANCE

* INSTITUTE OF ENGINEERING UNIV. GRENOBLE ALPES
46 avenue Félix Viallet, 38 000 GRENOBLE, France
{victor.romero, romain.pinque, frederic.noel}@grenoble-inp.fr

Résumé – Les acteurs de l’industrie du futur essayent de systématiquement numériser leurs produits et leurs processus. La numérisation systématique des objets techniques du bureau d’études reste néanmoins très inégalitaire. En effet, les outils qui supportent la conception détaillée sont plus sophistiqués que les outils de spécification ou d’architecture. Notons tout de même la récente évolution d’une ingénierie système basée sur les modèles plutôt que sur les documents. Cependant, les modèles assimilables à des diagrammes 2D inertes, ne facilitent pas assez la détection des erreurs de spécification (validation) et de conception (vérification) en amont des tests d’intégration et des essais de qualification. Nous proposons alors un processus holistique de vérification assistée par ordinateur. Celui-ci permet de s’assurer qu’une solution de conception satisfait ou non aux exigences, et ce, dans des conditions opérationnelles dont l’amélioration progressive du niveau de réalisme assure une continuité entre le système virtuel et le système réel. Notre processus de vérification assistée par ordinateur repose, d’une part, sur la cosimulation d’un modèle de spécification et d’un modèle de conception et, d’autre part, sur une immersion dans une scène virtuelle qui permet de mieux apprécier les propriétés du système, et ce, avant de progressivement substituer le virtuel par le réel au moyen de simulations Hardware-In-the-Loop.

Abstract – Although companies systematically strive for a full digitalisation of their products and their processes, the design phase shows that the quality of models is very unequal. Indeed, detailed design benefit from much more sophisticated methods and tools than the specification and architecture activities. However, we should note the recent paradigm shift from document-based to model-based system engineering. However, these models, which are mainly static 2D diagrams, remain too poor to facilitate the early detection of design errors. Thus, to detect most errors that occur during the design phase, companies have no other alternative than to wait up to the testing phase which occurs after several years for complex systems. Thus, we propose a holistic user-centred computer-aided verification process to ensure that the design meets the requirements under realistic operational conditions. The verification process provides a progressive immersion into the virtual system before seamlessly transitioning from the virtual to the real system. Our work relies on state-of-the-art MBSE methods such as the Property Model Methodology, which enables systems engineers to co-simulate specification models and design models. We improve such MBSE methods by increasing the level of realism that experiences the end-user during the verification of a design by the original combination of Model-In-the-Loop, Immersive Model-In-the-Loop, Human-In-the-Loop, and Hardware-In-the-Loop simulation strategies.

Mots clés – Ingénierie Système, Vérification, Réalité Virtuelle, Modélisation et Simulation, Retours Haptiques

Keywords – Model-Based Systems Engineering, Verification, Virtual Reality, Human-In-the-Loop Simulation

1 INTRODUCTION

1.1 Context

In the past decades, the design of increasingly complex engineered systems forces the designers to replace the traditional document-based systems engineering approach with Model-Based Systems Engineering (MBSE). In its systems engineering vision 2020 (Crisp, 2007), the International Council On Systems Engineering (INCOSE) defined Model-Based Systems Engineering (MBSE) as “*the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases*”. To date,

we perceive the motivation for engineering systems using models through two different lenses.

On the one hand, some practitioners argue that the use of models to support an engineering activity is an age-old solution since models have been used for decades to engineer technological systems. Indeed, the roots of MBSE can be traced back to the pioneer work achieved by Wayne Wymore. In his book entitled “*Model-Based Systems Engineering*” (Wymore, 1993), Wymore has shown the crucial role of mathematics in the specification and design of systems. By theorising systems engineering concepts, he paved the way for MBSE, but his visionary mathematical theory of systems engineering was so far ahead of systems engineers that it took several decades before his contribution was recognized (Bahill, 2011).

On the other hand, the model-based approach is in the limelight today because there is another community of engineers who see MBSE as the salvation of systems engineering. The intellectual Renaissance of MBSE results from academic researchers and industrialists who have worked out a set of theoretical and methodological concepts related to systems engineering, a branch that originally was more an irrational practice rather than a rigorous method (Suh, 2001). MBSE rests on a four-activity process: specification, validation, design, and verification (INCOSE, 2015, sec. 9.2.1). Our study focuses on the verification activity. It is important not to confuse verification that intends to ensure that the “product is built right” and validation that intends to ensure that the “right product is built” (INCOSE, 2015). According to the INCOSE, there exists four verification techniques: inspection, demonstration, testing, and analysis. Here we concentrate on the analysis.

“Analysis – This technique is based on analytical evidence obtained without any intervention on the submitted element, using mathematical or probabilistic calculation, logical reasoning (including the theory of predicates), modelling, and/or simulation under defined conditions to show theoretical compliance. Mainly used where testing to realistic conditions cannot be achieved or is not cost-effective.”

The recent development of cutting-edge computational capabilities has opened new vistas for the analytical verification of design solutions (Hoppe, Engel, & Shachar, 2007). Carried out early in the design process, analytical verification using executable models enables companies to drastically reduce costs and time-to-market while improving quality (Becquet et al., 2018; Fabre, Micouin, Gaurel, & Pandolfi, 2020).

Nevertheless, when we observe the verification activity in the industry, we notice that there are several – more or less advanced – practices including, but not limited to, the sharing of 2D captures of models, dynamic co-simulation of specification and design models, immersive 3D design reviews, hardware-in-the-loop simulation, physical testing, etc., but there is no holistic computer-aided verification process that aims at logically articulating them.

1.2 Problem(s)

Document-Based System Engineering (DBSE) is limited for the design of complex engineered systems. Indeed, the content of documents is difficult to maintain, synchronise, and access (INCOSE, 2015). When looking for design errors during the verification activity, engineers do not have other alternatives than to get together for a pragmatic design review where they share screenshots of their models (SysML, CAD, physics-based simulation, testing reports, etc.) supporting a relatively subjective argumentation process that aims at justifying the satisfaction of requirements stored in a large and ambiguous document-based specification.

To make the verification activity more objective, recent MBSE methods propose to verify whether the design meets the requirements by the co-simulation of design models and specification models (Becquet et al., 2018). Requirements are not natural language statements but logical assertions that monitor the state of design variables in near real-time. Although these new MBSE methods and tools are essentials for the early detection of design errors, the lack of visualisation

and interactions of MBSE virtual environments do not enable stakeholders to experience the virtual system in realistic conditions (Figure 1).

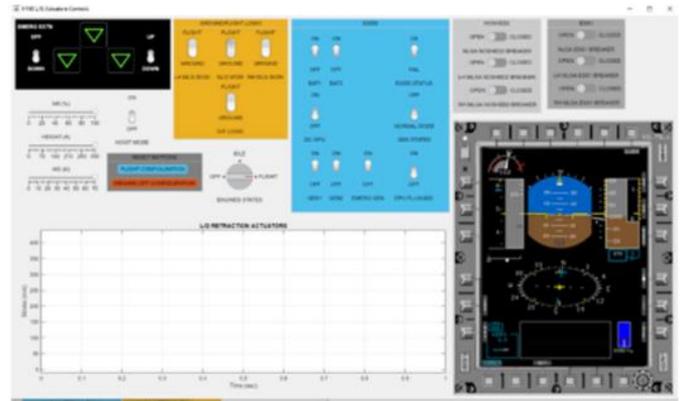


Figure 1: Visualisation virtual cockpit in the MathWorks suite (Fabre et al., 2020)

There are numerous ways to experience a virtual system in realistic conditions, but there is no integrated and continuous verification strategy. For instance, to increase the appreciation of structural properties, we can immerse stakeholders in an immersive environment using Virtual Reality (VR) (Han & Black, 2011). However, the isolation of the specification and the design also leads to a subjective design review. Indeed, designers look for potential design errors in immersion without a reliable strategy since requirements are stored in external ambiguous documents. Regarding the behavioural properties of a system, we can replace models of sub-systems with real components and carry out Hardware-In-the-Loop simulations to get rid of some modelling assumptions and therefore be more realistic (Sarhadi & Yousefpour, 2015), but the specifications remain isolated from the design too.

Those limits motivate us to ask the question:

How to logically combine advanced V&V technologies (MBSE, Simulation, VR, HIL, etc.) to perform an objective verification with a user-centred process, in realistic conditions?

1.3 Proposal

To connect the dots, we propose a holistic computer-aided verification process for continuously verifying that the design satisfies the requirements with objective evidence in realistic conditions. Our computer-aided verification process is also a continuous model-centric integration strategy from the specification phase to the first prototype.

This process relies on the use of virtual reality to visualise and interact with the simulation, to improve the quality of the visualisation and add natural interactions with the simulated system adding then operational randomness.

2 LITERATURE REVIEW

In this section, we start with a fundamental definition of the verification activity and briefly introduce model- and simulation-based systems engineering methods that will serve as a starting point for our research study.

2.1 Verification

According to the main systems engineering standards: “The purpose of the verification process is to provide objective evidence that a system or a system element fulfils its specified requirements and characteristics” (ISO/IEC/IEEE 15288, [6.4.9.1]) (INCOSE, 2015).

Except for analysis, all the verification types require a physical prototype. Building a physical prototype is excessively expensive for complex systems. To pass by this issue, the MBSE approach emerged in the past decades.

The starting point of the verification activity is the formalisation of the requirements, to allow an objective comparison between the actual system and what is expected. For this purpose, standardised graphical annotations such as SysML (Karban, Hauber, & Weilkens, 2015; OMG, 2017) and Capella (Roques, 2016) help to develop and manage the requirements but do not natively provide any formal validation and verification capabilities. However, the next section will discuss recent studies that focus on the use of simulation software to formally model computable requirements that fulfil a monitoring function. Tight integration of specification models and design models in a co-simulation environment helps to identify design errors early on.

2.2 Model- and Simulation-Based Verification

Engineers use modelling and simulation technologies for verifying complex engineered systems in a very subjective way. Indeed, as Suh (Suh, 2001) reports “Often designers find that the precise description of “what we want to achieve” is a difficult task. Many designers deliberately leave them implicit rather than explicit and then, start working on design solutions even before they have clearly defined their design goals. They measure their success by comparing their design with the implicit design goals they had in mind. They spend a great deal of time to improve and iterate the design until the design solution and “what they had in mind” converge, which is a time-consuming process at best.” (Suh, 2001). Today, the most common verification approach consists in modelling and simulating the properties of the system and organising design reviews to approve or reject a design candidate based on the results of the simulations without a rigorous definition of the requirements and their integration with design solutions.

To satisfy Suh’s recommendation “To be efficient and to generate the design that meets the perceived needs, we must specifically state the design goals in terms of “what we want to achieve” and begin the design process.” (Suh, 2001), recent MBSE methods and tools propose formal specification and verification activities. For instance, Micouin proposed the Property Model Methodology (PMM) (Micouin, 2013). PMM was successfully applied to an academic case study using Modelica (Pinquie et al., 2016) before to be evaluated on an operational Airbus Helicopter program (Becquet et al., 2018; Fabre et al., 2020). The PMM process is based on formally defined requirements using the concept of Property-Based Requirements (PBR) (Micouin, 2010). After the specification, validation, and design activities comes the verification thanks to the co-simulation of design models and specification models. PBR can be supported by various simulation language such as Modelica (Buffoni & Fritzon, 2016; Nguyen, 2014; Otter et al., 2015) or the MathWorks suite (Nilsson, 2014).

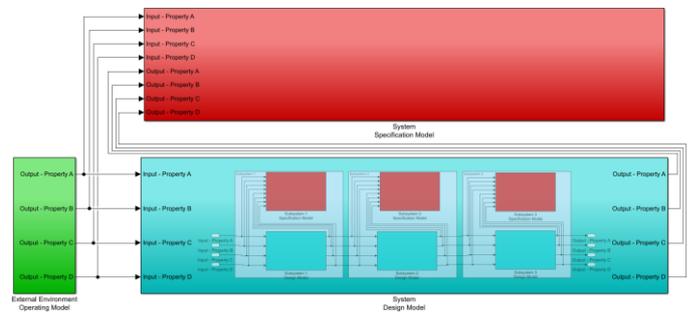


Figure 2: With the Property Model Methodology, the model- and simulation-based verification activity requires: the top-down modelling of (1) the system specification model (block in red), (2) the system design model (block in blue), (3) the sub-systems specification models (blocks in red within the design model), (4) the sub-systems’ design models (block in blue within the design model), (5) the operating scenarios that stimulates the specification and design models (block in green).

The limits of these model- and simulation-based verification approaches pointed out during the introduction are the weak visualisation and interaction capabilities of the physics-based simulation tools that support the method, and which lead to a low level of realism for the end-user experience.

2.3 Virtual Reality-Based Verification

Casenave & Lugo (2018) demonstrated that Virtual Reality technologies, which support the 3D representation of a system, give much more natural access to the product for the end-user, improving the understanding of the data. This encourages the industries to adopt these technologies for various applications including the verification activity (Berg & Vance, 2017; PWC, 2015), but the experience is still limited to the inspection-like verification. Indeed, the immersive environments increase the end-user experience when analysing structural properties, but they hardly integrate a physics-based simulation for emulating the behavioural properties and do not integrate formal specification models. Therefore, the design remains separated from the specification.

The literature review shows that there are various methods and tools for verifying designs early on. However, there is no holistic computer-aided verification process that logically integrates advanced practices.

3 A HOLISTIC COMPUTER-AIDED VERIFICATION PROCESS

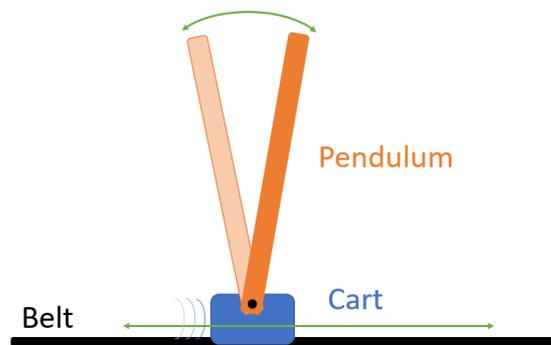


Figure 3: Inverted Pendulum

To illustrate our computer-aided verification process, we will use an inverted pendulum (Figure 3), a system that is not

complex but sufficiently rich to illustrate our proposal as it contains five multi-engineering subsystems (Figure 4).

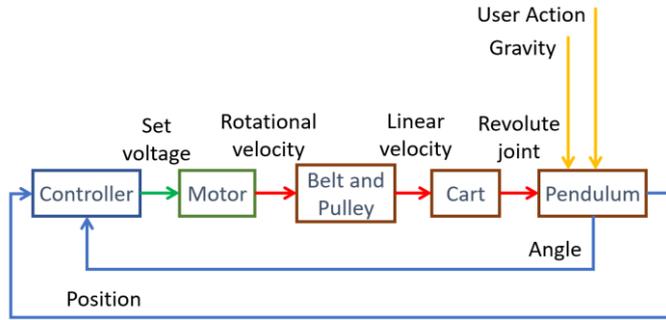


Figure 4: Architecture of the inverted pendulum

We assume that an efficient computer-aided verification process (Figure 5), which provides increasing analytical evidence in realistic conditions, starts with a state-of-the-art analytical verification activity: the co-simulation of design models with specification models – **Model-In-the-Loop Verification**.

The second activity of the computer-aided verification process consists of immersing the end-user in a virtual world to improve his experience in terms of visualisation – **Immersive Model-In-the-Loop Verification**.

Always with the desire to improve the degree of realism, especially the interactions with the virtual world, the third activity integrates the human to realistically stimulate the models in near real-time. – **Immersive Model- and Human-In-the-Loop Verification**.

The fourth activity removes modelling assumptions by a continuous substitution of virtual building blocks from the design model by real ones up to the first real prototype of the system - **Immersive Model-, Human- and Hardware-In-the-Loop Verification**.

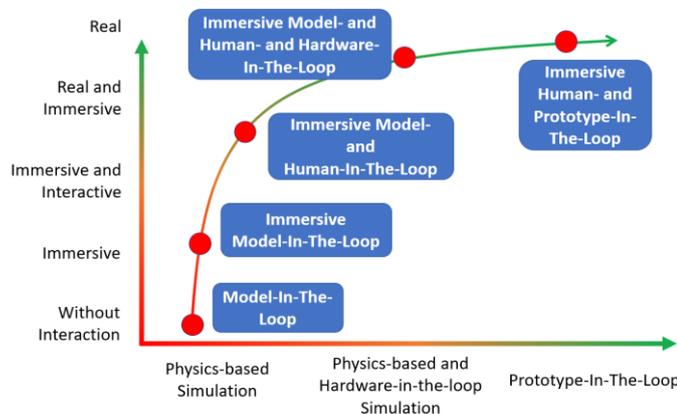


Figure 5: A continuous and holistic computer-aided verification process from virtuality to reality

3.1 Model-In-The-Loop Verification

3.1.1 Definition

We start from the state-of-the-art MBSE methods by using the Property Model Methodology (Micouin, 2013) for its Model-In-the-Loop verification strategy implemented with the MathWorks suite (Becquet et al., 2018; Fabre et al., 2020). Model-In-the-Loop verification consists of the co-simulation of a formal specification model and a discrete and/or continuous design model. The specification model monitors the design models (Figure 6). We assume that these models are

already created in a previous design step and we expect to check that the design matches the specifications. Currently, our process starts at this design stage.

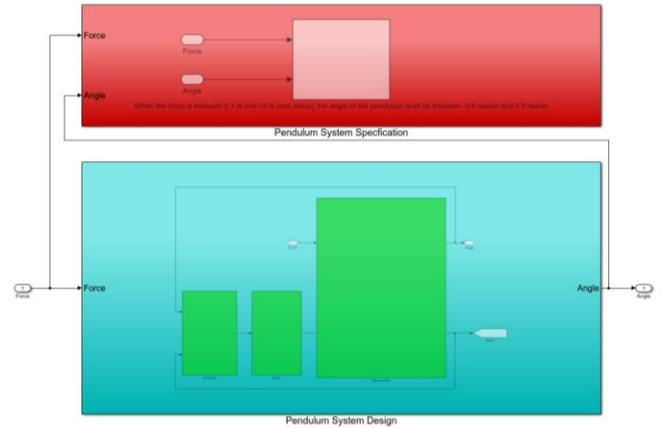


Figure 6: The specification model (in red) monitors the design model (in blue for the system and in green for the sub-systems)

3.1.2 Implementation

The quality of the implementation depends on the capabilities of software that supports the PMM method. We use the MathWorks suite to co-simulate our specification models and our design models because past experiences (Fabre et al., 2020; Pinquie et al., 2016) demonstrated that its maturity outperforms existing alternatives such as Modelica-based solutions. As an example, Figure 7 shows a specification model (in orange) containing a formal requirement. The inputs of the specification model correspond to the external stimulus – i.e. the force applied to the pendulum – and the intended effect – i.e. the angular position of the pendulum. The specification, which is simplified to a single requirement here, makes sure that the design behaves as expected under specific conditions, that is, the pendulum remains in equilibrium if the magnitude of the force acting on the pendulum belongs to the specified interval. Here we assume that the equilibrium state corresponds to an angular position between -0,5 and 0,5 radian. Figure 7 show that the specification of the system is derived into sub-systems specifications by satisfying the prime contractor theorem which makes sure that sub-systems requirements are equally or more constraining than the systems requirements.

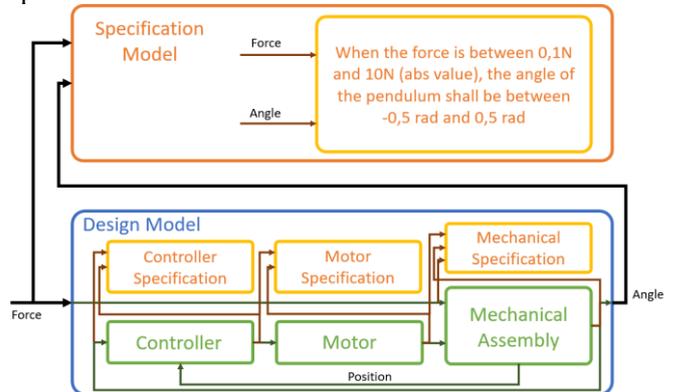


Figure 7: Representation of the Model-In-The-Loop verification in the MathWorks suite.

3.1.3 Limits

The modelling scheme of Model-In-the-Loop verification is limited to predefined operational scenarios, whereas sound

verification requires extensively to stimulate the system. There is therefore a need to define numerous and heterogeneous operational scenarios to make sure the design meets the requirements. The definition of operational scenarios is limited to the nominal and well-known critical scenarios while missing potential unintended situations. Furthermore, the visuals are graphs, tables, or 2/3D animations, whereas simulation outputs require an intuitive visualisation thanks to immersive visualisation techniques such as virtual reality (Han & Black, 2011; Pausch, Proffitt, & Williams, 1997). By integrating virtual reality into PMM, the level of realism would be higher because of the improvement of many visualisation performance criteria, including but not limited to stereoscopy, field of view, display size and resolution, frames per second (Bowman & McMahan, 2007).

3.2 Immersive Model-In-The-Loop Verification

3.2.1 Definition

Reminding that our goal is to be closer to a test in *realistic conditions*, we expect to improve the realism of the Model-in-the-Loop verification with visual immersion. Thus, to perform a much more intuitive visualisation of the system properties, we conserve the co-simulation of specification models and design models, but we replace the native visualisation capabilities of the MathWorks suite with an immersive environment giving birth to the Immersive Model-in-the-Loop verification strategy (Figure 8).

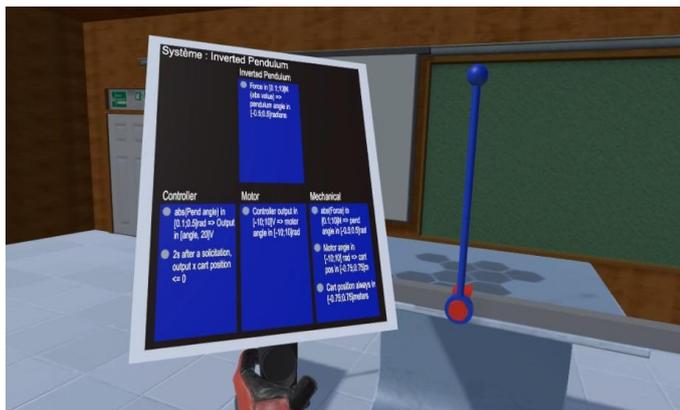


Figure 8: User view of the inverted pendulum in a virtual environment and the requirements on a pad.

3.2.2 Implementation in the use case

Immersive Model-in-the-Loop verification requires a VR engine to communicate with the Model-In-the-Loop simulation. This interconnection uses UDP network communication because there is no verification from the server of the data received by the client leading to faster data exchanges. The VR visualisation module was developed within the Unity3D framework (Unity Technologies, 2020). We use the Simulink block SendUDP to send the position of the cart and the rotation of the pendulum around the cart axis to another computer in which Unity3D is running (Figure 9).

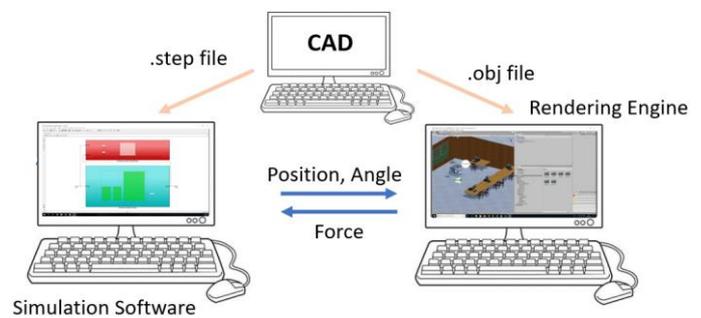


Figure 9: Immersive Model-In-The-Loop verification

In this second software, to animate the system in the virtual environment created in the Unity3D graphic window, we can open the appropriate ports and read the received data thanks to C# scripts. Regarding the visual metaphors for the specification model, we use a pad that the user activates and deactivates using a key of the left-hand remote controller. This pad contains a representation of the system architecture with the system at the top and the sub-systems at the bottom (Figure 10). Each system-of-interest contains a textual representation of the requirements contained in its specification model. The virtual lamp that stands next to each requirement turns green when the requirements are tested and satisfied, red when the requirements are tested and unsatisfied, and remains grey when it is untested.

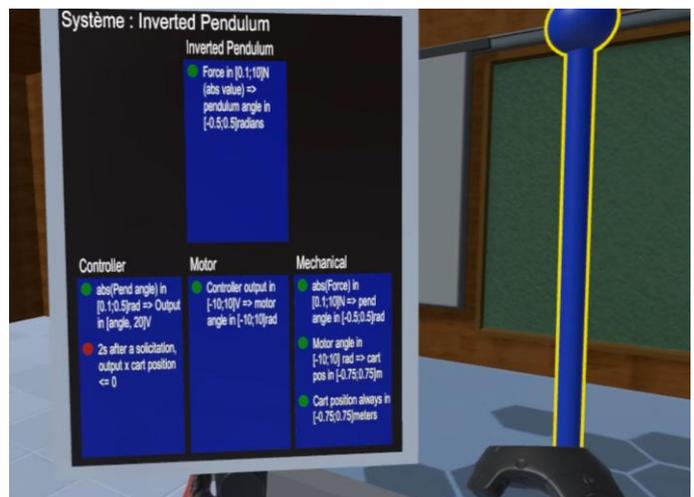


Figure 10: User view of interaction with the pendulum (highlighted when touching with the user's virtual hand)

3.2.3 Limits

Immersive Model-in-the-Loop improves the visual appreciation of the structural and behavioural properties. However, we still have a lack of realism in the interactions provided by this simulation. Indeed, the user cannot directly interact with the system limiting the virtual experience to planned (models or recordings embedded in the behavioural model) operational scenarios without unintended events.

3.3 Immersive Model- and Human-In-The-Loop Verification

3.3.1 Definition

In the last section, we improved the realism during the verification activity, but VR must not be restricted to 3D visual perception as it offers advanced interactive capabilities. The integration of Human-In-the-Loop simulation into the Immersive Model-In-the-Loop Verification aims at providing

the end-user with the opportunity to naturally operate the system. We name this new verification strategy Immersive Model- and Human-In-the-Loop verification (Figure 10, Figure 11 and Figure 12).

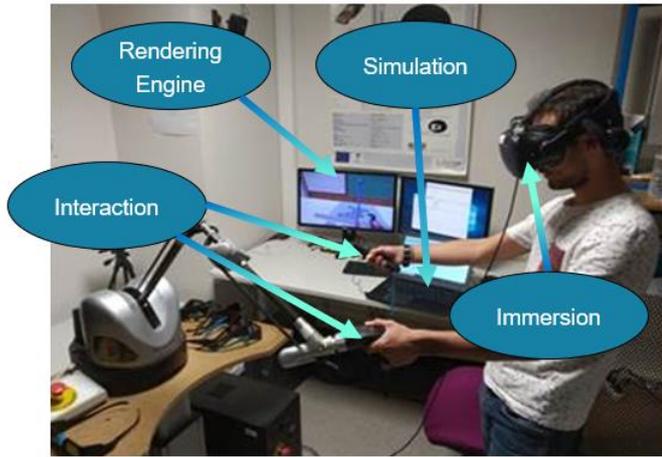


Figure 11: Example of an Immersive Model- and Human in the loop verification

3.3.2 Implementation in the use case

To achieve an Immersive Model- and Human-In-the-Loop verification, we need changes in the simulation and visualisation software. The simulation will have a new input standing for the end-user stimulus. We can use the same communication protocol as for the visualisation but in the other direction: the VR software sends the interaction data to the physics-based simulation through the communication network using a UDP protocol and the simulation software receives it and uses it as an input of the system. These interactions are captured by the remote controllers provided with the HMD HTC Vive. We also use a haptic arm, the Virtuoso 6D 35-45, which captures the position of the hand and gives force feedback when the user touches the virtual pendulum. When providing significant force feedback, the user develops a certain apprehensiveness, not to say fear, to touch the pendulum again leading to an increase of the level of presence, that is, the sense of “being in” the scene and the user starts behaving as he was experiencing the system in the real world.

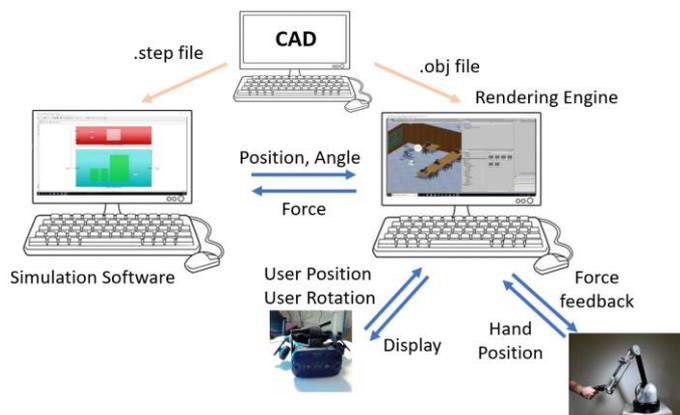


Figure 12: Immersive Model- and Human-In-The-Loop verification including haptic feedbacks

3.3.3 Limits

The natural interaction with realistic behaviour and structural properties improves the early verification activity. However, as the system is fully virtual, modelling assumptions may lead to unintended events.

3.1 Immersive Model-, Human-, and Hardware In-The-Loop Verification

3.1.1 Definition

Immersive Model, Human- and Hardware-In-the-Loop Verification consists of substituting virtual building blocks from the design model by real ones up to the first real prototype instrumented with sensors that feed the system specification model. The simulation of the system under design embedding real components enable designers to get rid of modelling assumptions leading to potential design errors.

3.1.2 Implementation in the use case

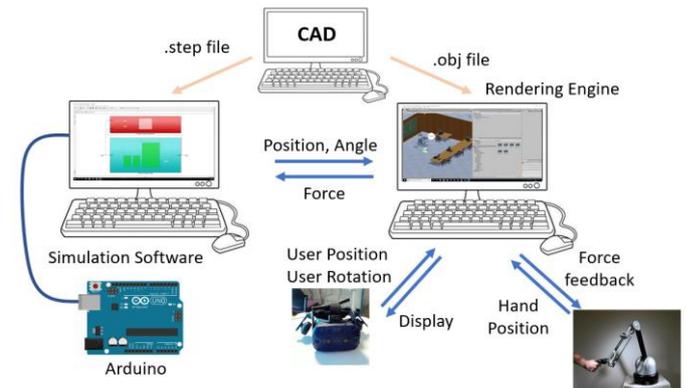


Figure 13: Immersive Hardware- and Human- and Model-In-The-Loop verification

In our case study, we can start replacing the design model of the “Controller” sub-system (Figure 4) with the Arduino Uno card (Figure 13). With this strategy, we can proceed to a step-by-step verification of the real components before integrating them together up to the verification of the full prototype instrumented with sensors (Figure 14).

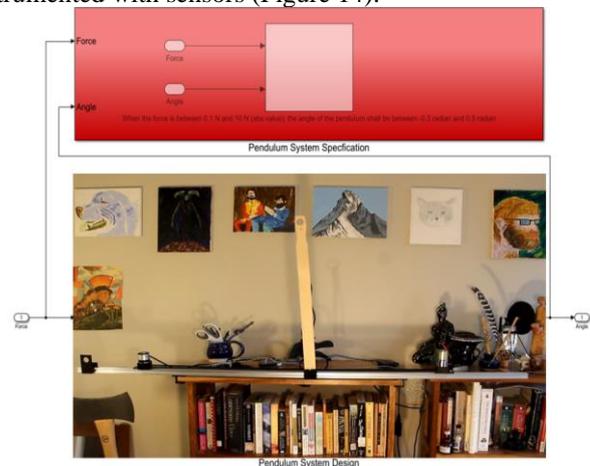


Figure 14: Verification of the first functional prototype with the system specification model in the loop

4 CONCLUSION

Model-based requirements and systems engineering, virtual reality, physics-based simulation, and hardware-in-the-loop simulations are key technologies for the early verification of engineered systems. However, there was a need for continuous integration of these technologies into a holistic computer-aided verification process for verifying engineered systems in realistic conditions.

We demonstrated that it is possible and relevant to integrate VR technologies with an MBSE method to recreate realistic operational conditions while keeping the added-value of the

co-simulation of specification models and design models. Also, following the growing maturity of the designed system by progressively substituting the virtual building blocks with real ones, the designer can be confident that the structural and behavioural properties of the first prototype will satisfy the requirements, avoiding then extra costs due to design errors.

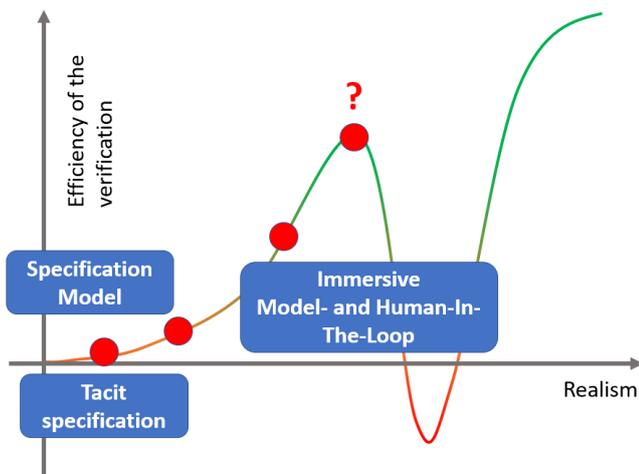


Figure 15: An attempt to conceptually define the uncanny valley of immersive verification

As future works, we plan to develop a reporting dashboard in Unity 3D containing indicators related to the verification activity. We will also test the proposed computer-aided verification process in an industrial context. Moreover, analytical verifications strategies strive to integrate numerous technologies and develop immersive virtual environments to experience more and more realistic conditions. Nevertheless, we hardly know: how much realism is enough to verify a design? We may therefore carry out new research studies to derive sound and parsimonious recommendations that would help companies to reach the right level of realism according to the verification goals and the available resources.

To provide more objective criteria, we may borrow the phenomenon of the uncanny valley suggested by the researches applied to humanoids (Mathur & Reichling, 2016). The uncanny valley translates the unexpectedly negative dislike reactions provoked by imperfect human-likeness in human robots (Mori, 1970). By analogy, we may assume that more sensorial feedback does not systematically increase the level of realism. For example, studies show that is the case for haptics feedback (Berger, Gonzalez-Franco, Ofek, & Hinckley, 2018). We may therefore assume that, at a certain point, more realism does not automatically increase the efficiency of the verification process. Figure 15 illustrate a new research hypothesis stating that this phenomenon may exist in our virtual environment for systems verification. Therefore, as future work, we can try to demonstrate the existence or not of the uncanny valley phenomenon in the virtual phases of our computer-aided verification process to finally be able to characterise the “right need” of the user, that is, the perfect balance between the time and cost spent for developing a realistic virtual environment and its efficiency in terms of verification (e.g. the number of design errors identified).

5 REFERENCES

- Bahill, T. (2011). In Memoriam A. Wayne Wymore. *INSIGHT*, 14(2), 57–61. <https://doi.org/10.1002/inst.201114257>
- Becquet, R., Micouin, P., Fabre, L., Guérin, F., Paper, P., & Razafimahefa, T. (2018). Property Model Methodology: A Landing Gear Operational Use Case. *INCOSE International Symposium*, 28(1), 321–336. <https://doi.org/10.1002/j.2334-5837.2018.00484.x>
- Berg, L. P., & Vance, J. M. (2017). Industry use of virtual reality in product design and manufacturing: a survey. *Virtual Reality*, 21(1). <https://doi.org/10.1007/s10055-016-0293-9>
- Berger, C. C., Gonzalez-Franco, M., Ofek, E., & Hinckley, K. (2018). The uncanny valley of haptics. *Science Robotics*, 3(17), eaar7010. <https://doi.org/10.1126/scirobotics.aar7010>
- Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7), 36–43. <https://doi.org/10.1109/MC.2007.257>
- Buffoni, L., & Fritzson, P. (2016). Expressing Requirements in Modelica. *SNE Simulation Notes Europe*, 25(3–4), 21–22. <https://doi.org/10.11128/sne.25.tn.10314>
- Crisp, H. (2007). *INCOSE systems engineering vision 2020*.
- de Casenave, L., & Lugo, J. E. (2018). Effects of Immersion on Virtual Reality Prototype Design Reviews of Mechanical Assemblies. *Volume 7: 30th International Conference on Design Theory and Methodology*, 7, 1–11. <https://doi.org/10.1115/DETC2018-85542>
- Fabre, L., Micouin, P., Gaurel, C., & Pandolfi, P. (2020). *Advances in Property Model Methodology (PMM®)*.
- Han, I., & Black, J. B. (2011). Incorporating haptic feedback in simulation for learning physics. *Computers and Education*, 57(4), 2281–2290. <https://doi.org/10.1016/j.compedu.2011.06.012>
- Hoppe, M., Engel, A., & Shachar, S. (2007). SysTest: Improving the verification, validation, and testing process—Assessing six industrial pilot projects. *Systems Engineering*, 10(4), 323–347. <https://doi.org/10.1002/sys.20082>
- INCOSE. (2015). *Systems engineering handbook A guide for system life cycle processes and activities*.
- Karban, R., Hauber, R., & Weilkiens, T. (2015). MBSE in Telescope Modeling. *INSIGHT*, 12(4), 24–31. <https://doi.org/10.1002/inst.200912424>
- Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22–32. <https://doi.org/10.1016/j.cognition.2015.09.008>
- Micouin, P. (2010). Toward a Property Based Requirements Theory: System Requirements Structured as a Semilattice. *Wiley InterScience*, 14(3), 305–326. <https://doi.org/10.1002/sys>
- Micouin, P. (2013). Property-Model Methodology: A Model-Based Systems Engineering Approach Using VHDL-AMS. *Wiley Online Library*, 14(3), 305–326. <https://doi.org/10.1002/sys>
- Mori, M. (1970). The uncanny valley. *Energy*, 7, 33–35.
- Nguyen, T. (2014). FORM-L: A MODELICA Extension for Properties Modelling Illustrated on a Practical Example. *Proceedings of the 10th International Modelica Conference, March 10-12, 2014, Lund, Sweden*, 96, 1227–1236. <https://doi.org/10.3384/ecp140961227>
- Nilsson, V. (2014). *Model-Based Testing with Simulink Design Verifier*. Chalmers University of Technology.
- OMG. (2017). OMG SysML Home | OMG Systems Modeling Language. Retrieved from <http://www.omgsysml.org/>
- Otter, M., Thuy, N., Bouskela, D., Buffoni, L., Elmqvist, H., Fritzson, P., ... Tundis, A. (2015). Formal Requirements

- Modeling for Simulation-Based Verification. *Proceedings of the 11th International Modelica Conference, Versailles, France, September 21-23, 2015*, 118, 625–635. <https://doi.org/10.3384/ecp15118625>
- Pausch, R., Proffitt, D., & Williams, G. (1997). Quantifying immersion in virtual reality. *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '97*, (August 1997), 13–18. <https://doi.org/10.1145/258734.258744>
- Pinquie, R., Micouin, P., Veron, P., Segonds, F., Pinquié, R., Véron, P., ... Segonds, F. (2016). Property Model Methodology : A Case Study with Modelica. *Proceedings of Tools and Methods for Competitive Engineering*, (April), 1–12. <https://doi.org/10.13140/RG.2.1.1925.1603>
- PWC. (2015). For US manufacturing, virtual reality is for real. Retrieved May 23, 2019, from 2015 Disruptive Manufacturing Innovations Survey website: <https://www.pwc.com/us/en/industrial-products/next-manufacturing/augmented-virtual-reality-manufacturing.html>
- Roques, P. (2016). MBSE with the ARCADIA Method and the Capella Tool. *8th European Congress on Embedded Real Time Software and Systems (ERTS 2016)*.
- Sarhadi, P., & Yousefpour, S. (2015). State of the art: hardware in the loop modeling and simulation with its applications in design, development and implementation of system and control software. *International Journal of Dynamics and Control*, 3(4), 470–479. <https://doi.org/10.1007/s40435-014-0108-3>
- Suh, N. P. (2001). *Axiomatic design : advances and applications*. Oxford University Press.
- Unity Technologies. (2020). Multiplatform - VR-AR. Retrieved March 10, 2020, from <https://unity3d.com/fr/unity/features/multiplatform/vr-ar>
- Wymore, A. (1993). *Model-based systems engineering*.

CIGI 2021

DGA – V&V anticipée pour les SDS

FABIENNE HAVY, ARNAUD BERNIER

¹ DGA MAITRISE DE L'INFORMATION
BP 7 35998 Rennes Cedex 9, France

Application dans le cadre du Programme de transformation de l'Armée de Terre, d'une approche optimisée et anticipée des opérations de Vérification et Validation :

Méthodes / Outils

- Utilisation de la modélisation Système de Systèmes pour construire des scénarios de validation bout en bout
- Développement d'un outil interfacé avec MEGA HOPEX, pour répartir les scénarios d'essais Système de Systèmes et les actions associées à des participants
- Apparition de la notion de méta système, permettant de regrouper de manière opportuniste des systèmes constituants, parties prenantes des scénarios de validation

Techniques

- Conception et mise en place d'une plateforme technique, permettant la virtualisation du SI du programme SCORPION (Synergie du COntact Renforcé par la Polyvalence et l'InfovalorisatiON)
- Interfaçage avec les systèmes contributeurs (notamment de moyens de communication Radio)
- Simulation d'un GTIA (Groupement Tactique Inter Armes) (400 véhicules) en environnement virtualisé
- Dé-risquage des opérations d'expérimentation technique sur le terrain, par déroulement en avance de phase des scénarios bout en bout
- Développement d'un Système dédié à la collecte, et l'exploitation, des résultats de tests a posteriori

Mots clés – Modélisation et Validation / Simulation / Virtualisation / Collaboratif.

Within the framework of the Army Transformation Program, this article presents an optimized and anticipated approach for Verification and Validation activities:

Methodologies and Tools

- Use of System of Systems Model, to build end-to-end validation scenarii
- Development of a tool interfaced with MEGA HOPEX, to distribute SDS test scenarii and associated actions with participants
- Emergence of the « meta system » concept, in order to group in an opportunistic manner some of the component systems, stakeholders of validation scenarii

Technical

- Design and implementation of a technical platform, allowing the virtualization of the SCORPION IS
- Interfacing with contributing systems (including radio communication)
- Simulation of an Inter-Arms Battle Group (400 vehicles) in a virtualized environment
- De-risking of real life technical experimentation operations, by processing ahead of time the end-to-end scenarii
- Development of a related information system, tasked with collection and exploitation of test results

Keywords – Model and Validation / Simulation / Virtualization / Collaborative work.

1 INTRODUCTION

La Direction Générale de l'Armement, a entamé fin 2018 une transformation d'envergure de son fonctionnement, et notamment des méthodes de conduite des Opérations d'Armement.

La démarche de V&V présentée dans cet article s'inscrit dans cette transformation, et a pour objectif de concrétiser ses enjeux stratégiques dans la conduite des Programmes d'Armement :

- Anticipation au plus tôt des opérations de V&V
- Passerelle [modélisation → V&V]
- Apport de la simulation / virtualisation dans les opérations de V&V
- Dérisquage en amont des expérimentations terrain

Après une première partie consacrée au contexte de cette transformation engagée au sein de la DGA, nous traiterons de l'illustration effective de la mise en place de ces nouvelles approches dans le cadre du Programme SCORPION.

2 CONTEXTE : L'INSTRUCTION MINISTERIELLE 1618

Le constat de départ, repris mi 2018 par la Ministre des Armées, fait état de 2 grands enjeux, qui doivent guider les travaux de conduite des Opérations d'Armement et des Opérations d'Investissement :

1. Travailler en équipe
2. Travailler plus vite

2.1 Travailler en équipe

« Les programmes pourraient se parler davantage, et les dialogues entre l'état-major et la DGA devraient être plus directs et pragmatiques »

Il est nécessaire de regrouper les équipes *« pour développer au quotidien l'axe capacitaire, de manière progressive sur l'ensemble du spectre d'activités »*, d'organiser les travaux *« à l'échelle d'une capacité et non plus isolément programme par programme »*

2.2 Travailler plus vite

« Nous ne voulons plus de sur-spécification, de délais non justifiés »

« La formalisation du besoin doit être élaborée conjointement avec la spécification technique du produit à développer, de manière à combiner les aspects techniques, industriels, financiers et calendaires »

« Réduire les délais et faciliter le processus de qualification et de mise en service des matériels sur la base de dispositions partagées et d'un programme commun d'essais et d'expérimentations »

2.3 Déclinaison dans les méthodes de conduite

L'instanciation des enjeux associés à ce constat se formalise concrètement par une nouvelle méthode de conduite des Opérations d'Armement et d'Investissement, synthétisée au travers de différents mots clés, leviers de cette démarche :

- Capacitaire
- Collaboratif
- Rapidité
- Agilité
- Incrémental
- Innovation
- Coopération
- Exportabilité

3 IMPACTS SUR LES TRAVAUX D'INGENIERIE

Au sein de la DGA, l'enjeu principal des travaux d'Ingénierie menés dans le cadre des conduites d'opération, est et doit rester, la maîtrise technique de nos systèmes.

Cette maîtrise technique s'articule autour de 2 grandes priorités :

1. Mieux définir et mieux spécifier nos systèmes
2. Renforcer la robustesse de nos qualifications

3.1 Mieux définir et mieux spécifier nos systèmes

Cette priorité se décline de la manière suivante :

- Améliorer la prise en compte des besoins multi-systèmes et la performance globale de nos moyens de défense
 - o Consolider l'approche globale par une utilisation entre programmes des mêmes outils et des mêmes pratiques
 - o Systématiser les tâches d'analyse « transverses »
- Optimiser collectivement, notamment avec l'Industriel, les dimensions coût / délai / performance

3.2 Renforcer la robustesse de nos qualifications, notamment du point de vue multi-systèmes

Sont entendues pour parvenir à cet enjeu, les activités suivantes :

- Planifier et organiser les campagnes d'essais de façon concertée
- Interconnecter les moyens, pour des essais bout en bout

3.3 Vision « Système de Systèmes »

Un des objectifs majeurs de cette nouvelle Instruction Ministérielle est de traiter de la cohérence multi programmes.

La dimension « outillage » pour atteindre cet objectif s'articule autour de 3 axes :

1. Modélisation
2. Simulation
3. Expérimentation et qualification

3.3.1 Outils de modélisation

Les outils de modélisation doivent permettre d'agréger des représentations graphiques de nos systèmes, que ce soit au niveaux des vues fonctionnelles comme des architectures techniques.

Cette attente nécessite de définir et partager des règles communes. La démarche centrée modèle (MBSE) peut être un élément de réponse.

3.3.2 Outils de simulation

Les outils de simulation doivent permettre de :

- Mieux évaluer des solutions et architectures avant la phase de réalisation
- Anticiper au plus tôt les phases de validation de nos systèmes.

Outre l'apport en phase de conception du système d'intérêt, les outils de simulation doivent désormais être pensés pour constituer un environnement simulé pour les systèmes connexes.

De plus, l'inter fonctionnement des outils de simulation doit permettre d'enrichir globalement le réalisme de nos campagnes de simulation.

3.3.3 Outils / moyens d'expérimentation / qualification

Les outils de simulation doivent permettre de construire et mettre à disposition des plateformes d'interopérabilité,

indispensables pour accompagner les nouveaux concepts de combat collaboratif. Cette dimension nécessite de mettre en place une organisation spécifique au sein des équipes, afin d'être en mesure d'adresser ces essais multi-systèmes.

L'illustration présentée dans la suite de cet article instancie concrètement ces dimensions multi systèmes / collaboratives / apport de la simulation et virtualisation dans les phases de Vérification et Validation d'une partie du programme de transformation de l'Armée de Terre.

4 ILLUSTRATION : V&V SDS SCORPION

SCORPION, opération d'armement majeure de l'armée de terre, se doit d'être proactif dans la mise en place de méthodes et outils, en particulier dans la phase de V&V. L'enjeu principal est de permettre à la DGA de fournir des éléments tangibles aux forces, leur permettant de décider d'un déploiement à venir.

Les paragraphes suivants, après présentation du contexte programmatique SCORPION, décrivent ce qui a été mis en place concrètement lors de la préparation et lors de la réalisation de cette phase de V&V.

4.1 Contexte programmatique SCORPION

SCORPION est un programme de transformation d'ampleur de l'armée française, avec pour objectif :

- De moderniser le parc des moyens de l'armée de terre avec en particulier l'arrivée de nouveaux véhicules,
- D'harmoniser les systèmes d'information et les systèmes de communication,
- De s'appuyer sur des socles communs d'échanges de données,
- De déployer des nouvelles capacités collaboratives telles que la tenue de situation amie (le BFT : *Blue Force Tracking*) ou la protection collaborative.

La réalisation cadencée de ces différents enjeux, permettra alors de répondre aux objectifs opérationnels suivants :

- Produire les justes effets,
- Limiter :
 - o les phases de contact violent dans le temps,
 - o l'espace et les volumes de forces engagées,
- Accroître la survivabilité des forces,
- Maîtriser l'empreinte logistique.

Pour mesurer l'ampleur des travaux à mener en V&V d'un tel programme, il est à noter que l'ensemble étudié correspond en langage militaire au déploiement de 400 véhicules et de 1200 hommes : Groupement Tactique Inter-Armes (GTIA). C'est en intégrant cette dimension opérationnelle et les capacités souhaitées, que la démarche de V&V a été pensée.

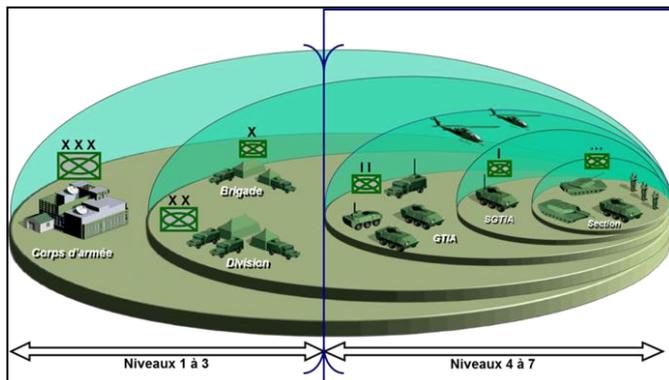


Figure 1. GTIA – Niveaux hiérarchiques 4 à 7

4.2 La spécificité du SDS & introduction des travaux V&V

Le GTIA SCORPION est un Système de Systèmes (SDS), dont le cycle de vie substitue un cycle de travail en « V imbriqués » au traditionnel V du développement des systèmes simples.

Les travaux de spécifications de niveau SDS puis de niveau système ont été réalisés via une démarche d'ingénierie système, incluant l'ingénierie des exigences et la modélisation d'architecture. Cela a permis de définir les niveaux capacitaires souhaités, les performances attendues, les systèmes concernés par telle ou telle atteinte de performance.

En préparant les travaux de V&V du SDS SCORPION, l'idée de décliner la démarche d'ingénierie système pour la mener de bout en bout s'est de fait imposée. La première étape a permis de lier les exigences définies en amont avec les scénarios d'essais, ce afin de valider ces exigences de la façon la plus automatisée possible.

La seconde étape a consisté à définir les moyens nécessaires à l'exécution des scénarios d'essais. Or, avec le niveau SDS attendu, ces moyens doivent être dimensionnés à la hauteur de l'enjeu du GTIA à déployer. C'est là que les notions de stimulation, simulation et virtualisation des systèmes à évaluer apparaissent comme des solutions potentielles.

4.3 Modélisation au profit des essais de V&V

La démarche IS de bout en bout a été retenue de par la complexité des essais de niveau SDS à mettre en place durant les phases de V&V. D'autant plus que cette démarche avait été initiée dans la phase amont du programme SCORPION, en particulier au travers de l'utilisation de modèles découpant le besoin en vues fonctionnelles et organiques.

Cette démarche de modélisation initiée dans la descente du V a été reprise pour plusieurs raisons :

- la modélisation s'avère pertinente pour représenter des scénarios de bout en bout faisant participer un nombre conséquent de systèmes de différents niveaux hiérarchiques au sein du GTIA. Les fiches d'essai (Word ou autre) ne permettent pas facilement de suivre l'histoire du bout en bout. Ce manque est alors comblé par cette représentation modèle sous forme de diagramme de séquences qui devient alors la base du déroulé du scénario.
- La modélisation permet également, via des fonctionnalités spécifiques proposées par les outils utilisés, de lier des exigences à telle ou telle action prévue dans les scénarios de bout en bout.

Il faut ensuite prévoir que ces données soient récupérées lors de l'exécution des dites actions. Pour cela, l'exploitation d'un export global des scénarios, pour en générer automatiquement des fiches d'essai, a été mis en place.

Afin de parvenir à cette ultime étape de génération de fiches d'essais, l'étape précédente de représentation de diagrammes de séquence doit suivre avec rigueur des règles de modélisation définies.

Le schéma ci-dessous représente les éléments principaux de cette démarche centrée modèle.

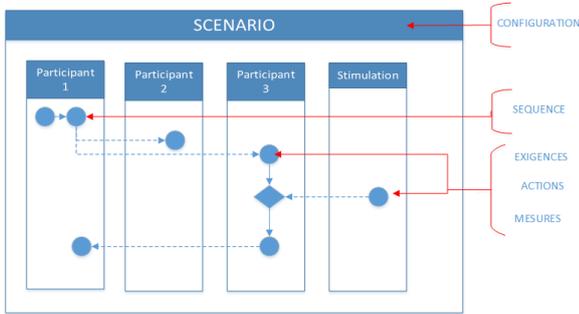


Figure 2. Terminologie modèle orienté test

Via une terminologie orientée test, les scénarios modélisés permettent de décliner les fiches d'essais, par participant. Chaque fiche d'essai est composée :

- De différentes **séquences** de test,
- **D'actions** à réaliser dans chacune des séquences en interne ou bien en interface avec un autre participant (réception/envoi de messages par exemple),
- De **mesures** à vérifier à chaque action,
- **D'exigences** à qualifier sur certaines actions

La **configuration** du scénario de bout en bout décrit les configuration scénaristiques, portées par la simulation pour préciser les différentes contraintes potentielles liées au scénario, pouvant avoir un impact sur par exemple les mesures observées. Cette configuration permettra alors de définir les moyens plates-formes à mettre en place.

Ci-dessous un résumé de la démarche :

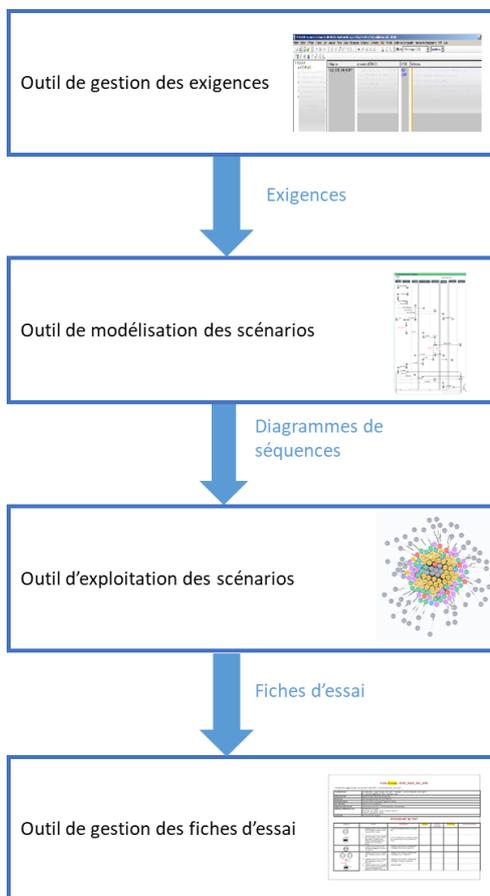


Figure 3. Démarche IS de bout en bout

Cette approche globale a permis notamment :

- De lier les données des différents outils d'IS (des exigences aux fiches d'essais),
- De disposer dans un outil de modélisation d'une base d'essais de bout en bout.

4.4 Démarche de conception de la plate-forme

Les essais, abordés dans cet article, sont réalisés après différentes phases de V&V systèmes, ayant permis de qualifier les systèmes unitairement, puis interface par interface. Ces essais sont exécutés par des équipes de la DGA assistés des opérationnels.

Des équipes SDS interviennent en complément des équipes système pour procéder aux essais de bout en bout.

L'objectif pour les plates-formes d'essais SDS est de permettre l'évaluation des performances d'un GTIA SCORPION avant son déploiement opérationnel.

Pour répondre à cet objectif, trois plates-formes de niveau SDS ont alors été définies, l'une relative au test d'un ou quelques véhicules SCORPION équipés (jusqu'à 5), l'autre mettant en œuvre un nombre important de véhicules sur le terrain (de 30 à 80).

Manque la représentativité du GTIA complet, en particulier en terme de charge réseau, point de départ de la réflexion menée pour la 3^{ème} plate-forme. Celle-ci a été pensée avec l'ambition de représenter en terme de systèmes d'information et de communication, le GTIA dans sa globalité : soit 400 éléments, interconnectés.

Il est évident que sur une plate-forme d'essais, ce volume important d'éléments ne peut être mis en œuvre. La démarche a alors consisté à déterminer le nombre minimum d'éléments réels (en systèmes d'information et radio) complété par des éléments non réels, permettant une évaluation suffisamment représentative d'un déploiement opérationnel.

Dans la partie suivante, cette notion d'éléments non réels (simulés et virtualisés) est détaillée.

4.5 Stimulation, Simulation & Virtualisation

Pour pouvoir évaluer les éléments réels dans un contexte proche de l'opérationnel, il est nécessaire de les alimenter avec des données correspondantes à un déploiement complet (qualitativement et quantitativement): ceci est possible grâce à la **stimulation**, à la **simulation** et à la **virtualisation**.

La **stimulation** tout d'abord, sur une plate-forme initialement statique, permet de générer des trames de position, envoyées à chacun des systèmes d'information, tels que le ferait un système de navigation.

Cela « fait croire » au système réel qu'il suit une trajectoire particulière comme s'il était dans un véhicule en mouvement. Cette stimulation est mise en place au travers de scénarios techniques et/ou opérationnels, définis dans des outils de simulation (par exemple : VBS (Virtual Battle Space) de Bohemia ou VRForces de MAK technologies). Chacun des véhicules est représenté dans ces outils, puis relié au système d'information. Ce système intègre les informations envoyées par la stimulation, met à jour sa position, et transmet celle-ci aux autres éléments du déploiement.

Vient ensuite la **simulation** :

- des systèmes d'informations,
- des systèmes de communication.

Les systèmes d'information simulés correspondent uniquement à des émetteurs de messages selon les normes appliquées par les systèmes d'informations réels.

Cela permet de générer autant de messages de position que le ferait un système d'information réel. Ces éléments simulés sont également connectés à la stimulation, et partagent des trames de position avec les autres éléments du déploiement.

Les systèmes d'information réels en liaison avec ces éléments simulés (via des radios simulés), affichent alors ces éléments sur leur IHM, rendant cette situation tactique perçue de plus en plus proche de ce que verrait un GTIA complet.

Pour aller au-delà de ce générateur de messages, très limité en représentativité, la **virtualisation** du système d'information a été étudiée. Après une comparaison entre différentes architectures possibles (*GPU, empilement de machines virtuelles sur un serveur, type de serveur*), une architecture permettant d'empiler jusqu'à 80 machines virtuelles sur un serveur a été retenue. Deux serveurs ont été déployés, faisant atteindre à la plate-forme PS³ l'équivalent de 160 systèmes virtualisés.

Ces machines virtuelles sont reliées à la stimulation ainsi qu'au reste du déploiement, via des radios simulés.

Cette solution permet finalement d'atteindre l'ambition de **représenter un GTIA sur une plate-forme de 50 m²** (à comparer avec les 600 km² nécessaires pour déployer effectivement un GTIA).

Ce concept de virtualisation permet :

- Une montée à l'échelle du nombre d'entités gérées par la plateforme,
- Une reconfiguration plus rapide de la plateforme avec le pool de VM reconfigurable en série,
- Un coût réduit par rapport à l'équivalent réel (pour les plates-forme de la DGA, et la solution est applicable également pour les centres formation).

Maintenant que le GTIA est représenté, il faut être capable d'évaluer le bout en bout d'un tel déploiement. Pour ce faire, un système de collecte et d'analyse est mis en œuvre.

Avec ce système, il est possible de collecter les données de toutes les composantes de la plate-forme, de la stimulation, jusqu'aux éléments réels. Après la collecte d'un volume de données conséquent, il faut être en mesure d'exploiter ces données.

Pour la partie analyse des données, une architecture de type *big data* est donc implémentée, permettant de procéder via des scripts prédéfinis à des analyses précises. Celles-ci peuvent répondre aussi bien à des objectifs techniques (avec des observations de phénomènes particuliers) qu'à des objectifs d'évaluation de performances de bout en bout.

Actuellement, la première fonction du combat collaboratif qu'est le BFT est en cours d'évaluation. L'outil d'analyse permet de comparer la « vérité terrain » émise par la stimulation, à ce qu'en retranscrit le système d'information réel : la **perception de soi**. Cette perception est partagée via les systèmes de communication aux autres éléments de la plate-forme. Ensuite, de nouvelles analyses sont faites pour comparer la vérité terrain d'un élément avec la réception du message de position par ses « voisins » : la **perception des autres**. Evidemment, ces analyses sont menées en prenant en compte un certain nombre de critères, telles que la vitesse de

déplacement, la présence ou non de la phonie, les échanges autres que le partage de position.

Ces analyses permettent avant déploiement terrain de fournir des éléments chiffrés aux forces évaluant telle ou telle fonction.

Il est à noter que ce système de collecte et d'analyse n'est pas uniquement déployé lors des essais plates-formes. Il est en effet possible de le mettre en œuvre dans un certain nombre de véhicules réels lors des essais terrain pré-déploiement. Après analyse de la DGA, des éléments chiffrés sont fournis aux opérationnels, prenant en compte les contraintes et aléas liés au terrain, se rapprochant alors des performances d'un déploiement réel.

Organisationnellement, la mise en place de ces moyens a rapproché les opérationnels des équipes de la DGA, à tel point que le rôle de la DGA dans la phase de V&V a été revu. Les opérationnels, conscients du potentiel d'un tel type de plate-forme ont émis un nouveau besoin : le « dérisquage » des essais terrain.

Les essais terrain mettent en œuvre un nombre important de véhicules et donc d'équipages associés, ce qui représente un coût conséquent. Il est important de définir en amont le périmètre fonctionnel précis de ces essais, d'être au fait des limitations éventuelles sur les scénarios opérationnels, afin d'optimiser ces phases.

Avant de partir en essais terrain, les équipes de la DGA préparent en amont une plate-forme permettant de jouer en avance de phase les scénarios, avec le concours des forces. Cette étape de dé-risquage valide ce que percevront les opérateurs des systèmes et remonte le cas échéant les problèmes susceptibles de se reproduire sur le terrain.

Après deux années de mise en place, ces opérations de dé-risquage amont se sont avérées essentielles, et sont devenues au sein du programme SCORPION une démarche quasi systématique. L'apport a été prouvé, les essais plate-forme et terrain permettant de corroborer les observations, et de statuer plus aisément sur l'état réel des systèmes.

5 PERSPECTIVES

Les travaux réalisés pour SCORPION dans le cadre de ces activités ouvrent de réelles perspectives pour la conduite des Opérations d'Armement au sein du Ministère des Armées :

1. Au niveau de la démarche, un travail de capitalisation est en cours, afin de pouvoir proposer une méthodologie de réflexion applicable à d'autres programmes d'armement. Certains systèmes de systèmes d'envergure ont d'ores et déjà fait la démarche de s'inscrire dans cette réflexion, et d'utiliser les retours d'expérience proposés
2. La plateforme mise en œuvre est centrée autour de l'Armée de Terre, et des systèmes de communication. Des études connexes sont actuellement en cours pour imaginer une interconnexion de cette plateforme avec d'autres plateformes d'essais (disponibles ou en cours de conception), afin de disposer d'environnement potentiellement inter armées, permettant de remonter encore dans l'approche capacitaire.

6 CONCLUSION

L'apport d'une telle démarche peut se mesurer sur différents axes :

1. En termes de temps, anticiper et dérisquer les évaluations opérationnelles permet d'être plus efficace sur le terrain, et de mobiliser les Forces sur un temps utile optimisé.
2. En termes de coût, plusieurs facteurs de gain sont avérés :
 - a. Coût en Ressources Humaines opérationnelles, en ciblant beaucoup plus précisément les attendus et sollicitations pendant les phases d'évaluation
 - b. Coût en Ressources Matérielles, en maximisant l'utilisation en amont des éléments virtuels / simulés
3. En termes de Qualité des Systèmes :
 - a. en permettant l'identification des anomalies au plus tôt,
 - b. en associant les opérationnels suffisamment en amont pour dérisquer l'adoption des systèmes livrés,
 - c. en évaluant les niveaux de performance des systèmes au plus tôt dans la phase de validation.

Au-delà de ces gains effectifs, la démarche d'anticipation des phases de V&V pour SCORPION, a permis d'initier des réflexions sur les processus de gestion des opérations d'Armement de niveau Système de Systèmes, autour d'enjeux majeurs tels que :

- l'exécution automatisée de scénarios d'essais opérationnels, sur des environnements représentatifs et maîtrisés,
- les moyens à mettre en place, et la réflexion autour de la valeur ajoutée d'environnements virtualisés / simulés.

Enfin, retenons au niveau organisationnel, la synergie obtenue entre les Forces et les équipes de la DGA, permettant une réelle approche collaborative, et une optimisation des sollicitations pendant les différentes phases d'essai,

Session 4.2 - Session spéciale 8 :
Transformation du secteur
agroalimentaire à l'ère de l'industrie 4.0
Chairman : Robert Pellerin - Maurice Pillet

Prédiction de la qualité de la farine après mouture

LOIC PARRENIN^{1,2}, CHRISTOPHE DANJOU^{1,2}, ROBERT BEAUCHEMIN³, BRUNO AGARD^{1,2}

¹ Laboratoire en Intelligence des Données
Département de mathématiques et génie industriel,
École Polytechnique de Montréal, CP 6079, succursale Centre-Ville, Montréal, Québec, Canada
loic.parrenin@polymtl.ca, christophe.danjou@polymtl.ca, bruno.agard@polymtl.ca

² Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT)

³ La Meunerie Milanaise, Saint-Jean-sur-Richelieu, Québec, Canada

Résumé – Dans la production de farine biologique, différentes qualités de farine sont exigées par les industriels ou artisans en fonction du type de produit à fabriquer. La qualité de ces grains, matière première vivante, varie au cours du temps en fonction de nombreux facteurs, entraînant par conséquent des variations sur la qualité de la farine biologique. Pour obtenir la qualité de farine souhaitée, il est possible de réaliser sur la chaîne de transformation des mélanges de blé ainsi que des ajustements opérés sur les machines de productions. L'article a pour objectif de proposer un outil de prédiction de la qualité de la farine produite, à travers un modèle de classification. Le modèle utilise les données collectées sur les mélanges ainsi que la qualité des grains de blé disponibles en amont de la chaîne de production. L'outil permet d'identifier les caractéristiques influentes sur la qualité de la farine et d'anticiper certains choix dans la transformation des grains de blé.

Abstract - In the production of organic flour, different qualities of flour are required by industrialists or craftsmen depending on the type of product to be produced. The quality of these grains, a living raw material, varies over time depending on many factors, thus leading to variations in the quality of organic flour. In order to obtain the desired flour quality, it is possible to make wheat mixes on the processing line as well as adjustments on the production machines. The article aims to propose a tool for predicting the quality of the flour produced, through a classification model. The model uses the data collected on the mixes as well as the quality of the wheat grains available upstream of the production chain. The tool allows to identify the characteristics that influence flour quality and to anticipate certain choices in wheat grains processing.

Mots clés – Industrie 4.0, Grains biologiques, Prédiction, Qualité, Classification.

Keywords – Industry 4.0, Organic grains, Prediction, Quality, Classification

1 INTRODUCTION

L'accroissement de la population mondiale et l'augmentation des niveaux de vie entraînent des besoins alimentaires plus importants. L'agriculture biologique est une méthode spécifique de production alimentaire utilisant des substances et des procédés naturels. Cette méthode encourage une utilisation responsable des ressources naturelles, le maintien et l'amélioration de la biodiversité, la qualité de l'eau et la fertilité des sols (MAPAQ, 2019). Le secteur céréalier biologique affiche en 2017 une croissance significative de la quantité de surface dédiée à la culture de céréales biologiques dans le monde avec un gain de 6% (représentant 280 000 hectares) par rapport à 2016 (Willer & Lernoud, 2019). D'après (Future Market Insights, 2020), on note une augmentation de la demande de farine de blé biologique, en Amérique du Nord et dans les pays d'Europe occidentale, notamment chez les fabricants de produits alimentaires, en raison de consommateurs plus soucieux de leur santé.

La farine est produite à partir du broyage de grains de blé. Le broyage (ou mouture) se déroule typiquement dans un moulin. L'objectif est de séparer les différentes parties du grain de blé et de les réduire en plus petits morceaux. Deux principales techniques existent pour la mouture du blé (Cappelli, Oliva & Cini, 2020) : la mouture sur meule et la mouture sur cylindre. Chaque technique présente ses avantages et inconvénients.

Toutefois, la mouture sur cylindre est une technique plus récente et la plus répandue actuellement dans les moulins pour des raisons de rendements élevés. Elle permet d'obtenir des farines très blanches.

Il existe différentes catégories de blé :

- Le blé dur et le blé tendre sont des blés de dureté différente. Le blé dur possède une plus grande concentration en gluten, ce qui le rend très dur. Sa flexibilité et sa force en font un blé idéal pour les pâtes alimentaires. Le blé tendre est le blé le plus semé et le plus utilisé dans l'alimentation. Il est cultivé pour faire de la farine panifiable.
- Le blé de printemps et le blé d'hiver sont des blés semés à différentes périodes de l'année en fonction des zones climatiques. Un blé semé au printemps sera dit « blé de printemps » tandis qu'un blé semé à l'automne sera dit « blé d'hiver ».

Chaque type et variété de blé présente des caractéristiques différentes. De plus, les propriétés intrinsèques d'un type de grains de blé varient énormément entre les régions, les conditions météorologiques et climatiques, les périodes de récoltes et la rotation des cultures (Borghi et al., 1995; Johansson & Svensson, 1998; Triboi et al., 2000).

La production industrielle de farine biologique est quelque peu différente de la farine conventionnelle. Pour la production de farine conventionnelle, les volumes de production sont très importants, et très peu de mélanges sont réalisés. De plus, pour maintenir une qualité de production stable tout au long de l'année, il est possible d'avoir recours à des agents chimiques tels que le chlore ou le peroxyde de benzoyle pour le blanchissement de la farine, l'acide ascorbique pour améliorer la qualité boulangère et la couleur de la farine, ainsi que des enzymes pour uniformiser l'indice de chute (Gouvernement du Canada, 2020). Alors que pour la production de farine de blé biologique, les grains de blé utilisés sont cultivés sans aucune utilisation de produit chimique, de pesticides et d'engrais synthétiques (MAPAQ, 2019). Lors de la transformation, aucun ajout d'agent chimique n'est autorisé au regard de la production alimentaire biologique. Il faut alors utiliser la bonne proportion de différentes sortes de grains de blé, qui ensemble permettront d'obtenir le produit final voulu.

Dans ce contexte, **comment produire de la farine de blé biologique, de qualité constante tout au long de l'année, lorsque la qualité des matières premières fluctue selon la période de l'année, en fonction des conditions climatiques et météorologiques ?**

Dans le présent article, nous proposons une méthodologie qui vise à prédire différentes caractéristiques de la farine après mouture, en fonction de la qualité des grains en entrée et des proportions des grains de blé pour chaque farine produite. La méthodologie proposée s'appuie sur un arbre de décision.

La structure de l'article se décompose de la manière suivante. Dans l'état de l'art (section 2), nous présentons les éléments essentiels de la production de la farine (2.1), notamment les caractéristiques intrinsèques des grains de blé (2.1.1), les principales caractéristiques recherchées de la farine (2.1.2) et les principales étapes de la production de farine (2.1.3). Puis, l'état de l'art présente les outils existants dans le domaine de la meunerie pour prédire la qualité de la farine (2.2). Finalement, nous présentons ensuite les éléments essentiels de l'outil de prédiction utilisés (2.3). La section 3 présentera l'outil proposé. La section 4 montrera une validation de l'outil dans un contexte réel. Finalement, la section 5 rappelle les résultats et limites de la méthode proposée et propose quelques perspectives.

2 ÉTAT DE L'ART

Bien que le domaine de la transformation des céréales en général soit très bien documenté, il n'existe que très peu de références scientifiques qui décrivent la production de la farine biologique.

2.1 Production de la farine biologique

D'après une étude sur la production alimentaire biologique, la plus grande superficie consacrée aux produits céréaliers dans le monde était en 2017 pour le grain de blé (32%) suivi des grains de maïs (14%) (Willer & Lernoud, 2019). La demande de farine biologique augmente, ceci est d'autant plus vrai en période de crise liée au covid-19 (Fournier, 2020).

2.1.1 Caractéristiques intrinsèques des grains de blé

Différentes variétés de grain de blé existent et sont définies par des caractéristiques physiques spécifiques. Tel que schématisé dans la figure 1, le grain de blé est situé dans un épi de blé et se compose de 4 parties distinctes : le **germe** (≈ 3 % de la masse

totale du grain); l'**albumen** ou **amande farineuse** ou endosperme (84%) ; les **enveloppes** (13%) et la brosse (masse négligeable).

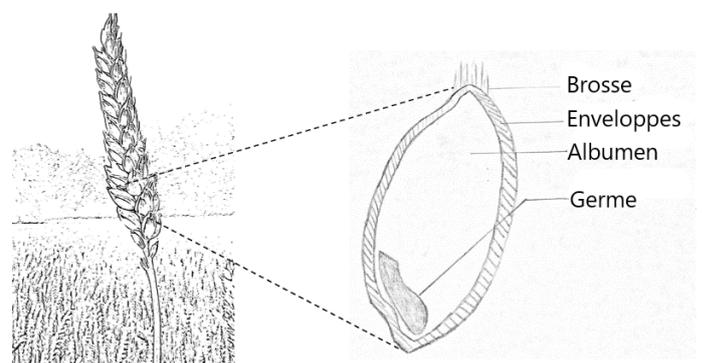


Figure 1. Schéma de l'anatomie d'un grain de blé (coupe longitudinale d'un grain de blé)

Le germe représente la future plante, elle est riche en matière grasse et en vitamines B et E. L'amande farineuse est principalement constituée de glucides (amidon) et de protéines (dont le gluten). L'amidon est un glucide complexe (sucre lent) utile à la croissance de la nouvelle plante. Le **gluten est une protéine qui donne aux pains l'élasticité** suffisante pour assurer un gonflement du pain. Les caractéristiques de chaque variété de blé varient en fonction du lieu de récolte et de la saison.

Pour évaluer la qualité d'un grain de blé, des analyses sont réalisées sur des échantillons. Les analyses mesurent notamment le **taux de protéine**, le **taux d'humidité**, l'**indice de chute**, le **taux de cendres**, le **BEM** (Brabender Energy Max) et le **PMT** (Peek Max Time). Le taux de protéine indique le pourcentage de protéine dans le grain. Le taux d'humidité précise la quantité d'eau présente dans le grain. L'indice de chute mesure l'activité enzymatique du grain qui se développe dans le grain dès le début de la germination (Côté, 2018), cette mesure informe sur le stade de germination du grain. Le taux de cendre mesure la quantité de minéraux présents dans un grain ou dans la farine. Le BEM et PMT sont des mesures obtenues par le test Glutopeak. L'appareil analyse les capacités d'agrégation du gluten en phase liquide (Gresle, 2013). Le BEM définit le couple maximum exprimé en unité brabender (UB) (Gresle, 2013) et précise la force boulangère de la pâte. Le pic de la courbe obtenue par le Glutopeak correspond au BEM (Gresle, 2013). Le PMT indique la durée nécessaire pour atteindre le couple maximum (Gresle, 2013).

2.1.2 Caractéristiques recherchées de la farine

Plusieurs types de farine existent en fonction du produit alimentaire final recherché : pâtisserie, biscuit, pain, pâte à pizza, pâte alimentaire, semoule...

Un premier choix est réalisé au niveau de la **dureté du blé**. Un blé dur est favorable pour produire une farine dédiée à la production de semoule ou de pâte alimentaire. Tandis qu'un blé tendre est utilisé pour produire de la farine panifiable (farine dédiée à la production de pain, pâtisserie, biscuit).

Au niveau du blé tendre, un premier paramètre de qualité s'oriente autour de la **proportion en gluten** (une protéine) présent dans le grain de blé. Un blé tendre plus riche en protéine sera adapté pour la fabrication de pain. Une farine plus riche en protéine, et donc en gluten, permettra d'obtenir les qualités d'élasticité et

d'extensibilité de la pâte lors de la confection du pain. Tandis qu'un blé tendre plus faible en protéine sera plus approprié pour la pâtisserie et les biscuits, où les propriétés d'une pâte cassante et d'absorption d'eau plus faible sont recherchées.

Un deuxième paramètre concerne la **proportion de « son »** contenu dans la farine. Cette proportion est mesurée à l'aide de l'indicateur du taux de cendre et constitue le principal critère de classement des farines de blé pour le grand public. Plus le taux de cendre est élevé, plus la farine contient du son. Le son regroupe l'enveloppe et le germe du blé. Une farine dite intégrale contient tous les éléments nutritifs présents dans un grain de blé (enveloppe, albumen, germe). Une farine complète contient seulement les parties de l'enveloppe et de l'albumen du grain de blé. Tandis qu'une farine blanche sera composée seulement de la partie de l'albumen.

Comme on a pu le mentionner, plus le taux de protéine est élevé, plus la proportion en gluten dans la farine est importante. Toutefois, le **taux de protéine ne reflète pas la qualité du gluten**. C'est pour cela que le BEM représente une mesure de la qualité du blé plus fiable que le taux de protéine. Le BEM semble être corrélé au **taux d'absorption** en eau qui est un paramètre important pour déterminer les propriétés fonctionnelles de la farine (Fu, Wang & Dupuis, 2017). Une farine avec une plus grande absorption en eau, est préférée pour la production du pain (propriété de manipulation de la pâte améliorée, pain de meilleure qualité au niveau du goût et de la structure) - (Fu et al., 2017). Tandis qu'avec un taux d'absorption plus faible, il est préférable de produire de la pâtisserie et des biscuits.

2.1.3 Les étapes de production de la farine de blé

La transformation des grains de blé en farine se déroule au moulin en plusieurs étapes. Les principales étapes de la production de farine de blé sont représentées sur la figure 2.

(1) **Les blés** achetés sont **livrés** par lot, par camions, au cours de l'année. Les blés développent des caractéristiques qui leur sont propres en fonction de leur environnement et de leur type. La grandeur des champs peut être également une source de variabilité de la qualité du blé avec des zones qui ont possiblement pu être moins irriguées que d'autres. Pour s'assurer de la qualité des blés réceptionnés, une prise d'échantillon des grains de blé est réalisée pour être ensuite analysée au laboratoire. Les résultats d'analyse

permettent d'identifier le blé et de sélectionner le silo dans lequel les grains de blé seront entreposés. Afin de préserver la qualité de chaque type de blé et contrôler à un certain degré les mélanges, les blés de qualité similaire sont entreposés dans un même silo. La présence de multiples silos permet d'entreposer davantage de grain de blé, mais également de séparer les grains de blé de qualité différente. (2) Avant que les grains de blé soient entreposés, ils subissent une **étape de pré-nettoyage** (les criblures de blé désignent les impuretés) de manière à éliminer les gros déchets (paille, terre) ainsi que les objets métalliques pouvant endommager les machines de production.

Le talent du meunier consiste à bien **choisir les blés et les mélanger dans de bonnes proportions** afin de réaliser de bon mariage de blés et obtenir le type de farine souhaité, **ces mélanges correspondent à des recettes**. Les recettes sont réalisées en fonction de la qualité et de la disponibilité des matières premières. La recette définie, (3) les grains de blé sont **nettoyés**, l'objectif est de retirer toutes les impuretés (grains cassés, pierres, pailles...) pour ne garder que les grains de blé. (4) Les grains nettoyés passent au processus de mouillage dans le but d'**humidifier** les grains de blé et faciliter la séparation de l'albumen du reste des composants lors du processus de mouture. Les grains mouillés sont finalement envoyés au processus de mouture. L'opération de mouture **permet de séparer les différentes parties du grain de blé et de réduire en farine l'amande farineuse** du grain pour obtenir de la farine. La mouture se décompose en 4 étapes : le broyage ; le claquage ; le convertissage et le blutage. (5) **Le broyage** permet d'ouvrir le grain, à l'aide de cylindres cannelés, et d'obtenir de grosses particules des différentes parties du grain (semoules). (7) **Le claquage** réduit les semoules blanches (grosses particules d'amande farineuse). (8) **Le convertissage** réduit les semoules vêtues (particules dont une partie de l'amande farineuse est toujours collée à l'enveloppe). La réduction des particules se fait à l'aide de cylindres lisses. (6) À chaque étape, les particules sont triées en fonction de leur granulométrie, à partir de plusieurs tamis présents dans un *plansichter* pour être redirigées vers une nouvelle étape de broyage, de réduction ou dans un silo de produit transformé (farine ou son), c'est **l'étape de blutage**. Ces étapes sont réalisées de manières successives, d'où la présence dans le moulin de plusieurs machines de broyage, de claquage et de convertissage.

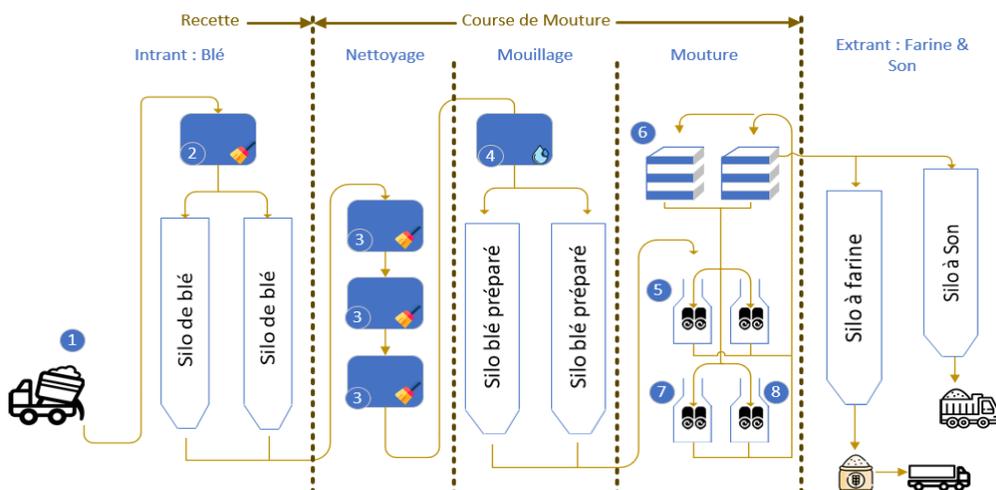


Figure 2. Processus de production de la farine de blé

Une multitude de paramètres doivent être contrôlés et certains ajustés en conséquence, tant au niveau des recettes, du mouillage que de la mouture afin de produire de la farine biologique. Cependant, l'influence et l'importance de chacun des paramètres sur l'ensemble du processus rendent leurs ajustements complexes.

2.2 Prédire la qualité de la farine

Actuellement dans le secteur de la farine, le meunier a pour rôle de sélectionner les mélanges de blés en entrée du processus de transformation et d'ajuster les paramètres de production pour obtenir une qualité de farine souhaitée. La qualité de la farine est mesurée généralement à l'aide d'un farinographe (temps de stabilité), alvéographe (index alvéographique (W, P/L)) et d'un proche infrarouge (protéine, humidité, taux de cendre) - (Cocchi et al., 2005). Des moyennes sur la qualité des grains de blé mélangé ou un échantillon représentatif d'une recette peuvent être analysés avant de débiter une course de mouture de manière à estimer la qualité de la farine produite.

(Fu et al., 2017) montre que le couple max (BEM) du « Glutopéak » et le taux d'absorption d'eau de la farine présentent une forte corrélation positive avec un coefficient de détermination (r^2) de 0.97 sur 83 échantillons de grains de blé différents. Le test Glutopéak est donc être un outil rapide et utile pour prédire le taux d'absorption d'eau de la farine et déterminer sa qualité.

Des méthodes d'apprentissage machine sont aujourd'hui utilisées pour prédire la qualité ou les paramètres de certaines transformations d'aliment (Adebayo, Hashim, Abdan, Hanafi & Mollazade, 2016; Bao et al., 2014; Barbon et al., 2016; El-Bendary, El Hariri, Hassanien & Badr, 2015; Liu, Yang & Deng, 2015).

Dans le milieu de la farine, de nombreuses méthodes de prédiction existent également pour prédire la qualité de la pâte (Torbica, Blazek, Belovic & Hajnal, 2019) ou l'usage final de la farine (Cocchi et al., 2005) en fonction de sa qualité. Des prédictions de la qualité de la farine ont été réalisées à l'aide d'un modèle de réseaux neurones en utilisant comme donnée d'entrée les résultats (ondes) de la spectroscopie du proche infrarouge de différentes farines (Mutlu et al., 2011). L'objectif est différent du travail présenté ici, car (Mutlu et al., 2011) traitent des données d'analyse de farine pour prédire les caractéristiques de la farine (taux de protéine, taux d'absorption, etc.) dans le but de simplifier les analyses et réduire les temps d'analyse, alors que la présente étude traite des données d'analyses de grains de blé pour prédire une classe de qualité de farine.

Parmi les nombreuses méthodes étudiées, aucune ne fait le lien entre la qualité des matières premières et les recettes de transformation réalisées. La qualité des grains de blé varie au cours du temps (transport, entreposage) comme tout produit alimentaire interagissant avec son environnement extérieur. Toutefois des grains de blé de qualité différente sont parfois mélangés pour obtenir une certaine qualité de farine désirée. De plus, ces grains de blé subissent des opérations de nettoyage, mouillage et de mouture impactant la qualité de la farine. La qualité de la farine dépend donc de nombreux paramètres dans l'ensemble du processus de production. La manipulation de matière première vivante le long du processus de production rend la prédiction de la qualité de la farine complexe.

2.3 Outil de prédiction

Il existe différents outils de prédiction, dont les réseaux neurones (RN). Les RN sont devenus une technique populaire dans les sciences biologiques en raison de leur qualité prédictive (Alvarez, 2009; Dubey, Bhagwat, Shouche & Sainis, 2006). Les réseaux neurones ont pu être utilisés avec succès dans l'industrie alimentaire tels que : la classification de grains de céréales à l'aide de caractéristiques morphologiques (Visen, Paliwal, Jayas & White, 2002), la prédiction de la conductivité thermique des produits boulanger (Sablani, Baik & Marcotte, 2002), la prédiction de paramètres de qualité de la farine à partir des spectres du proche infrarouge (Mutlu et al., 2011), la prédiction de la qualité des petits pois en fonction du temps de cuisson (Xie & Xiong, 1999), etc.

Bien que les réseaux neurones offrent de bons résultats de prédiction, ce type de modèle a besoin d'avoir de nombreuses données d'entraînement se définissant par des données traitées et utilisées pour l'apprentissage du modèle, de manière à atteindre un niveau de prédiction précis (Wuest, Weimer, Irgens & Thoben, 2016). De plus, le modèle des réseaux neurones est difficilement interprétable, souvent caractérisé comme une « boîte noire » qui n'explique pas les relations et l'influence des paramètres sur les résultats de prédiction (donnée de sortie). Un modèle de régression non linéaire ou d'**arbre de décision** semblent de bonnes alternatives au modèle des réseaux neurones afin de pouvoir **interpréter et comprendre ces modèles de prédiction et de classification**.

Un arbre de décision est un graphique orienté acyclique utilisé pour prendre des décisions (Burkov, 2019a). Il est composé de nœuds, de branches et de feuilles. Un nœud teste un attribut. Une branche correspond à la valeur d'un attribut suite à un test réalisé au niveau d'un nœud. Une feuille indique la classe à laquelle l'exemple appartient. L'arbre de décision permet à partir des données d'entraînement de tester plusieurs attributs et d'établir certaines règles, précisées à chaque nœud, pour classifier chaque exemple. Les données tests qui sont de nouvelles données indépendantes des données d'entraînement valideront le modèle de classification suivant les règles établies.

Dans une approche similaire, (Ronowicz, Thommes, Kleinebudde & Krysinski, 2015) ont étudié la qualité des pastilles pharmaceutique. Pour cela, ils ont exploré à l'aide d'un modèle d'arbre de décision, les relations de cause à effet entre les caractéristiques de la formulation (composition des granulés et des paramètres du procédé) et la sphéricité des granulés (attribut de qualité sélectionné du produit final). La découverte de ces relations de cause à effet entre les différentes variables offre la possibilité d'optimiser le processus de fabrication des granulés. Cette étude montre la pertinence d'une approche par arbre de décision dans le milieu pharmaceutique.

3 OUTIL DE PREDICTION DE LA FARINE A PARTIR DE LA QUALITE DES GRAINS DE BLE

L'État de l'art montre que les grains de blé sont une matière « vivante » définie par des propriétés intrinsèques évoluant avec le temps. La connaissance des caractéristiques de la farine produite est importante pour déterminer l'usage final du produit et reprendre précisément au besoin du client.

Dans le milieu biologique, l'ajout d'agents chimiques pour atteindre les caractéristiques recherchées ne peut être envisageable au regard des normes (MAPAQ, 2019). Pour cela il est nécessaire de travailler avec la qualité des grains de blé récolté pour obtenir une qualité de farine donnée. Cependant, d'après l'état de l'art (section 2.2), il n'y a actuellement pas d'outils pour prédire la qualité de la farine à partir des matières premières. Le problème fait intervenir de nombreux facteurs impactant pour chacun d'entre eux la qualité de la farine à des degrés différents. La pratique se base actuellement sur l'expérience du meunier.

Nous proposons de **développer un outil de prédiction** afin de déterminer les caractéristiques de la farine qui sera produite après le processus de mouture. L'outil de prédiction utilise la moyenne de la qualité des grains de blé présent dans un silo (protéine, humidité, taux de cendre, BEM, PMT) ainsi que les pourcentages des lots de grains de blé de chaque silo. À partir de ces données d'entrées, l'outil indique la classe de la qualité de la farine associée au BEM de la farine.

Pour construire l'outil de prédiction, une méthode telle que représentée sur la figure 3 est suivie. Cette méthode se compose de 4 phases.

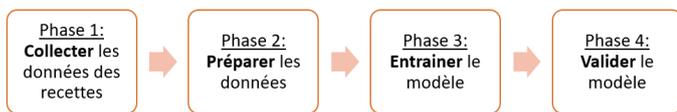


Figure 3. Méthodologie de prédiction

3.1 Phase 1 : collecter les données des recettes

Dans cette première phase, il est nécessaire de **définir les besoins**. Cette définition précise les données utiles pour le développement de l'outil de prédiction et délimite notre zone d'étude. Dans notre cas, nous souhaitons collecter les données précisant la qualité des grains présents dans chaque silo ainsi que leur quantité pour chaque recette réalisée.

La qualité des grains de blé est mesurée à l'aide d'un proche infrarouge (protéine, humidité, taux de cendre) et d'un test Glutopeak (BEM, PMT). Dans chaque silo, différents types de grain de qualité similaire peuvent être mélangés. Pour cela des moyennes sur le taux de protéine, le taux d'humidité, le taux de cendre ainsi que le BEM et le PMT sont calculés pour préparer une recette. Il est alors nécessaire de collecter les moyennes de la qualité des blés pour chaque silo ainsi que les pourcentages des quantités de blé de chaque silo utilisé pour la recette.

3.2 Phase 2 : préparer les données

Afin d'entraîner le modèle, il est nécessaire de préparer les données. Pour cela, nous rassemblons l'ensemble des données dans une **table de donnée** (« dataframe »). Un aperçu de l'entête de la table de donnée est montré sur le tableau 1.

RECETTE (R)	A1	A2	...	Y
R1	a _{1,1}	a _{1,2}	...	y ₁
R2	a _{2,1}	a _{2,2}	...	y ₂
R3	a _{3,1}	a _{3,2}	...	y ₃

Tableau 1. Table de donnée

Chaque ligne de cette table représente une **recette réalisée (R)**. Nous retrouvons ainsi en colonne, les informations collectées au niveau des recettes précisant la qualité des grains et les pourcentages des mélanges utilisés. Les attributs désignent les **caractéristiques mesurées (A)** pour identifier la qualité des grains de blé. À partir des proportions de quantité de blé utilisées pour chaque silo, une **moyenne sur chaque caractéristique (a_{i,j})** est calculée. La **donnée de sortie (Y)** indique la classe de la qualité de la farine. La classe est identifiée à partir du BEM provenant des résultats d'analyses du test du Glutopeak. L'analyse est réalisée sur un échantillon de farine après le processus de mouture. La donnée de sortie doit être associée à la recette réalisée pour être en mesure d'entraîner le modèle. Ce type d'apprentissage est un **modèle d'apprentissage supervisé**, c'est-à-dire que les données sont annotées par une étiquette (« label ») afin que l'algorithme puisse être capable, une fois entraîné, de prédire l'étiquette sur de nouvelles données non annotées.

Pour certains modèles, il est nécessaire de **normaliser** les valeurs de manière à éviter qu'une entité domine sur les autres en raison d'une plus grande distribution ou plage de valeur (Burkov, 2020). Les arbres de décision ne sont pas sensibles face à ces disparités entre chaque entité, pour cela les données n'ont pas été normalisées.

Dans certains cas, il est possible que les **données soient déséquilibrées** et que les valeurs de **certain attributs ne soient pas disponibles ou manquent dans l'ensemble de données** (Wuest et al., 2016). En fonction du modèle algorithmique choisi et du problème, différentes solutions existent et sont disponibles dans la littérature scientifique (Lakshminarayan, Harp, Goldman & Samad, 1996; Marlin, 2008; Zhang & Hu, 2014). Dans notre cas, en utilisant un modèle d'arbre de décision pour la prédiction et de la classification, le déséquilibre a un impact sur la structure de l'arbre décisionnel. Il est possible de rebalancer les données en infligeant une pénalité pour certaines erreurs de classification. La **classification pénalisée** impose un coût supplémentaire au modèle pour les erreurs de classification commises sur la classe minoritaire de l'ensemble de données (Datascientest, 2020). Ces pénalités biaisent le modèle afin qu'il accorde une plus grande importance à la classe minoritaire.

Parfois les ensembles de données présentent des données manquantes pour certains attributs, le plus souvent lorsque les données sont rentrées manuellement. Plusieurs approches sont possibles : supprimer les recettes dont les données sont manquantes ; utiliser un modèle d'apprentissage pour remplacer les valeurs manquantes ; utiliser une technique d'imputation des données. Avec un ensemble de données suffisamment grand, ne contenant peu de données manquantes, il a été choisi de **supprimer les recettes ayant des données manquantes** dans le but de garder des informations précises et fiables.

Pour être en mesure d'entraîner et valider notre modèle, nous séparons notre ensemble de données en **trois sous-ensembles**. On utilise **70%** des données pour les **données d'entraînement**, **15%** pour la **validation** et **15 %** pour le **test** (Burkov, 2019b). Ces proportions varient en fonction de la dimension des données (Burkov, 2019b). Afin d'entraîner correctement notre modèle et éviter de créer des biais sur notre ensemble de données tests, nous gardons une proportion représentative de chaque classe de sortie dans les 3 sous-ensembles. Pour cela, il est possible de réaliser un échantillonnage **stratifié**.

3.3 Phase 3 : entraîner le modèle

Plusieurs algorithmes d'arbres de décision existent : ID3, C4.5, C5.0, CART (Classification and Regression Trees) - (scikit-learn developers, 2020). L'algorithme d'arbre de décision le plus connu et utilisé est l'algorithme C4.5. Il accepte les attributs discrets et continus, gère les exemples incomplets et résout les problèmes de surapprentissage par « élagage ». **L'outil de prédiction utilise une version optimisée de l'algorithme CART à l'aide de la librairie scikit-learn** sous le langage Python (Pedregosa et al., 2011).

Lors de l'entraînement du modèle, il est important d'éviter d'avoir un surapprentissage ou sous-apprentissage du modèle. On parle de surapprentissage si la performance du modèle sur les données d'entraînement est considérablement meilleure que sur les données de test. Lorsque le modèle fait du **surapprentissage, la variance est élevée** car le modèle aura appris le bruit présent dans les données d'entraînement. Le bruit sont des données ayant des fluctuations aléatoires ou des décalages par rapport aux valeurs réelles au niveau des attributs et de la variable de réponse. Ce bruit peut masquer la relation réelle entre les attributs et la variable de réponse. Dans le cas où le modèle réalise de nombreuses mauvaises prédictions sur les données d'entraînement, le modèle fait du **sous-apprentissage et possède un biais élevé**. Les raisons peuvent être multiples : un modèle trop simple pour les données étudiées ; un nombre d'attributs qui ne fournit pas suffisamment de données assez informatives pour enrichir le modèle et représenter la réalité (Burkov, 2019b).

Pour optimiser la performance du modèle, il est nécessaire d'affiner les paramètres de l'algorithme. Ces paramètres appelés hyperparamètres sont ajustés de manière expérimentale. Avec un modèle d'arbre de décision, il est possible d'ajuster les hyperparamètres par des critères d'arrêt (pré-élagage) ou des méthodes post-élagage. Les critères d'arrêt regroupent des techniques telles que : limiter la profondeur maximale de l'arbre; préciser le nombre minimum d'exemples nécessaires pour réaliser une séparation; indiquant le nombre minimum d'objets nécessaires pour avoir une feuille. Les méthodes post-élagage sont par exemple : fixer une valeur minimale à atteindre pour la diminution de l'impureté lors d'une séparation à un nœud; ajouter une fraction de poids à respecter au minimum pour la création d'une feuille et de ses branches. **Le bon ajustement de ces paramètres doit permettre de trouver un juste équilibre entre le biais et la variance.**

De plus, face à des données déséquilibrées au niveau de la variable de sortie, une classification pénalisée est réalisée comme décrit dans la phase 2. Une analyse de sensibilité du coût en fonction de la précision (« accuracy ») est examinée afin de déterminer les pénalités optimales à attribuer pour chaque type de mauvaise classification.

3.4 Phase 4 : valider le modèle

À partir de l'ensemble de **données tests**, il sera possible de valider la performance du modèle créé et ajusté avec les données d'entraînement et de validation. L'évaluation du modèle de classification est réalisée à partir d'une **matrice de confusion** de manière à connaître les résultats de classification et la précision du modèle à prédire la bonne classe de sortie. La matrice de confusion renseigne sur les bonnes classifications, les faux positifs et les faux négatifs.

Le tableau ci-dessous montre une matrice de confusion 2x2, représenté par seulement 2 classes de sorties.

Classe réelle	Classe prédite		
		Classe 1	Classe 2
	Classe 1	Vrai Positif	Faux Négatif
Classe 2	Faux Positif	Vrai Négatif	

Tableau 2: Matrice de confusion 2x2

On mesure la performance du modèle à partir de la métrique de précision. La précision (« accuracy ») est calculée de cette manière :

$$\text{Précision} = \frac{\text{Vrai Positif} + \text{Vrai Négatif}}{\text{Vrai Positif} + \text{Faux Positif} + \text{Faux Négatif} + \text{Vrai Négatif}}$$

L'objectif de l'étude est de développer un outil de prédiction pour prédire la qualité de farine produite à partir de la qualité des grains de blé utilisée. La revue de littérature ne fournit pas d'outil semblable qui réponde à cet objectif. Pour cette raison, un outil utilisant des algorithmes d'apprentissages machines a été développé. L'outil de prédiction se base sur les données d'analyses d'échantillons de blé ainsi que des recettes pour classer la qualité de la farine dans une catégorie adaptée pour certaines industries.

4 CAS D'ETUDE/VALIDATION

Afin de valider l'outil proposé, nous allons le tester avec un cas d'étude. Il n'existe présentement d'après la revue de littérature aucune méthode ou outil offrant des résultats en lien avec l'objectif étudié pour comparer les résultats de l'outil proposé. Pour cela, **la moyenne de la valeur du BEM des grains de blé pour chaque recette servira de référence.**

4.1 Contexte

L'outil a été testé sur les données d'une entreprise de transformation de grains céréaliers biologiques au Canada. L'entreprise connaît depuis quelques années une forte croissance des ventes. Cette croissance à deux chiffres implique une augmentation de la production et passe par un volume de production plus important et/ou par une meilleure productivité. En raison de machine très coûteuse et en cherchant à minimiser les pertes lors de la transformation ainsi que les temps de production, une meilleure productivité est recherchée. La prédiction de la qualité de la farine permettra éventuellement de **comprendre certains facteurs influençant la qualité finale de la farine produite et anticiper certains choix dans la transformation comme l'ajustement des recettes** pour atteindre la qualité souhaitée. Cette anticipation offrira des gains de productivité sur la chaîne de production.

Les données mises à disposition par le partenaire industriel regroupent : les analyses d'échantillons de grains de blé réceptionnés à l'usine ; les recettes ; les analyses d'échantillons de farine prélevés à la fin de la mouture et les courses de mouture (provenant du système ERP – Entreprise Resource Planning).

Le partenaire industriel possède 12 silos pour l'entreposage de grain de blé. Chaque silo contient plusieurs tonnes de grains de blé pouvant être mélangées avec d'autres lots de grains de blé de fournisseurs différents. Les grains de blé de qualité similaire sont regroupés dans un même silo.

Chaque recette est sauvegardée dans un fichier Excel sous un dossier sur le serveur de la compagnie. Le fichier recette précise les valeurs des caractéristiques moyennes des grains de blé présent dans chaque silo, les quantités des grains de blé et proportions utilisées de chaque silo ainsi que la quantité totale de blé à utiliser. Les caractéristiques moyennes des grains de blé sont calculées automatiquement à partir du fichier Excel répertoriant les résultats d'analyses d'échantillons de grains de blé réceptionnés à l'usine.

4.2 Mise en œuvre de la démarche

4.2.1 Phase 1 : collecter les données des recettes

De nombreuses données sont enregistrées et disponibles à travers différentes interfaces et systèmes informatiques. Dans cette première phase, nos besoins se porteront sur les recettes ainsi que les données d'analyses réalisées au laboratoire pour identifier la qualité des blés ainsi que la qualité de la farine produite. La collecte de ces données et l'étiquetage de ces données sont primordiales pour avoir des exemples et fournir un contexte afin que le modèle d'apprentissage machine puisse en tirer des enseignements. L'étiquetage est rendu possible à partir du numéro de course. Chaque nouvelle recette est liée à un numéro de course créée automatiquement à l'aide du système ERP. Toutefois, certaines courses de mouture ne possèdent pas de recette si aucun mélange n'est nécessaire. Les liens entre les recettes et les numéros de courses n'étant pas sauvegardés sur le système ERP sont retrouvés à travers les courriels du partenaire industriel ainsi que des fichiers Excel.

L'ensemble des données au niveau des recettes, des analyses et des courses de production sont inscrites manuellement pouvant accroître le risque de données manquantes ou erronées. Pour cela, après l'étiquetage des données, il est important de vérifier la cohérence des données avec plusieurs sources d'information en prélevant un échantillon de l'ensemble de données. En raison d'un temps d'étiquetage des données important, une durée de 6 mois est sélectionnée. **170 recettes sont collectées sur une période de 6 mois (janvier 2020 à juin 2020).**

4.2.2 Phase 2 : préparer les données

Dans la phase 2, nous rassemblons les données des recettes dans une table de données à l'aide d'un script Python. Le script lit chaque fichier Excel utilisant un même template pour collecter l'information présente au niveau des recettes et l'inscrire dans une table de données sous une ligne. La figure 4 montre une partie de l'entête de la table obtenue :

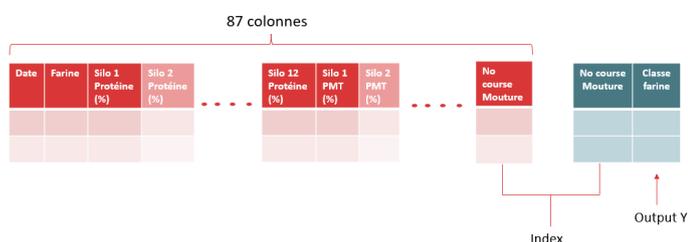


Figure 4. Table de donnée des recettes

La figure 4 présente la dimensionnalité de la table regroupant les 6 caractéristiques (A) des grains de blé ainsi que les proportions utilisées pour les 12 silos. Le numéro de course de mouture étant un numéro unique pour chaque recette est identifié comme index.

Le calcul de la moyenne de chaque caractéristique sur les 12 silos en fonction des proportions utilisées permet de réduire le nombre de colonnes de 87 à 9. La moyenne de chaque caractéristique ($a_{i,j}$) est calculée par cette formule :

$$\sum_{k=1}^{12} (\text{Silo } k A_j) * (\text{Silo } k \text{ Mélange}) = a_{i,j}, j = 1, \dots, 6 \quad (1)$$

Les recettes sont identifiées par la date et le type de farine. Ces deux données permettent de vérifier sur un échantillon d'exemple collecté l'exactitude des informations et d'ajouter le numéro de course de mouture associé à la recette. Le numéro de course permet ensuite l'étiquetage de la donnée de sortie sur l'ensemble de données. Finalement la date et le type de farine sont supprimés de la table de donnée et la donnée de sortie y est ajoutée afin que le modèle s'entraîne uniquement à partir des données décrivant la qualité des grains de blé.

Dans l'ensemble de données, s'il n'est pas possible d'identifier la recette à son numéro de course en raison d'un manque d'information ou d'information non cohérente, les recettes sont supprimées de la table de données. Les doublons sont également retirés. Au total 15 recettes ont été supprimées de la table de donnée.

À partir de l'expertise du partenaire industriel, la qualité de la farine est répartie en 7 classes. Seulement en raison d'une forte disparité du nombre d'exemples présente dans chacune des classes, le nombre de classes a été réduit à 4. Les 4 classes sont notées sur une échelle de 0 à 3 avec la classe 0 représentant une farine possédant un BEM faible et la classe 3 représentant une farine possédant un BEM très élevé.

Finalement en présence d'un débalancement de nos données, 30 recettes identifiées spécifiquement pour les catégories débalancées (classe 0 et 3) sont ajoutées à l'ensemble de données. Ces nouvelles recettes ont été réalisées durant la période entre juillet et octobre 2020. Suite à la réduction des classes et de l'ajout des recettes, la répartition des exemples se présente comme suit : 32 pour la classe 0, 49 pour la classe 1, 63 pour la classe 2 et 37 pour la classe 3.

4.2.3 Phase 3 : entraîner le modèle

L'entraînement du modèle de l'arbre de décision est réalisé à l'aide de la librairie Scikit-learn sur python. L'ensemble de données préparé est divisé en deux ensembles de la manière suivante : **85% pour les données d'entraînement et 15% pour les données tests.** Cette division est réalisée par un échantillonnage stratifié de la variable de sortie. Les paramètres énoncés dans la section 3.3 ont été optimisés par validation croisée avec 5 groupes différents échantillonnés de l'ensemble de données d'entraînement. **La fonction « GridSearchCV » est utilisée pour tester une multitude de paramètres souhaités** rapidement et connaître celles qui offrent la meilleure performance. Pour rebalancer l'ensemble des données, des poids sont ajoutés pour pénaliser la classification erronée sur des classes minoritaires (classe 0 et 3). Les paramètres optimaux obtenus pour le modèle sont : « max_depth » = 5 ; « min_samples_leaf » = 6 ; « min_samples_split » = 2 et « class_weight » = {'a' : 5}.

4.2.4 Phase 4 : valider le modèle

Le modèle d'apprentissage est validé à partir des données tests. La précision du modèle sur les données d'entraînement est d'environ 58% avec un écart type d'environ 5.6% entre les 5 groupes de la méthode de la validation croisée. Ce faible écart type indique que le modèle est assez robuste. **La précision du modèle sur les données de test est d'environ 64%**, ce qui est meilleur que la précision obtenue sur les données d'entraînement. Cela indique que le modèle ne fait pas de surapprentissage.

4.3 Analyse des résultats

Nous obtenons **un arbre à 5 niveaux avec 15 feuilles. Les 2 variables les plus importantes** d'après l'arbre obtenu sont **la moyenne du BEM** ainsi que **le taux de protéine moyen** de chaque recette. La figure 5 montre les résultats de classification du modèle à partir des données tests dans une matrice de confusion.

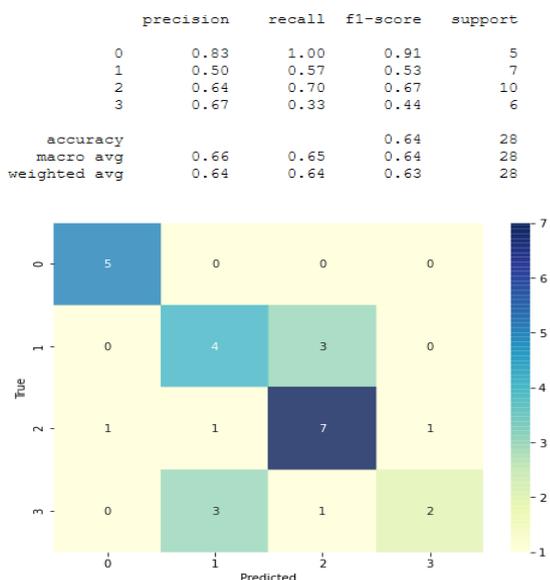


Figure 5. Matrice de confusion du modèle de l'arbre de décision

On observe d'après la figure 5 que le modèle classe très bien les farines de la classe 0 avec un BEM faible. Il classe relativement bien les farines de la classe 1 et 2. Toutefois les farines ayant un BEM élevé (classe = 3) sont assez mal classées et réalisent de mauvaises prédictions. Sur les 6 exemples présents appartenant à la classe 3, seulement 2 sont correctement prédits par le modèle.

En utilisant à partir des données tests seulement la moyenne des BEM (moyenne pouvant être calculée lors des recettes), la méthode offre une précision d'environ 57%, ce qui est légèrement inférieur aux résultats obtenus par l'arbre de décision.

Sur le cas d'étude, **l'outil de prédiction a permis de comprendre les variables d'intérêts telles que : le BEM, le taux de protéine, le taux d'humidité et le taux de cendre.** L'arbre permet également de visualiser et comprendre comment la classification est réalisée. Dans le cas d'étude, **le modèle performe légèrement mieux qu'une classification réalisée par la moyenne des BEM.** Pour de meilleures performances, d'autres modèles d'apprentissages tels que les réseaux neurones ou les arbres de type random forest peuvent être envisagés. Cependant, ces modèles se traduisent par une perte d'interprétabilité des résultats.

5 CONCLUSION

La qualité de la farine biologique varie tout au long de l'année selon la qualité des matières premières dont les caractéristiques changent au cours du temps en fonction des conditions environnementales. Pour cela, l'objectif de cet article était de proposer un outil capable de prédire la qualité de la farine après la mouture à partir des données de la qualité des grains présents dans des silos. Cet outil s'appuie sur 4 phases. Celles-ci sont (1) collecter les données, (2) préparer les données collectées, (3) entraîner le modèle d'apprentissage à partir des données préparées et finalement (4) valider le modèle. La démarche a été testée sur un cas d'étude. Les résultats ont été significatifs, car on note une performance légèrement supérieure à un simple calcul de la moyenne du BEM. Toutefois, on observe d'après les résultats la présence de mauvaises classifications dans chaque classe et plus particulièrement dans la dernière classe. Ces mauvaises classifications peuvent s'expliquer par un manque d'information au niveau des paramètres de production concernant la transformation des grains de blé. Le modèle pourrait être amélioré en incluant certains paramètres clés de production impactant la qualité de la farine (l'écartement des rouleaux d'écrasement, la durée de trempage des grains, le débit massique des grains acheminés vers les rouleaux d'écrasement sont des exemples de paramètres intéressants à intégrer au modèle).

En dehors de l'ajout de paramètres, une comparaison de la performance du modèle avec une approche neuronale entraîné sur un volume de donnée plus important semble pertinente. De plus, l'utilisation de l'ensemble des données de chaque silo, autrement que par le calcul de la moyenne pondérée de chaque caractéristique, doit être étudiée. La moyenne pondérée engendre une perte d'information, pouvant possiblement être utile à l'apprentissage du modèle et le rendre plus robuste.

La prédiction de la qualité de la farine offre la possibilité d'optimiser les recettes en termes de cout et de volume selon la qualité et quantité de grains de blé disponibles. Une piste qui sera étudiée plus en détail avec le partenaire industriel La Milanaise. La qualité de la farine est un paramètre important par la suite pour obtenir une fermentation optimale et des propriétés d'élasticité, de ténacité et d'extensibilité de la pâte propre à chaque client en fonction de la qualité du pain recherchée.

6 REMERCIEMENTS

Nous tenons à remercier notre partenaire industriel, La Milanaise, d'avoir participé au développement de cet outil ainsi que le MAPAQ (projet IA119053) pour leur soutien financier.

7 REFERENCES

Adebayo, S. E., Hashim, N., Abdan, K., Hanafi, M., & Mollazade, K. (2016). Prediction of quality attributes and ripeness classification of bananas using optical properties. *Scientia Horticulturae*, 212, pp. 171-182.

Alvarez, R. (2009). Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *European Journal of Agronomy*, 30(2), pp. 70-77.

Bao, Y. D., Liu, F., Kong, W. W., Sun, D. W., He, Y., & Qiu, Z. J. (2014). Measurement of Soluble Solid Contents and pH of White Vinegars Using VIS/NIR Spectroscopy and Least

- Squares Support Vector Machine. *Food and Bioprocess Technology*, 7(1), pp. 54-61.
- Barbon, A. P. A. C., Barbon, S., Mantovani, R. G., Fuzyi, E. M., Peres, L. M., & Bridi, A. M. (2016). Storage time prediction of pork by Computational Intelligence. *Computers and Electronics in Agriculture*, 127, pp. 368-375.
- Borghini, B., Giordani, G., Corbellini, M., Vaccino, P., Guermanni, M., & Toderi, G. (1995). Influence of Crop-Rotation, Manure and Fertilizers on Bread-Making Quality of Wheat (*Triticum-Aestivum* L). *European Journal of Agronomy*, 4(1), pp. 37-45.
- Burkov, A. (2019a). Chapter 3: Fundamental Algorithms. The Hundred-Page Machine Learning Book. pp. 36-55
- Burkov, A. (2019b). Chapter 5: Basic Practice. The HundredPage Machine Learning Book. pp. 68-87
- Burkov, A. (2020). Chapter 4: Feature Engineering. Machine Learning Engineering. True Positive Inc. pp. 79-124
- Cappelli, A., Oliva, N., & Cini, E. (2020). Stone milling versus roller milling: A systematic review of the effects on wheat flour quality, dough rheology, and bread characteristics. *Trends in Food Science & Technology*, 97, pp. 147-155.
- Cocchi, M., Corbellini, M., Foca, G., Lucisano, M., Pagani, M. A., Tassi, L., & Ulrici, A. (2005). Classification of bread wheat flours in different quality categories by a wavelet-based feature selection/classification algorithm on NIR spectra. *Analytica Chimica Acta*, 544(1-2), pp. 100-107.
- Côté, M.-H. (2018). Évaluation de différents taux d'humidité de la récolte du blé panifiable tout en considérant la qualité boulangère et l'aspect économique. Groupe multiconseil agricole Saguenay-Lac-Saint-Jean.
- Datascientest. (2020). *Classification pénalisée*. Tiré de: <https://datascientest.com/glossary/classification-penalisee>
- Dubey, B. P., Bhagwat, S. G., Shouche, S. P., & Sainis, J. K. (2006). Potential of artificial neural networks in varietal identification using morphometry of wheat grains. *Biosystems Engineering*, 95(1), pp. 61-67.
- El-Bendary, N., El Hariri, E., Hassanien, A. E., & Badr, A. (2015). Using machine learning techniques for evaluating tomato ripeness. *Expert Systems with Applications*, 42(4), pp. 1892-1905.
- Fournier, M.-E. (2020). Vendre de la farine, ce n'est pas de la tarte! . *La Presse*. Tiré de: <https://www.lapresse.ca/affaires/entreprises/2020-04-18/vendre-de-la-farine-ce-n-est-pas-de-la-tarte>
- Fu, B. X., Wang, K., & Dupuis, B. (2017). Predicting water absorption of wheat flour using high shear-based GlutoPeak test. *Journal of Cereal Science*, 76, pp. 116-121.
- Future Market Insights. (2020). *Organic Wheat Flour Market: Global Industry Analysis 2012 - 2016 and Opportunity Assessment; 2017 - 2027*.
- Règlement sur les aliments et drogues (C.R.C., ch. 870), (2020). Tiré de: https://laws-lois.justice.gc.ca/fra/reglements/c.r.c.,_ch._870/page-65.html
- Gresle, E. (2013). Du Farinograph au Glutopeak. *Tribune des Constructeurs*, pp. 27-33. <https://www.aemic.com/uploads/pdfs/IdC184.27.pdf>
- Johansson, E., & Svensson, G. (1998). Variation in bread-making quality: Effects of weather parameters on protein concentration and quality in some Swedish wheat cultivars grown during the period 1975-1996. *Journal of the Science of Food and Agriculture*, 78(1), pp. 109-118.
- Lakshminarayan, K., Harp, S. A., Goldman, R., & Samad, T. (1996). *Imputation of missing data using machine learning techniques*. AAAI.
- Liu, C., Yang, S. X., & Deng, L. (2015). A comparative study for least angle regression on NIR spectra analysis to determine internal qualities of navel oranges. *Expert Systems with Applications*, 42(22), pp. 8497-8503.
- MAPAQ. (2019). *Agriculture biologique*. Tiré de: <https://www.mapaq.gouv.qc.ca/fr/Productions/Production/agriculturebiologique/Pages/alimentsbio.aspx>
- Marlin, B. M. (2008). *Missing Data Problems in Machine Learning*, University of Toronto. Tiré de: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.160.8408&rep=rep1&type=pdf>
- Mutlu, A. C., Boyaci, I. H., Genis, H. E., Ozturk, R., Basaran-Akgul, N., Sanal, T., & Evlice, A. K. (2011). Prediction of wheat quality parameters using near-infrared spectroscopy and artificial neural networks. *European Food Research and Technology*, 233(2), pp. 267-274.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, pp. 2825-2830.
- Ronowicz, J., Thommes, M., Kleinebudde, P., & Kryszynski, J. (2015). A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm. *European Journal of Pharmaceutical Sciences*, 73, pp. 44-48.
- Sablani, S. S., Baik, O. D., & Marcotte, M. (2002). Neural networks for predicting thermal conductivity of bakery products. *Journal of Food Engineering*, 52(3), pp. 299-304.
- Scikit-learn developers. (2020). *Decision Trees*. Tiré de: <https://scikit-learn.org/stable/modules/tree.html>
- Torbica, A., Blazek, K. M., Belovic, M., & Hajnal, E. J. (2019). Quality prediction of bread made from composite flours using different parameters of empirical rheology. *Journal of Cereal Science*, pp. 89
- Triboi, E., Abad, A., Michelena, A., Lloveras, J., Ollier, J. L., & Daniel, C. (2000). Environmental effects on the quality of two wheat genotypes: 1. quantitative and qualitative variation of storage proteins. *European Journal of Agronomy*, 13(1), pp. 47-64.
- Visen, N. S., Paliwal, J., Jayas, D. S., & White, N. D. G. (2002). AE—Automation and Emerging Technologies: Specialist Neural Networks for Cereal Grain Classification. *Biosystems Engineering*, 82(2), pp. 151-159.
- Willer, H., & Lernoud, J. (2019). *The world of organic agriculture - Statistics & Emerging trends 2019*. IFOAM – Organics International & Research Institute of Organic Agriculture FiBL.
- Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, 4(1), pp. 23-45.
- Xie, G., & Xiong, R. (1999). Use of hyperbolic and neural network models in modelling quality changes of dry peas in long time cooking. *Journal of Food Engineering*, 41(3-4), pp. 151-162.
- Zhang, X. W., & Hu, B. G. (2014). A New Strategy of Cost-Free Learning in the Class Imbalance Problem. *Ieee Transactions on Knowledge and Data Engineering*, 26(12), pp. 2872-2885.

Implantation d'une architecture de CPS permettant l'interopérabilité des systèmes de production en contexte de PME

JEAN-REMI PIAT^{1,2}, CHRISTOPHE DANJOU^{1,2}, BRUNO AGARD^{1,2}, ROBERT BEAUCHEMIN³

¹ Laboratoire en Intelligence des Données
Département de mathématiques et génie industriel,
École Polytechnique de Montréal, CP 6079, succursale Centre-Ville, Montréal, Québec, Canada
jean-remi.piat@polymtl.ca, christophe.danjou@polymtl.ca, bruno.agard@polymtl.ca

² Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT)

³ La Meunerie Milanaise, Saint-Jean-sur-Richelieu, Québec, Canada

Résumé : L'industrie 4.0 est une stratégie basée sur les nouvelles technologies pour accroître les capacités des entreprises. Des standards et des méthodes ont été développés pour, et par, les grandes entreprises afin d'intégrer les technologies de l'industrie 4.0 aux systèmes de productions et créer une industrie intelligente. Cependant, l'intégration de ces technologies reste problématique dans beaucoup d'environnements industriels et notamment dans les petites et moyennes entreprises (PMEs). D'une part, l'hétérogénéité physique et informatique des systèmes d'entreprise complexifie la création d'un cadre d'interopérabilité, et d'autre part, les limites en ressource financière, en connaissances et en culture technologique contraignent les PME dans l'adoption d'une stratégie 4.0. L'article présente une architecture de CPS (système cyber-physique) pour rendre les systèmes de pilotage plus réactifs et mettre en place un cadre d'interopérabilité. Une méthodologie d'implantation agile est proposée pour adapter cette architecture au contexte des PME. Un cas d'étude sur une meunerie biologique de la région de Montréal est introduit pour valider la méthodologie et présenter les premiers résultats.

Abstract: Industry 4.0 is a strategy based on emerging technologies to improve capabilities in enterprises. Standards and methods have been introduced for, or by, large firms for their own benefits to integrate industry 4.0 technologies in production systems and develop a smart factory. However, the integration of these systems is still an issue in many industrial environments, especially in small and medium enterprises (SMEs). First, physical and digital heterogeneity increases the complexity of interoperability introduction, on the other hand, the financial resources, knowledge resources and technology awareness limitations constrain SMEs in strategy 4.0 adoption. The paper presents a CPS (cyber-physical system) architecture to make the control systems more reactive and establish an interoperability frame. An agile methodology is proposed to adapt this framework in the SMEs and a case study of a Montreal biological flour milling is described to validate the methodology and present the first results.

Mots clés – Interopérabilité, Système Cyber-Physique, Industrie 4.0, PME, Architecture.

Keywords – Interoperability, Cyber-Physical System, Industry 4.0, SME, Architecture.

1 INTRODUCTION

L'industrie est en constante évolution et a déjà connu trois révolutions industrielles par le passé pour s'adapter au rythme des attentes de notre société (Drath et Horch, 2014). De nos jours, l'industrie traditionnelle doit satisfaire de nouveaux standards en fournissant des produits et des services personnalisés de haute qualité tout en maintenant le profit d'une production de masse (Lasi et al., 2014).

L'industrie 4.0 est une stratégie apparue en 2011 pour répondre à ce défi (Drath et Horch, 2014). L'industrie 4.0 s'appuie sur un ensemble de technologies et de concepts pour établir une connectivité ubiquitaire et réorganiser la chaîne de valeur (Hermann et al., 2015). L'objectif est de transformer les processus, les produits et les services par une prise de décision décentralisée en temps réel pour développer de nouvelles capacités (Danjou et al., 2017). L'industrie intelligente est

l'achèvement de ce système complètement connecté qui opère en autonomie grâce à la génération, au transfert et à l'analyse de données (Lasi et al., 2014).

Un des concepts essentiels de l'industrie 4.0 est l'interopérabilité (Hermann et al., 2015). L'interopérabilité entre des systèmes représente leur capacité à communiquer et à coopérer entre eux malgré leur hétérogénéité et leur peu de connaissances qu'ils ont les uns des autres (ISO-2382-1, 1993). Les systèmes considérés peuvent comprendre des machines de productions spécialisées, des capteurs, des produits ou des systèmes d'information. Chaque entreprise doit sélectionner des technologies en fonction de ses propres systèmes et de ses besoins pour mettre en place son cadre d'interopérabilité et développer sa stratégie d'industrie 4.0.

Les standards et les protocoles sont des facteurs clés dans l'architecture d'une industrie intelligente (Hermann et al., 2015).

L'utilisation de standards permet d'identifier et de positionner une architecture par rapport aux autres mais aussi de comparer des concepts, des principes, des méthodes, des modèles et des outils avec d'autres domaines (Vernadat, 2010). Cependant, les standards sont multiples donc les systèmes d'une même entreprise peuvent utiliser des standards et des protocoles différents. L'hétérogénéité des systèmes d'une entreprise rend l'implantation d'architecture standard couteuse et requiert des compétences multidisciplinaires.

Le contexte des petites et moyennes entreprises (PMEs) accentue les difficultés d'implantation d'une stratégie d'industrie 4.0 (Masood et Sonntag, 2020). Les méthodes, les architectures et les technologies utilisées dans les grandes entreprises doivent être adaptées pour répondre aux contraintes financières, en ressources humaines et en connaissances des PME, tout en s'accordant avec leur environnement hétérogène qui comprend beaucoup de données dispersées et isolées.

Ceci pose la question de comment peut-on implémenter un système de pilotage réactif et flexible dans un environnement de PME de transformation agroalimentaire.

La structure de l'article se décompose de la manière suivante : la section 2 présente l'état de l'art. Tout d'abord, nous présentons les niveaux et les standards de l'interopérabilité (section 2.1), ensuite, les systèmes cyber-physiques (Cyber-physical system CPS) sont définis et une architecture de CPS est introduite (section 2.2). La section 2.3 porte sur les difficultés d'implantation de CPS dans le contexte des PME et la section 3 présente la méthodologie d'implantation proposée. La section 4 expose un cas d'étude dans le domaine agroalimentaire. Nous présentons le contexte du cas d'étude (4.1) et montrons la mise en œuvre détaillée de la méthode proposée sur un exemple (4.2). Finalement, la section 5 rappelle les résultats et limites de la méthode proposée et propose quelques perspectives.

2 ÉTAT DE L'ART

Dans cette section, une architecture de système cyber-physique est décrite pour mettre en place un cadre d'interopérabilité. Le contexte des petites et moyennes entreprises (PMEs) est ensuite étudié pour comprendre leurs contraintes spécifiques.

2.1 Interopérabilité

L'interopérabilité est une notion essentielle dans l'ère de l'industrie 4.0 et fait l'objet d'études et de projets académiques pour en définir des structures et des méthodes d'implantation. L'institut des ingénieurs électriciens et électroniciens (IEEE) l'a défini comme étant « la capacité de deux ou plusieurs systèmes ou éléments à échanger des informations et à utiliser les informations ainsi échangées » (Geraci, 1990). Des structures et des standards ont ensuite été introduits pour comprendre, analyser et guider la mise en œuvre d'un cadre d'interopérabilité (Valle et al, 2019).

2.1.1 Niveaux d'interopérabilités

L'interopérabilité peut être décomposée en trois niveaux par (EIF, 2004) :

- L'interopérabilité organisationnelle qui concerne les interactions entre les groupes d'affaires, les processus d'affaires et les personnes au travers de toute l'organisation. Elle représente la capacité de chaque unité à fournir, recevoir

et utiliser des services pour assurer le fonctionnement global de l'entreprise.

- L'interopérabilité sémantique qui assure que le partage d'information et de service préserve le flux sémantique. L'objectif est de s'assurer que toutes les informations échangées sont compréhensibles par toutes les applications même si celles-ci n'ont pas été développées dans ce but. Chaque système, même hétérogène, doit être capable de découvrir des données, de se les représenter et de leur donner un contexte pour interpréter les données partagées de façon cohérente.
- L'interopérabilité technique qui couvre l'ensemble des problèmes techniques pour connecter les systèmes informatiques. Son but est de faciliter les communications en utilisant des protocoles permettant les échanges de données et de messages entre les applications du système. Ce type d'interopérabilité est associé avec les composants physiques, informatiques et réseaux permettant la communication inter-machine.

Ces trois niveaux d'interopérabilité sont les plus récurrents dans la littérature, ils traitent des différents aspects du système et sont nécessaires à l'intégration d'un cadre d'interopérabilité (Valle et al., 2019; Vernadat, 2010).

2.1.2 Standard

Une architecture orientée service (Service oriented architecture SOA) est une des technologies pour établir l'interopérabilité sémantique (Valle et al., 2019). Elle utilise des interfaces logicielles, appelées services, pour créer une infrastructure facilitant les interactions et les échanges d'informations entre les systèmes. Ces interfaces spécifient les opérations (ce que fait le service), les paramètres partagés lors d'une opération, et les protocoles qui définissent l'échange des paramètres. L'interaction entre les services dans une architecture SOA est confirmée si le receveur du service comprend les données transmises par le fournisseur du service (X. Chen et Voigt, 2020).

Une architecture SOA conserve le flux sémantique du système et fournit une forte interdépendance entre les services des trois groupes fonctionnels (la production, la qualité et le personnel) d'une entreprise (Paniagua et al., 2019). L'interdépendance de ces groupes fonctionnels forme l'intégration horizontale des systèmes d'une entreprise (ANSI/ISA-95, 2013). L'intégration verticale porte sur la connexion et la communication des systèmes à tous les niveaux pour donner à la direction des informations de production à jour et au bon moment. Les intégrations horizontale et verticale sont des notions introduites par la norme ISA-95 (ANSI/ISA-95, 2013). La norme ISA-95 contient cinq parties, publiées au début des années 2000, et définit l'interopérabilité entre le management corporatif et le système de production. La norme décrit une architecture pyramidale standard en cinq niveaux pour représenter les fonctionnalités de l'entreprise. Des systèmes d'information, dont les ERP, les MES et les SCADA, ont été développés pour fournir les services correspondant aux fonctionnalités de chaque niveau et assurer l'intégration verticale et horizontale. Ces systèmes spécialisés apportent une grande diversité et complexifient la mise en place d'un cadre d'interopérabilité. L'hétérogénéité des systèmes comprend leur fournisseur, leurs protocoles de communication, leur structure de données, leurs interfaces d'accès, leur fréquence de

communication et le volume de données échangées (Kajati et al., 2019).

L'Open Platform Communications Unified Architecture (OPC-UA) est un standard (IEC-62541-1, 2020) utilisant une architecture SOA et assurant l'intégration verticale des systèmes de production présents dans la norme ISA-95 (Riedl et al., 2014). L'OPC-UA a été développé pour répondre aux problématiques d'interfaces propriétaires, des traitements asynchrones et des flux d'informations complexes entre les systèmes. Les communications effectuées entre l'ensemble des interfaces sont standardisées pour limiter les traducteurs de chaque système. Autrement dit, l'OPC-UA définit une structure d'interface permettant à tous les systèmes de fournir des services et de comprendre les services des autres afin de créer des interactions. Le receveur d'un service, i.e. un client, et le fournisseur du service, i.e. un serveur, communiquent en s'échangeant des informations sur des objets qui modélisent les systèmes physiques de l'entreprise.

L'architecture pyramidale de l'ISA-95, est une représentation générique des interactions entre les différents systèmes d'une industrie (Colombo et al., 2017; Sinha et Roy, 2020). Elle définit des services nécessaires, mais pas suffisants pour les nouvelles capacités de l'industrie 4.0. De plus, la structure rigide des systèmes d'information implique une prise de décision hiérarchisée et distribuée qui n'est pas adaptée aux capacités flexibles, évolutives et réactives des nouveaux systèmes de l'industrie 4.0.

2.2 Système Cyber-Physique (CPS)

Dans une stratégie d'industrie 4.0, les systèmes d'entreprise sont soutenus par les nouvelles technologies, dont l'intelligence artificielle, le nuage informatique (cloud computing), l'internet des objets, les systèmes cyber-physiques, la cybersécurité, les machines autonomes, la communication inter-machines, les systèmes de simulations et la réalité augmentée (Danjou et al., 2017). L'augmentation de l'utilisation de ces technologies en industrie transforme l'environnement industriel en un système caractérisé par une forte coordination entre des éléments informatiques et physiques qui forme un système cyber-physique (Cyber-physical system CPS) industriel (Colombo et al., 2017). Ce système est le pilier de la connectivité permettant une prise de décision en temps réel et décentralisée.

2.2.1 Définition

Le terme CPS a été introduit en 2006 par la fondation nationale des sciences (National Science Foundation NSF) américaine pour désigner l'intégration des systèmes informatiques avec les processus physiques (Colombo et al., 2017; Sinha et Roy, 2020). Ce nouveau système enrichit les interactions entre les services des systèmes industriels pour augmenter l'intégration verticale et horizontale. Le CPS permet d'aplatir l'architecture pyramidale définie par l'ISA-95 (Riedl et al., 2014). Les relations entre les systèmes d'information de la norme ISA-95 sont réutilisées pour permettre des communications directes entre les systèmes qui devaient auparavant passer par d'autres niveaux, créer un environnement évolutif pouvant accueillir de nouveaux systèmes et enfin fournir de nouveaux services en explorant les données de tous les systèmes.

Les performances d'un CPS dépendent de sa capacité à collecter efficacement et analyser une grande quantité de données provenant de différentes sources pour surveiller, contrôler et optimiser l'environnement industriel de façon durable (Colombo et al., 2017). Les principaux défis sont l'hétérogénéité et la complexité des composants ainsi que les interactions entre le domaine virtuel et physique (Sinha et Roy, 2020). L'objectif est d'introduire une négociation entre les systèmes connectés permettant de reconfigurer les processus et les ressources en fonction de la dynamique de production. Les ressources humaines de l'entreprise doivent être capables d'interagir avec le CPS pour prendre des décisions, comprendre et optimiser la planification, les installations, la production, les services et la maintenance (Colombo et al., 2017 ; Sinha et Roy, 2020).

2.2.2 Architecture d'un CPS

Une architecture en quatre niveaux, physique, réseau, nuage informatique et terminal de contrôle, peut être considérée pour étudier les différentes technologies et problématiques d'un CPS (B. Chen et al., 2018; Wan et al., 2016; Wang et al., 2016a; Wang et al., 2016b).

Les équipements présents sur la chaîne de production forment le niveau physique. Ces systèmes sont connectés avec les équipements de communication performants du niveau réseau qui utilise des technologies informatiques pour transmettre des informations au niveau du nuage informatique. Ces technologies regroupent les protocoles SOA comme l'OPC-UA, les réseaux industriels de capteurs sans-fils (Industrial Wireless Sensors Networks IWSNs), les réseaux définis par logiciels (Software-Defined Networks SDNs), les bus de terrain et d'autres technologies informatiques. Le réseau doit assurer une qualité de service, des communications fiables et la coopération entre les équipements pour fournir de faibles délais de communication, une grande densité d'accès, une faible consommation d'énergie et une synchronisation très précise.

Le nuage informatique regroupe les ressources informatiques pour gérer les données. Des techniques d'exploration de données sont utilisées pour créer les services nécessaires aux prises de décisions et à l'optimisation des processus. Les technologies d'exploration de données utilisent une grande quantité de données pour extraire de l'information à la l'aide d'algorithmes informatiques. Ces informations servent à prévoir et modéliser des événements pour maîtriser les problèmes de productions, de qualités et de maintenance. Les résultats sont communiqués aux autres systèmes ou au terminal de contrôle qui regroupe toutes les données importantes et leurs informations connexes sur une application web appelée interface homme-machine (IHM). Cette application fournit un support efficace, spécialisé et visuel pour gérer les interactions entre les systèmes et les opérateurs.

Le CPS traite des problématiques d'interopérabilité technique et sémantique grâce aux technologies informatiques. La représentation sémantique des différents concepts et leurs liens peuvent notamment être modélisés par une ontologie qui s'appuie sur des technologies de base de données pour enrichir les modèles classiques et gérer des systèmes plus complexes et intelligents. Le CPS permet aux ressources physiques de fournir des informations et des services au domaine virtuel, mais aussi d'en recevoir malgré

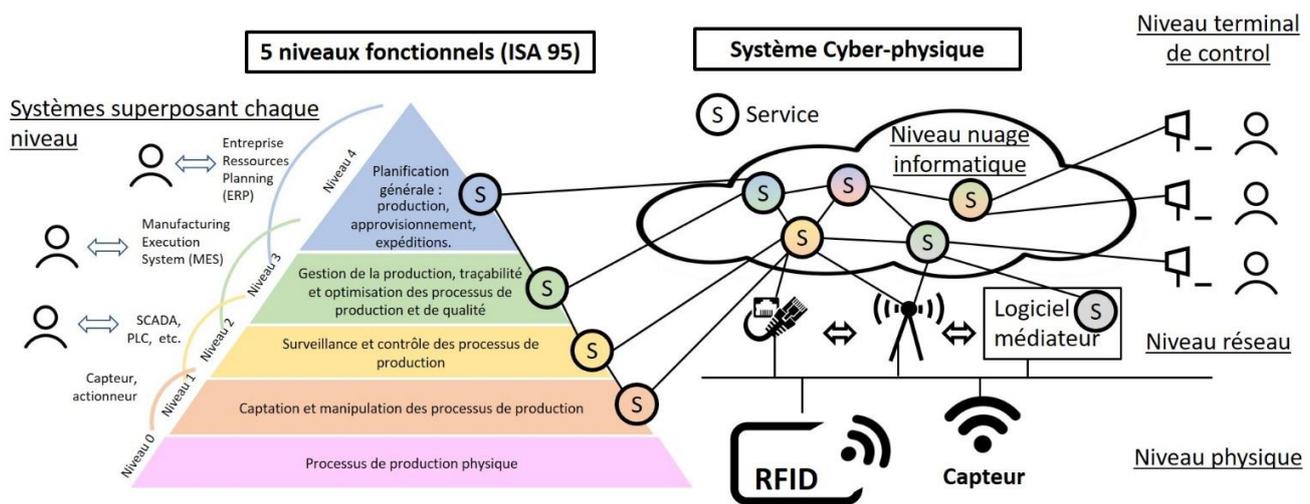


Figure 1 : Superposition d'un CPS sur l'architecture générique d'une entreprise

l'hétérogénéité des systèmes spécialisés de hautes performances du plancher de production.

Les systèmes d'information sont contraints dans leurs interfaces et ne peuvent pas accueillir tous les nouveaux systèmes alors que les CPS offrent une architecture flexible par construction. Les composants des niveaux physique et réseau ne sont pas contraints par les capacités des autres systèmes. Chaque équipement est responsable de ses services, et ceux-ci peuvent être mis en relation au niveau du nuage informatique.

Le CPS permet aussi d'implanter un système de pilotage plus réactif en distribuant le pouvoir décisionnel de façon hétérarchique (Trentesaux, 2009).

L'hétérarchie a été définie par Trentesaux grâce à la théorie des graphes par « un graphe orienté, composé de nœuds représentant des entités décisionnelles et d'arcs représentant l'interaction de maître-esclave d'une entité décisionnelle (maître) avec une autre entité (esclave), est appelé un graphe d'influence. Si chaque nœud peut être considéré comme maître et comme esclave, sans de lien de hiérarchie, alors le graphe est fortement connecté. Ces connexions définissent une hétérarchie. » Une architecture hétérarchique est donc l'inverse d'une hiérarchie mais les deux architectures peuvent être complémentaires pour former une hétérarchie au sens large (Trentesaux, 2007). Une hétérarchie au sens large représente une organisation hiérarchique comportant au moins une sous-organisation hétérarchique.

Une hétérarchie peut être mise en place grâce au terminal de contrôle du CPS en responsabilisant les opérateurs de façon à ne plus avoir de relations de supériorité ou de subordination. Cette nouvelle organisation permet d'augmenter la réactivité localement en éliminant les temps de communication. Dans un système de contrôle hiérarchique, la communication des informations entre les niveaux, la prise de décision et l'application des décisions génèrent des latences et des instabilités (Trentesaux, 2009). Le principe est donc de permettre aux entités décisionnelles de travailler ensemble pour réagir rapidement au lieu de demander une prise de décisions aux niveaux supérieurs.

2.3 Contexte PME

La majorité des recherches sur l'industrie 4.0 sont réalisées par et pour de grandes organisations et des entreprises multinationales (EMNs) (Mittal et al., 2018 ; Masood et Sonntag, 2020). Les stratégies d'industrie 4.0 utilisent des standards et des protocoles pour créer des systèmes complexes. Les modèles, les architectures et les technologies ainsi développés et utilisés sont liés entre eux pour permettre le développement de nouvelles capacités. L'utilisation non systématique des standards dans les PME crée des différences entre les PME et avec les EMNs au niveau de leur fonctionnement, de leur structure et de leurs ressources (Masood et Sonntag, 2020). Les environnements uniques des PME rendent la recherche de modèles et de méthodes difficile.

Les défis auxquels les PME font face ont été répertoriés dans la littérature (X. Chen et Voigt, 2020 ; Jordan et al., 2017 ; Masood et Sonntag, 2020). Premièrement, les PME rencontrent des limitations financières lors de la mise en place des projets technologiques complexes de l'industrie 4.0. Deuxièmement, ces projets nécessitent des connaissances pour personnaliser l'intégration des systèmes informatiques dans un environnement hétérogène. Des compétences générales en programmation, en ingénierie, en technologie et en management sont nécessaires pour évaluer le projet et des connaissances spécialisées sur les technologies utilisées et sur le domaine industriel sont ensuite requises. Troisièmement, la culture technologique de la PME joue un rôle important sur les difficultés et les bénéfices que va apporter le projet. Le manque d'expériences, de connaissances et de partage sur les principes d'industrie 4.0 en PME rend l'adoption et l'implantation difficile.

Une étude a été effectuée pour évaluer les facteurs influençant les défis rencontrés et les bénéfices apportés par la mise en place d'un CPS (Masood et Sonntag, 2020). Les bénéfices regroupent la réduction des coûts et l'augmentation de l'efficacité, de la qualité et de la flexibilité des opérations. L'étude montre que les bénéfices sont plus conséquents dans les plus grandes PME et dans celles ayant une conscience des principes d'industrie 4.0. En revanche, la complexité des processus de production et des technologies utilisées augmente les difficultés d'implantation du projet.

3 METHODOLOGIE D'IMPLANTATION

L'État de l'art montre que les stratégies et les technologies d'industries 4.0 se basent sur des standards pour implanter de nouveaux concepts, comme l'interopérabilité, et apporter de nouvelles capacités et de nouveaux services dans l'entreprise. Les systèmes cyber-physiques (CPS) sont une des technologies développées dans les grandes entreprises pour connecter l'ensemble des processus physiques au monde virtuel. Leur architecture permet de réutiliser les services des systèmes existants et de les coupler avec des services d'autres systèmes isolés pour développer un système de pilotage flexible et réactif. Dans les PME, le développement d'un CPS est contraint d'une part, par l'absence de certains standards et services des systèmes d'informations des grandes entreprises et d'autre part, par les limites en ressources financières, en connaissances et en culture technologique.

Nous proposons donc une méthodologie pour définir les services et les fonctionnalités d'une architecture standard ISA-95 en contexte de PME ainsi que pour mettre en place une structure de CPS évolutive, flexible et adaptée. La figure 1 schématise le CPS superposant l'architecture ISA-95 générique d'une entreprise. L'architecture ISA-95 de gauche explicite les cinq niveaux fonctionnels essentiels avec la prise de décision hiérarchique des systèmes d'information tandis que l'architecture de droite représente la vision hétéroarchitecturale et flexible de la stratégie 4.0. Une approche basée sur les besoins de l'atelier de production permet de faciliter la virtualisation des processus physiques (Jordan et al., 2017 ; Lu et Xu, 2018). La figure 2 présente les quatre phases d'implantations utilisées dans le cas d'étude. Cette méthodologie repose sur des pratiques agiles qui sont très utilisées dans le développement de logiciels informatiques (Lu et Xu, 2018).

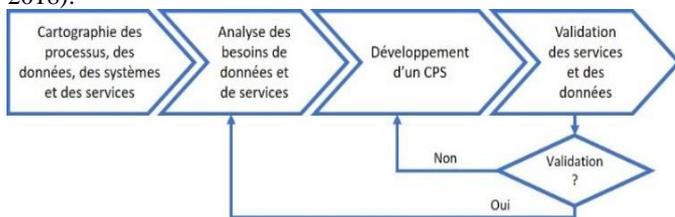


Figure 2 : Méthodologie d'implantation d'un CPS.

3.1 Phase 1 : Cartographie des processus, des données, des systèmes et des services.

Dans cette première phase, l'objectif est de cartographier les données et les services utilisés sur la chaîne de valeur. Cette démarche commence par une cartographie des processus pour définir la chaîne de valeur de l'étude. Les différentes données manipulées par les opérateurs et les systèmes sont ensuite répertoriées. Pour finir, chaque donnée est associée à un système qui sera étudié pour connaître ses standards, ses protocoles et les services qu'il offre.

3.2 Phase 2 : Analyse des besoins de données et de services

L'identification des besoins est un processus pour définir les ressources de production à virtualiser. La connaissance des différentes caractéristiques d'un CPS et la cartographie de l'existant permettent de déterminer les données et les services manquants pour répondre aux exigences du système.

L'utilisation d'une méthode de développement agile permet de répondre aux besoins de façon dynamique. Cette méthode permet de délivrer des interfaces logicielles continuellement pour répondre aux changements permanents des besoins (Abrahamsson, 2002). Chaque identification d'un besoin fait l'objet d'une période de développement et d'une mise en place rapide. Les modifications des besoins existants et l'ajout de nouveaux besoins sont pris en compte et implémentés de façon dynamique.

3.3 Phase 3 : Développement d'un système cyber-physique

La mise en place du système répond à un besoin en fournissant l'accès à de nouvelles données et en créant de nouveaux services. Une stratégie de migration permet de réutiliser les réseaux existants et de former de nouvelles communications entre les systèmes à certains points d'accès. Les boucles de contrôle existantes ne changent pas et de nouveaux services sont disponibles (Riedl et al., 2014).

L'objectif du système est de mettre en place un cadre d'interopérabilité grâce à une architecture en quatre niveaux d'un CPS. La connexion entre les systèmes aux niveaux physique et réseaux du CPS assure l'interopérabilité technique du système. Les interfaces logicielles doivent ensuite pouvoir coopérer et échanger des données pour assurer l'interopérabilité sémantique et créer des services dans le nuage informatique. La capacité du système à trouver l'information que l'utilisateur veut dans l'industrie, nécessite de faire la transition entre l'architecture de gauche de la figure 1 vers une architecture de CPS qui autorise les interactions entre l'ensemble des partenaires du système.

L'interopérabilité sémantique peut être construite à partir d'une ontologie ou plus simplement grâce à un diagramme de classe qui modélise les données partagées et leurs relations. Le diagramme de classe permet de décrire le contenu, la source et le but des données appartenant aux différents systèmes utilisés en entreprise et de créer une hiérarchie cohérente. Le format d'échange des données est défini pour qu'elles soient gérées de la même façon sur tous les équipements. L'utilisation des systèmes médiateurs embarqués permet de standardiser le format d'échange sans gérer les appareils individuellement (Riedl et al., 2014). Le diagramme est construit à partir de la cartographie de la première phase et prend en compte les problématiques des différents niveaux d'un CPS. Cette approche reste problématique pour des systèmes complexes avec beaucoup de systèmes hétérogènes (Chungoora et al., 2013).

Finalement, les services sont rendus accessibles aux employés à travers un terminal de contrôle décentralisé. Ce terminal permet la communication entre les systèmes et les employés pour assurer l'interopérabilité organisationnelle.

3.4 Phase 4 : Validation des données et des services

Des tests informatiques sont effectués pour vérifier l'ensemble des échanges de services entre les systèmes et valider la réponse aux besoins. Cette étape conduit soit à la phase trois de développement du CPS si le projet est rejeté (figure 2), soit à la phase deux d'identification des besoins pour définir un nouveau projet.

Le suivi d'indicateurs de performance permet de valider les caractéristiques du CPS à tous les niveaux. Deux indicateurs peuvent être considérés : le nombre de données enregistrées par

unité de temps et la latence du système. Le nombre de données permet de maîtriser l'espace informatique du système et de mettre en place une solution cohérente. La latence correspond au délai entre la demande d'un service et la réponse, elle est critique pour tous les systèmes intégrant des problématiques de réponse en temps réel (Trinks, 2019).

4 CAS D'ETUDE

4.1 Contexte

Le partenaire industriel dans lequel nous avons implanté un CPS est La Milanaise, une meunerie biologique dans la région de Montréal. L'entreprise appartient au secteur agroalimentaire qui impose de nombreuses réglementations sur la production pour garantir la qualité et la sécurité de ses produits (X. Chen et Voigt, 2020). De plus, les grains utilisés pour la production sont cultivés sans l'utilisation de produit chimique, de pesticides et d'engrais synthétiques (MAPAQ, 2019). Les caractères vivant et biologique des grains font que la qualité des matières premières fluctue en fonction de la période de l'année, des conditions climatiques et météorologiques. Une stratégie industrie 4.0 a donc été envisagée pour enrichir les systèmes actuels qui ne supportent pas la diversité des équipements de productions et les besoins informatiques des modèles de prédiction. L'objectif est d'optimiser et d'augmenter la réactivité du pilotage de production dans un environnement de PME hétérogène, fluctuant et rigoureux.

Le partenaire offre un contexte relativement favorable, parmi les PME, pour implanter un CPS d'après les critères de (Masood et Sonntag, 2020) vus en section 2.3. L'entreprise est considérée comme une grande PME avec plus de 50 employés et le processus de transformation est peu complexe avec un flux linéaire et continu (figures 3 et 4). Cependant, la culture technologique est encore peu mature d'après l'étude de Masood. D'une part, l'entreprise a des infrastructures récentes et très automatisées, et la direction encourage le développement technologique et les investissements dans le domaine. D'autre part, les employés sont peu formés aux technologies et aux concepts d'industries 4.0.

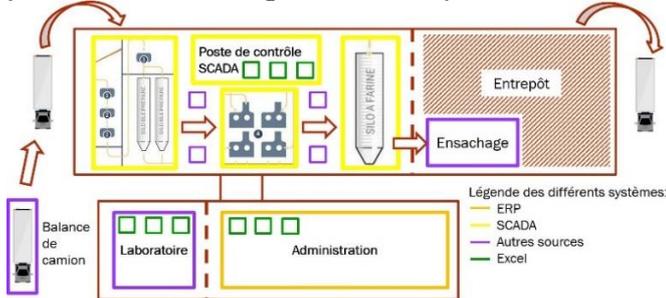


Figure 3 : Schématisation des systèmes utilisés pour le processus de transformation

L'entreprise contient plusieurs sources de données hétérogènes (figure 3) : le système de gestion ERP (en orange), le système de contrôle SCADA (en jaune), les fichiers Excels (en vert) et les autres sources dont les machines isolées (en violet). L'ensemble de ces systèmes forment un environnement complexe avec peu de coopération.

4.2 Mise en œuvre de la démarche

4.2.1 Phase 1 : Cartographie des processus, des données, des systèmes et des services.

Dans cette première phase, une cartographie des processus utilisant le formalisme de l'institut national américain des standards (American National Standards Institute ANSI) a été réalisée pour définir la chaîne de valeur (figure 4).

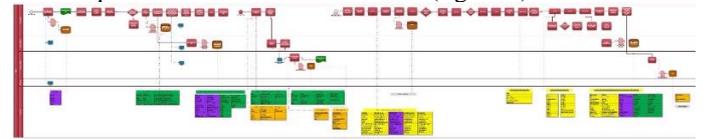


Figure 4 : Cartographie des processus enrichie avec les données manipulées et leur source

Le formalisme ANSI a été enrichi dans notre étude pour répertorier les données manipulées lors du processus de transformation sur le couloir du bas (figure 4) et pour la source des données avec un code couleur identique à la figure 3. Le nombre de données en vert montre que beaucoup de données étaient enregistrées sur des fichiers Excels. Ces données étaient isolées, statiques et dispersées dans l'entreprise.

4.2.2 Phase 2 : Analyse des besoins de données et de services.

La cartographie de la phase 1 nous a permis de comparer les services et les données utilisées dans notre cas d'étude avec les services fournis par l'architecture ISA-95 de la figure 1. Nous avons schématisé l'architecture au début du projet sur la figure 5.

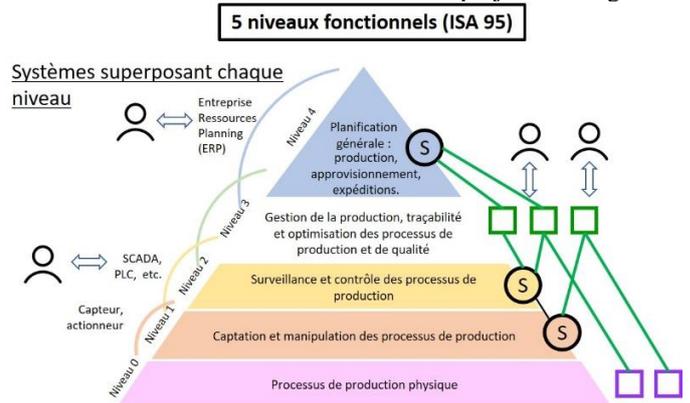


Figure 5 : Architecture des systèmes au début du projet.

On remarque que l'architecture introduite par la norme ISA-95 n'est pas présente dans la PME et que les services et les données d'un MES sont fournis par les fichiers Excel. Les services offerts par ces fichiers sont l'ordonnancement, la collecte des données de production, la gestion des ordres de production et le suivi de la qualité. Ces fichiers représentent les premiers besoins en données et en service à virtualiser, car les services développés dans un CPS nécessitent les données opérationnelles des systèmes d'information standards.

On a observé que beaucoup de données peuvent être virtualisées donc nous avons divisé le développement du système par besoin pour suivre une implémentation agile. Nous avons commencé par mettre en place un CPS sur le début du processus de transformation : le choix des fournisseurs pour les grains de blé. La figure 6 représente les différentes sources et le format des données manipulées pour cette première partie.

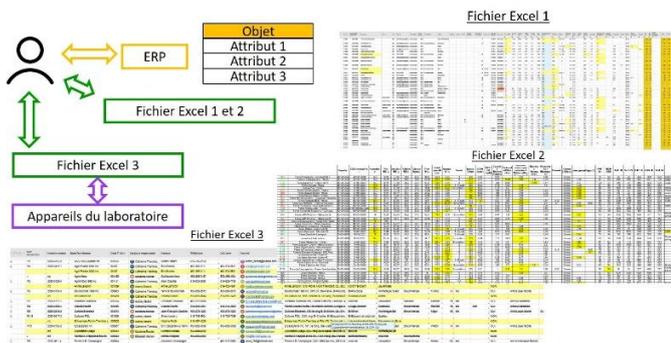


Figure 6 : Sources et format des données utilisées lors du choix des fournisseurs

La première source de données étudiée a été l'ERP de l'entreprise. Ses données cartographiées dans la phase un ont été associées avec les attributs des tables de la base de données relationnelle de l'ERP pour pouvoir les réutiliser dans le CPS. La deuxième source de données comprenait les fichiers Excel retenus et la troisième source regroupait les appareils de mesure du laboratoire. Ces appareils sont des machines spécialisées pour mesurer les caractéristiques d'un grain de blé. Les données et les services fournis par ces deux dernières sources étaient nos besoins à virtualiser.

4.3 Phase 3 : Mise en place d'un système cyber-physique

La mise en place d'un cadre d'interopérabilité pour ce besoin représente la communication entre les données des appareils de mesure du laboratoire, les données des fichiers Excels et des données de l'ERP.

L'interopérabilité technique a nécessité de connecter tous les appareils du laboratoire pour récupérer leurs données d'analyse. Les protocoles de sauvegarde de données ont été étudiés pour connaître le format des données et leur emplacement. Trois formats ont été rencontrés : soit l'appareil avait un format propriétaire nécessitant une interface particulière pour lire les données, soit les données étaient enregistrées dans une base de données relationnelle, ou soit l'appareil fournissait des données analogiques pouvant être analysées par un système médiateur.

L'architecture utilisée pour connecter les systèmes est basée sur une architecture traditionnelle comprenant des programmes informatiques embarqués sur les machines de production pour contrôler le flux de données sur le réseau, gérer les chemins de routage et transférer les données de production. Ces systèmes forment le niveau physique de notre CPS.

Le niveau réseau a été utilisé pour connecter les machines au domaine virtuel. La technologie utilisée pour notre problématique est le réseau Ethernet et le réseau WIFI de l'entreprise.

Le niveau nuage informatique contient la plateforme virtuelle supportant la base de données et les ressources informatiques pour les interfaces. Les données sont manipulées par des interfaces intermédiaires et des interfaces de haut niveau grâce à une architecture client-serveur RESTFUL développée avec Python. La complexité de l'environnement et le manque de ressource dans la PME ont orienté notre choix vers un diagramme de classe au lieu d'une ontologie. Les objets du diagramme ont été créés à

partir des fichiers Excel (figure 7) et des données des appareils de mesure.

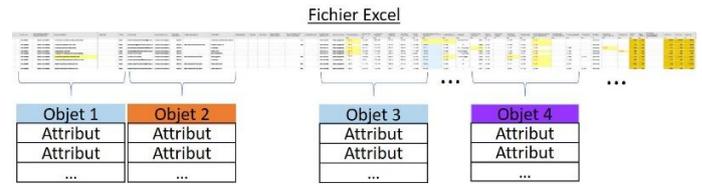


Figure 7 : Processus de création des objets.

La source des objets a été représentée par la couleur des tables : les classes appartenant à la base de données existante de l'ERP en orange, les données nouvellement virtualisées en bleu, et les données des autres sources utilisant des programmes embarqués ou d'autres bases de données en violet. Les relations entre les objets ont ensuite été définies sur le diagramme de classe (figure 8).

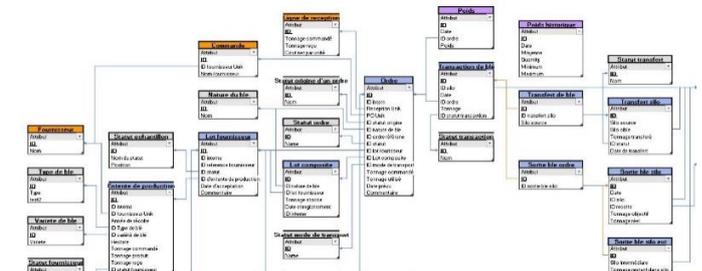


Figure 8 : Présentation d'une partie du diagramme de classe

Ce diagramme utilise le formalisme du langage de modélisation unifié (UML) et affiche seulement le nom utilisateur des objets et de leurs attributs pour conserver la lisibilité du diagramme. Des métadonnées ont ensuite été définies pour chaque attribut : leur nom informatique, leur type, leur modificateur de visibilité, leur capacité à être nul, leur valeur par défaut et un commentaire d'explication. Les métadonnées sur les objets ont été répertoriées dans un espace relié aux classes pour enrichir le formalisme.

Les interfaces gérant les objets de la base de données offrent des services pouvant communiquer avec les autres interfaces. Les communications sont basées sur des événements déclencheurs et peuvent être synchrones ou asynchrones en fonction du choix de design. Les événements permettent de réduire les données échangées en acceptant la communication uniquement sous certaines conditions (Riedl et al., 2014). Les communications synchrones sont principalement utilisées par les interfaces embarquées pour assurer la liaison en temps réel et signaler des événements comme des erreurs.

Le niveau du nuage informatique du CPS centralise les services des différents appareils de mesure et assure leur interopérabilité sémantique. Les analyses peuvent donc être gérées sur l'interface homme-machine de l'application web du niveau terminal de contrôle. L'accès aux résultats sur tous les appareils informatiques connectés aux réseaux de l'entreprise permet une prise de décision décentralisée et en temps réel.

4.4 Phase 4 : Validation des données

Trois boucles de développement ont été effectuées pour s'assurer que les services développés répondaient aux exigences et aux besoins. Les indicateurs de performances ont ensuite été calculés

pour maîtriser la solution. Aucune stratégie d'optimisation du nombre de données enregistrées n'a été développée car les résultats d'analyses sont des valeurs remarquables discrètes. Le délai de récupération des données n'a pas été calculé, car aucune problématique de temps réel n'a été retenue.

4.5 Analyse des résultats

La mise en place d'un CPS pour les choix des fournisseurs a permis à la PME de créer une base de données sur le nuage informatique de l'entreprise. L'information a ainsi pu être décentralisée de façon à avoir l'accès aux mêmes informations en temps réel et sur toutes les plateformes. Les résultats d'analyse des grains ont pu être étudiés pour classer les fournisseurs en fonction de leurs grains. Un seul système sera utilisé sur toute la chaîne de production pour pouvoir réutiliser les données virtualisées dans les prochaines interfaces pour créer des relations entre la qualité de la production et les grains des fournisseurs.

Finalement, l'implantation d'un CPS pour cette première partie a permis à l'entreprise de découvrir les limites et les possibilités de l'industrie 4.0 dans la PME. Les possibilités du système en termes d'optimisation des processus et de qualité ont confirmé la décision de la direction d'investir dans l'industrie 4.0. Le suivi du projet a amélioré la culture technologique et la culture d'industrie 4.0 des employés. Les principales limites sont des limites de cultures technologiques. D'une part, des limites technologies ont été spécifiées pour s'adapter aux connaissances et aux compétences des ressources humaines du projet. D'autres part, la direction a dû mobiliser l'ensemble des employés pour développer la culture technologique de l'entreprise et promouvoir l'implication des employés dans le projet et dans l'utilisation des systèmes développés.

5 CONCLUSION

L'objectif était de mettre en place une architecture de CPS pour assurer l'interopérabilité entre l'ensemble des systèmes d'une PME. Nous avons proposé une méthodologie agile pour adapter une architecture de CPS utilisée dans les grandes entreprises à un environnement de PME. La méthode s'appuie sur 4 phases et a été testée dans une meunerie biologique de la région de Montréal. Le projet est encore récent, mais a déjà apporté des résultats significatifs pour l'entreprise au niveau de leurs choix de fournisseurs et de l'analyse des matières premières. La définition de la méthode et le développement de la culture technologique sur une première problématique nous permettent d'étendre le CPS aux processus de production plus complexes. Une des perspectives est de généraliser l'architecture du CPS à l'ensemble de la chaîne de valeur pour avoir un système de pilotage réactif qui peut s'adapter à la qualité des matières premières. La direction a décidé de ne pas automatiser le CPS et de garder un pouvoir décisionnel humain mais le CPS permettra de responsabiliser les opérateurs de production pour créer une organisation hétéroarchitecturale au sens large, plus réactive localement.

Remerciements

Nous tenons à remercier notre partenaire industriel, La Milanaise, d'avoir participé au développement de ce travail ainsi que le MAPAQ (projet IA119053) pour leur soutien financier.

6 REFERENCES

- Abrahamsson, P. (2002). Agile software development methods: Review and analysis. *VTT Publications*(478), 3-107.
- ANSI/ISA-95. (2013). *Enterprise-Control System Intergration*. In.
- Chen, B., Wan, J., Shu, L., Li, P., Mukherjee, M., & Yin, B. (2018). Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges. *IEEE Access*, 6, 6505-6519. doi:10.1109/access.2017.2783682
- Chen, X., & Voigt, T. (2020). Implementation of the Manufacturing Execution System in the food and beverage industry. *Journal of Food Engineering*, 278, 109932. doi:10.1016/j.jfoodeng.2020.109932
- Chungoora, N., Young, R. I., Gunendran, G., Palmer, C., Usman, Z., Anjum, N. A., . . . Case, K. (2013). A model-driven ontology approach for manufacturing system interoperability and knowledge sharing. *Computers in Industry*, 64(4), 392-401. doi:10.1016/j.compind.2013.01.003
- Colombo, A. W., Karnouskos, S., Kaynak, O., Shi, Y., & Yin, S. (2017). Industrial Cyberphysical Systems: A Backbone of the Fourth Industrial Revolution. *IEEE Industrial Electronics Magazine*, 11(1), 6-16. doi:10.1109/mie.2017.2648857
- Danjou, C., Rivest, L., & Pellerin, R. (2017). *Douze positionnements stratégiques pour l'Industrie 4.0 : entre processus, produit et service, de la surveillance à l'autonomie*. Paper presented at the CIGI 2017 : 12ème Congrès International de Génie Industriel, Compiègne, France, 3-5 mai.
- Drath, R., & Horch, A. (2014). Industrie 4.0: Hit or Hype? [Industry Forum]. *IEEE Industrial Electronics Magazine*, 8(2), 56-58. doi:10.1109/mie.2014.2312079
- EIF. (2004). *European Interoperability Framework, Version 1.0*.
- Geraci A. (1991). *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries*. New York: Inst. Elect. Electron. Eng. Inc
- Hermann, M., Pentek, T., & Otto, B. (2015, 2016). *Design Principles for Industrie 4.0 Scenarios*.
- IEC-62541-1. (2020). International Electrotechnical Commission IEC TR 62541-1. OPC UNIFIED ARCHITECTURE – Part 1: Overview and concepts.
- ISO-2382-1. (1993). Information technology—Vocabulary—Part 1: Fundamental terms, International Organization for Standardization. In. Geneva (Switzerland).
- Jordan, F., Bernardy, A., Stroh, M., Horeis, J., & Stich, V. (2017, 9-13 July). *Requirements-Based Matching Approach to Configure Cyber-Physical Systems for SMEs*. Paper presented at the Portland International Conference on Management of Engineering and Technology (PICMET), Portland, USA.
- Kajati, E., Papcun, P., Liu, C., Zhong, R. Y., Koziorek, J., & Zolotova, I. (2019). Cloud based cyber-physical systems: Network evaluation study. *Advanced Engineering Informatics*, 42, 100988. doi:10.1016/j.aei.2019.100988
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering*, 6(4), 239-242. doi:10.1007/s12599-014-0334-4
- Lu, Y., & Xu, X. (2018). Resource virtualization: A core technology for developing cyber-physical production

- systems. *Journal of Manufacturing Systems*, 47, 128-140. doi:10.1016/j.jmsy.2018.05.003
- Masood, T., & Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, 121, 103261. doi:10.1016/j.compind.2020.103261
- MAPAQ. (2019). *Agriculture biologique*. Tiré de: <https://www.mapaq.gouv.qc.ca/fr/Productions/Production/agriculturebiologique/Pages/alimentsbio.aspx>
- Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2018). A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs). *Journal of Manufacturing Systems*, 49, 194-214. doi:10.1016/j.jmsy.2018.10.005
- Paniagua, C., Eliasson, J., & Delsing, J. (2019, 13-15 Feb). *Interoperability Mismatch Challenges in Heterogeneous SOA-based Systems*. Paper presented at the IEEE International Conference on Industrial Technology (ICIT), Australia.
- Riedl, M., Zipper, H., Meier, M., & Diedrich, C. (2014). Cyber-physical systems alter automation architectures. *Annual Reviews in Control*, 38(1), 123-133. doi:10.1016/j.arcontrol.2014.03.012
- Sinha, D., & Roy, R. (2020). Reviewing Cyber-Physical System as a Part of Smart Factory in Industry 4.0. *IEEE Engineering Management Review*, 48(2), 103-117. doi:10.1109/emr.2020.2992606
- Trentesaux, D. (2007). Les systèmes de pilotage hétéroarchitecturés : innovations réelles ou modèles stériles ?, *Journal Européen des Systèmes Automatisés (JESA)*, 41, 1165-1202
- Trentesaux, D. (2009). Distributed control of production systems. *Engineering Applications of Artificial Intelligence*, 22, 971-978. doi:10.1016/j.engappai.2009.05.001
- Trinks, S. (2019). *Image mining for real time fault detection within the smart factory*. Paper presented at the IEEE 21st Conference on Business Informatics (CBI), Moscow, Russia.
- Valle, P. H. D., Garcés, L., & Nakagawa, E. Y. (2019, Sept). *A Typology of Architectural Strategies for Interoperability*. Paper presented at the Proceedings of the XIII Brazilian Symposium on Software Components, Architectures, and Reuse, Salvador, Brazil.
- Vernadat, F. B. (2010). Technical, semantic and organizational issues of enterprise interoperability and networking. *Annual Reviews in Control*, 34(1), 139-144. doi:10.1016/j.arcontrol.2010.02.009
- Wan, J., Tang, S., Shu, Z., Li, D., Wang, S., Imran, M., & Vasilakos, A. (2016). Software-Defined Industrial Internet of Things in the Context of Industry 4.0. *IEEE Sensors Journal*, 1-1. doi:10.1109/jsen.2016.2565621
- Wang, S., Wan, J., Li, D., & Zhang, C. (2016a). Implementing Smart Factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, 12(1), 3159805. doi:10.1155/2016/3159805
- Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016b). Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Computer Networks*, 101, 158-168. doi:10.1016/j.comnet.2015.12.017

Méthodologie d'étude de la rotation de cultures à partir des données

AMBRE DUPUIS^{1,2}, CAMELIA DADOUCHI^{1,2}, BRUNO AGARD^{1,2}

¹ LABORATOIRE EN INTELLIGENCE DES DONNEES
DEPARTEMENT DE MATHÉMATIQUES ET GENIE INDUSTRIEL,
ÉCOLE POLYTECHNIQUE DE MONTREAL, CP 6079, SUCCURSALE CENTRE-VILLE, MONTREAL, QUEBEC, CANADA
AMBRE.DUPUIS@POLYMTL.CA, CAMELIA.DADOUCHI@POLYMTL.CA, BRUNO.AGARD@POLYMTL.CA

² CENTRE INTERUNIVERSITAIRE DE RECHERCHE SUR LES RESEAUX D'ENTREPRISE, LA LOGISTIQUE ET LE TRANSPORT (CIRRELT)

Résumé -

L'agriculture et la science des données fusionnent peu à peu pour promouvoir une agriculture durable : productive et conservatrice de l'environnement. La rotation de cultures est une pratique agricole aux multiples avantages économiques et environnementaux. Elle est fréquemment étudiée d'un point de vue agronomique cependant, les schémas de rotation de cultures sont encore peu analysés alors qu'ils représentent une composante importante de nombreux modèles agricoles bioéconomiques utiles, entre autres, à l'amélioration de politiques agricoles et au marketing technologique. Le présent article propose une méthodologie d'étude des rotations de cultures à partir des données historiques d'exploitation. Utilisant les principes du Process Mining et les Directly-Follows Graphs (DFG), la méthodologie vise à proposer une représentation et une analyse des successions de cultures, permettant aux décideurs de mieux appréhender les pratiques culturales les plus fréquentes. Les historiques de culture de 394 champs québécois ont été analysés de manière générale puis de façon plus précise en fonction du type d'exploitation et de la zone géographique dont ils sont issus. Notre méthodologie a permis de modéliser les successions directes de cultures les plus fréquentes comme la monoculture de maïs, l'alternance maïs-soja ou encore l'alternance laitue-céleri pour les exploitations maraîchères.

Abstract -

Agriculture and data science are gradually merging to promote sustainable agriculture: productive and environmentally conservative. Crop rotation is an agricultural practice with multiple environmental and economic benefits. It is frequently studied from an agronomic point of view however, crop rotation patterns are still poorly analyzed even though they represent an important component of many bio-economic agricultural models useful, among other things, for improving agricultural policies and technology marketing. This paper proposes a methodology to study crop rotations based on historical farm data. Using the principles of Process Mining and Directly-Follows Graphs (DFG), the methodology aims to propose a representation and an analysis of crops successions, allowing decision-makers to better understand the most frequent cropping practices. The culture history database of 394 Quebec fields was analyzed in a general way and then more precisely according to the type of exploitation and the geographical area from which they originate. Our methodology allowed to model the most frequent direct crop successions such as corn monoculture, corn-soya bean alternation or lettuce-celery alternation for vegetable farms.

Mots clés - Agriculture, process mining, rotation de cultures, DFG

Keywords - Agriculture, process mining, Crop rotation, DFG.

1 INTRODUCTION

Comment nourrir près de 9.8 milliards de personnes d'ici 2050 [United Nations, 2017]. C'est l'un des défis les plus importants que rencontre la société du XXI^e siècle [Pretty et al., 2010]. Une augmentation de 70% de la production mondiale de nourriture entre 2005 et 2050 semble nécessaire pour assurer une sécurité alimentaire mondiale [Kopittke et al., 2019]. Cependant, les pratiques agricoles actuelles ont un impact néfaste sur l'environnement. L'épandage de fertilisants organiques ainsi que l'apport de fertilisants minéraux et de pesticides dans les champs sont en partie responsables de la diminution de la biodiversité dans les écosystèmes par la pollution des eaux, des sols et de l'air. Le développement de nouvelles stratégies agricoles semble donc nécessaire [Ma, 2019]. Parmi elles, la rotation de cultures est un incontournable de l'agriculture durable [Cui et al., 2019], notamment pour une meilleure utilisation de l'épandage.

Dans le présent article, nous proposons une méthodologie qui vise à identifier les rotations de cultures les plus fréquemment utilisées à partir de l'analyse de données historiques. Cette modélisation pourra, dans un second temps, intégrer des modèles agricoles bioéconomiques plus complexes utiles, entre autres, à l'amélioration des politiques agricoles ou de marketing technologique. Cette méthodologie s'appuie sur les techniques de Process Mining et les historiques de cultures d'exploitations.

L'article se décompose comme suit : la section 2 présente l'état de l'art, dans lequel nous adresserons les fondamentaux agronomiques relatifs à la rotation de cultures (2.1), les recherches actuelles relatives à la modélisation des rotations de cultures (2.2), ainsi que les concepts préliminaires relatifs au Process Mining (2.3). Cette section sera conclue par une synthèse des grands principes proposés dans l'état de l'art (2.4). La section 3 présentera la méthodologie proposée afin de représenter les

rotations de cultures les plus fréquentes. La section 4 exposera un cas d'étude qui a servi à la validation de la méthode. Dans un premier temps, nous présenterons le contexte de l'étude de cas (4.1), dans un second temps l'analyse selon le type d'exploitation sera effectuée (4.2) et, dans un troisième temps, une analyse par zone géographique est réalisée (4.3). Finalement, la section 5 rappelle les résultats et limites de la méthode proposée et présente quelques perspectives.

2 ÉTAT DE L'ART

L'agriculture et la science des données fusionnent pour proposer une agriculture avec le minimum d'impact sur l'environnement tout en maximisant sa productivité [Liakos et al., 2018]. La pratique de la rotation de culture a elle aussi été fréquemment étudiée.

2.1 Rotation de cultures : fondamentaux agronomiques

La rotation de cultures est une pratique agricole consistant à répéter systématiquement, ou du moins fréquemment, une même séquence de culture durant plusieurs saisons de croissances [Reeves, 1994]. Le type de cultures utilisé dépend à la fois des facteurs écologiques (caractéristique physique et chimique des sols, interaction du sol avec l'écosystème, le climat, etc.) et des facteurs économiques [Porter, 2009]. Les avantages de cette pratique sont principalement la lutte contre l'érosion des sols, la conservation de la matière organique, la gestion des nutriments ainsi que le contrôle des maladies [USDA, s.d.]. Ces avantages peuvent être traduits en bénéfices pour les agriculteurs ; la rotation de cultures permet de « diminuer les coûts associés à la fertilisation des champs, réduire les besoins en pesticides et herbicides, améliorer le rendement des champs en valorisant la qualité des sols, prévenir l'érosion et conserver l'humidité des sols, protéger la qualité des eaux et aider à protéger la santé humaine. » [Clark, 2007]. L'agriculture conservatrice favorise ainsi la pratique de la rotation de cultures pour une agriculture durable et productive [Cui et al., 2019].

2.2 Modéliser les rotations de culture

La modélisation des processus permet de les décrire afin de détecter les situations problématiques. La reconception pour l'amélioration des processus modélisés devient alors possible [Biazzo, 2002]. La rotation de culture peut être vue comme un processus à améliorer. Cependant, alors qu'un grand nombre d'études tentent, par différents moyens, de prédire les rotations de cultures [Osman et al., 2015][Klein Haneveld et Stegeman, 2005][Detlefsen et Jensen, 2007][Salmon-Monviola et al., 2012][Le Ber et al., 2006][Zhang et al., 2019][Yaramasu et al., 2020][Kussul et al., 2017], très peu, à notre connaissance, cherchent à les modéliser à partir de données réelles. Cette observation est d'ailleurs confirmée par [Levavasseur et al., 2015] « peu de modèles proposent de reconstruire des séquences de cultures réelles et de simplifier leur diversité dans un nombre raisonnable de rotations de cultures. ». Pourtant, dès 1994, l'utilité de la modélisation des exploitations agricoles est reconnue par les scientifiques [Edward-Jones & McGregor, 1994] à des fins de marketing technologique ou d'évaluation de politique *ex ante*.

La connaissance des rotations de cultures et des interrelations entre les cultures est un des éléments les plus souvent utilisés dans les modèles agricoles bioéconomiques mécanistes. 27 des 48 études revues par [Janssen & van Ittersum, 2007] utilisent des contraintes liées aux rotations de cultures. Pour répondre à cette

problématique de modélisation, [Dogliotti et al., 2003] et [Bachinger et Zander, 2007] proposent des outils basé sur un ensemble de règles agronomiques alors que [Castellazzi et al., 2008] proposent une représentation mathématique des rotations de culture, soit des savoirs théoriques. [Levavasseur et al., 2015] proposent une caractérisation des rotations de cultures à l'aide de données réelles obtenues via des entretiens auprès d'exploitants locaux sur leurs habitudes culturelles ainsi qu'une modélisation dérivée des travaux de [Schönhart et al., 2011]. Ces derniers travaux ont permis de développer l'outil CropRota utilisant à la fois les règles agronomiques et des données réelles relatives aux historiques de cultures afin de faire ressortir les séquences de cultures les plus fréquentes [Schönhart et al., 2011].

2.3 Concepts préliminaires relatifs au Process Mining

Le Process Mining est à la croisée des chemins entre la fouille de données et la modélisation des processus d'affaires [van der Aalst, 2011]. L'objectif général est de trouver, à partir des données, des schémas de comportement récurrent [Cook et Wolf, 1995][van der Aalst, 2011]. Le Process Mining permet alors de représenter les principaux flux d'activités au sein d'un processus étudié [Nieves et al., 2020]. Le Process Mining se base sur trois composantes essentielles qui doivent impérativement se retrouver dans les données traitées soit :

- Une activité : événement ou étape bien définis dans le processus.
- Un cas : réalisation unique du processus.
- Un marqueur temporel : indication temporelle de la réalisation de l'activité dans un cas précis.

On appelle *Trace* la succession des activités dans un cas [van der Aalst, 2011].

Ce genre de données peut être utilisé dans 3 types de Process Mining [van der Aalst, 2016] :

- La découverte : permet de générer un modèle à l'aide des données sans informations a priori
- La conformité : permet la comparaison entre un modèle théorique et la réalité contenue dans les données
- L'amélioration : permet de modifier le modèle théorique en fonction de la réalité contenu dans les données

Le résultat obtenu est une représentation graphique (PetriNet, EPC...) de toutes les options possibles permettant de réaliser le processus en question. Il existe, entre autres, 3 grands algorithmes de process mining de type découverte. L'algorithme α , le plus simple, mais limité par ses capacités de détection de boucles courtes (moins de trois activités), d'activités en parallèle et est sensible au bruit. L'*heuristic-miner* prend en compte la fréquence des traces, ce qui permet de limiter l'impact du bruit et la détection des boucles courtes. Finalement, l'algorithme *inductive-miner*, le plus utilisé, permet de détecter les comportements peu fréquents et assure la cohérence du modèle [The Fraunhofer Institute for Applied Information Technology, 2020]. Cela dit, l'ensemble de ces algorithmes est utile lorsque les activités de début et de fin du processus sont identifiables.

Un outil permettant de représenter les activités et leurs relations directes, sans prendre en compte leurs places au sein du processus, est le *Directly Follow Graph* ou *DFG*. Ce graphe a pour nœud les différentes activités présentes dans la base de données et assigne un arc entre deux activités successives. Ainsi le DFG est un graphe orienté dont le poids, par exemple $w_{a,b}$, des arcs représente le nombre de fois dans lequel la succession des activités A puis B a

été enregistrée [Leemans et al., 2018]. Dans le cadre de la modélisation des relations directes entre les cultures, l'utilisation du DFG semble la plus appropriée.

2.4 Synthèse

L'état de l'art montre l'importance de la rotation de cultures pour préserver la qualité des sols et assurer la réduction de l'impact environnemental de l'agriculture sur le sol. De nombreuses méthodes, basées sur le savoir théorique et sur l'acquisition de savoir à partir des données, ont été élaborées afin de prédire les rotations de cultures utilisées par les agriculteurs. Cependant, très peu permettent de modéliser simplement les relations entre cultures, pourtant utiles à l'élaboration d'un modèle agricole bioéconomique plus complexe permettant une amélioration des politiques agricoles et du marketing technologique. À notre connaissance, aucune méthode n'utilise le Process Mining pour identifier les modèles de rotations de cultures. Pourtant, le Process Mining a pour objectif de trouver des schémas de comportements récurrents à partir de données événementielles. En assimilant la rotation de cultures à un processus, le type de culture planté pourrait être associé à un événement dont la date d'exploitation correspondrait au marqueur temporel nécessaire à l'utilisation des méthodes de Process Mining. Cette technique pourrait permettre de mieux appréhender les habitudes culturelles menant à l'exploitation d'un type de culture.

Pour ces raisons, nous proposons une méthodologie d'étude de la rotation de cultures. À partir de l'analyse des données historiques d'exploitation agricole, les successions de cultures sont mises en évidence, puis une analyse des différents graphes DFG permet de mieux comprendre les pratiques de gestion des agriculteurs en matière de rotation de cultures.

3 METHODOLOGIE D'ETUDE DES ROTATIONS DE CULTURES

Cette méthodologie peut être décomposée en 5 étapes détaillées ci-après.

Afin d'illustrer la méthodologie, nous considérons un ensemble de données qui décrit la réalisation successive de différents événements A, B, C, D et E. Par exemple :

Cas 1	A → A → A → A → B → B → A → D → B
Cas 2	A → A → A → A
Cas 3	B → B → D → A → B → B → A → A
...	

Le nombre d'événements peut varier d'un cas à l'autre, il peut y avoir des répétitions d'événements, tous les événements ne sont pas obligatoirement présents dans chaque cas.

3.1 Exploration statistique des données

Il est essentiel de comprendre la nature et la distribution des données pour pouvoir, par la suite, les préparer adéquatement. Cette étape permet également « de détecter des sous-ensembles intéressants pour formuler des hypothèses sur les informations cachées » [Pretorius et Matthee, 2006]. Pour connaître la nature des données, une visualisation des attributs disponibles et de leurs types peut être effectuée. Aussi, une visualisation des premiers enregistrements permet d'affiner la compréhension des informations présentes dans la base de données. Ainsi, les attributs relatifs aux activités, aux cas et aux marqueurs temporels (section 2.3) peuvent être identifiés. Par la suite, l'étude de la distribution des différentes activités dans la base de données peut être effectuée à l'aide d'un décompte du nombre d'occurrences.

3.2 Préparation des données

L'exploration statistique des données permet d'orienter la préparation des données. En effet, si le nombre d'activités uniques est trop important, elles peuvent être groupées en catégories plus générales. L'ensemble des données doit être contenu dans un tableau possédant au minimum 3 colonnes contenant les informations présentées dans la section 2.3. Ainsi, une première colonne contient l'identification unique du cas, une seconde colonne doit contenir l'information sur l'activité, une troisième colonne doit contenir la date d'occurrence de la combinaison des deux colonnes précédentes. L'exploration des données (section 3.1) a normalement permis d'identifier les attributs contenant les informations d'intérêt, mais il est possible que certaines transformations soient nécessaires afin d'atteindre le formatage exigé. La création d'un champ permettant l'identification unique des cas par l'association de deux attributs existants dans la base de données peut être nécessaire. Aussi, les données utilisées pour le marqueur temporel peuvent devoir subir une conversion au format « Date ». Finalement, si comme mentionné dans la section 3.1, des sous-ensembles d'intérêts sont identifiés, ils doivent être isolés par une sélection sur l'attribut d'intérêt. Ainsi le tableau 1 représente le 1^{er} cas de l'exemple, après l'étape de préparation des données.

Tableau 1: Exemple de représentation des données

case:concept:name	concept:name	time:timestamp
1	A	1
1	A	2
1	A	3
1	A	4
1	B	5
1	B	6
1	A	7
1	D	8
1	B	9

L'ensemble des cas sont transformés de la sorte et mis bout à bout.

3.3 Création des graphes (DFG)

Le graphe nommé le « Directly-follows graph » ou DFG (section 2.3) est utilisé pour sa facilité de représentation des relations directes existant entre les activités et les possibilités de transformation postérieure qu'il propose. La réalisation du graphe DFG sur les données brutes est proposée à la figure 1.

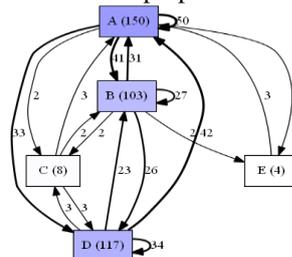


Figure 1: Graphe Directly-follows (DFG) brute

Ce premier graphe permet de représenter l'ensemble des activités ainsi que leurs interactions directes. On note que les activités A, B et D sont les plus fortement représentées dans la base de données étudiée avec respectivement 150, 103 et 117 alors que les événements C et E sont très marginaux.

Ce graphe permet de visualiser les différentes options possibles à la suite d'une activité. Avec les informations présentes sur la figure, les probabilités conditionnelles de successions de culture peuvent être déduites. Ici, si l'activité A survient, alors l'activité suivante sera l'activité A avec une probabilité de 39%, l'activité D avec une probabilité de 32%, l'activité B avec une probabilité de 26% ou les activités C ou E avec une probabilité de 1.5%.

3.4 Simplification du modèle

En général, la création du premier graphe à partir des données traitées est peu lisible, car l'ensemble des transitions sont représentées, et ce sans réelle distinction. Beaucoup d'informations s'enchevêtrent, aboutissant à un graphe surchargé et difficilement compréhensible. Pour pallier ce manque de visibilité, deux traitements sont proposés :

3.4.1 Regroupement des activités

L'illisibilité du graphe issu des données brutes nous pousse à traiter les données afin de limiter la complexité de leurs représentations. Le regroupement de culture consiste à sélectionner l'ensemble des cultures dont le nombre d'occurrences total dans la base de données étudiée est inférieur à un seuil et de les regrouper sous une appellation commune (e.g. « Autre »). Les événements dont le taux d'occurrence est inférieur à un seuil (ici fixé à 10%) sont regroupés. Ici les événements C et E représentent respectivement 2% et 1% des événements enregistrés. Ils sont donc regroupés dans une nouvelle catégorie d'événements appelée « Autre ». Cette nouvelle catégorie représente 3% des événements (12 enregistrements). Un nouveau graphe est réalisé avec ce regroupement (fig 2.).

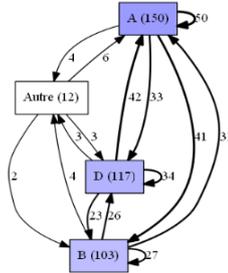


Figure 2 : DFG regroupé (seuil 10%)

Ce regroupement permet d'alléger le modèle en supprimant une case d'activité sans pour autant changer les différentes proportions des transitions entre les événements principaux.

On retrouve les mêmes probabilités de succession à l'activité A que précédemment, à l'exception des probabilités $P(A \rightarrow E)$ et $P(A \rightarrow C)$ qui sont transformés en $P(A \rightarrow \text{Autre}) = 3\%$.

3.4.2 Filtrage des arcs

Toujours dans l'optique d'améliorer la lisibilité des graphes DFG pour identifier les rotations de cultures les plus fréquentes, nous proposons d'effectuer une opération de filtrage. Le filtrage des arcs consiste à éliminer l'ensemble des arcs dont le taux de transition est inférieur ou égal à un seuil fixé.

Le taux de transition associé à l'arc i (Tt_i) est défini comme suit :

$$Tt_i = \frac{w_i}{\sum_{i=1}^n w_i}$$

Où : n = nombre d'arcs dans le graphe
 w_i = nombre de transitions associées à l'arc $\forall i \in n$

Puisqu'une transition relie toujours deux activités, on peut noter :

$$\sum_{i=1}^n w_i = nb \text{ activités} - nb \text{ entités}$$

D'où

$$Tt_i = \frac{w_i}{nb \text{ activités} - nb \text{ entités}}$$

Le nombre total d'activités est ici de 382 et le nombre total de cas enregistrés est de 53. Ainsi, si l'on fixe le seuil limite à 1%, seuls les arcs comptant plus de 4 transitions sont représentés (fig 3). Cette opération est appelée « filtrage ».

L'opération de filtrage permet de grandement simplifier le modèle puisque seules les activités les plus fréquentes sont représentées. Cependant, cette opération entraîne une légère modification dans les probabilités de successions puisqu'après filtrage : $P(A \rightarrow A) = 40\%$; $P(A \rightarrow B) = 33\%$ et $P(A \rightarrow D) = 27\%$. Cela étant dit, l'ordre est conservé.

Les opérations de regroupement et de filtrage peuvent être combinées comme le présente la figure 4. (section 3.4):

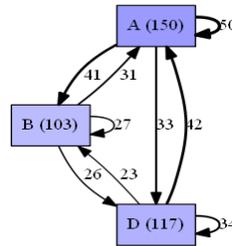


Figure 3 : DFG filtré (seuil 1%)

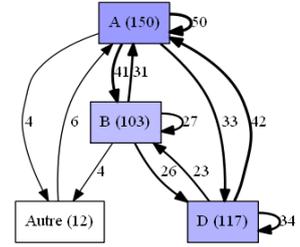


Figure 4 : DFG regroupé et filtré (seuil 10% et 1%)

La combinaison des opérations de regroupement et de filtrage permet de conserver les avantages de l'opération de filtration tout en minimisant son désavantage premier. En effet, ce double traitement permet d'une part de ne conserver que les transitions les plus fréquentes (éliminant de ce fait les activités fortement minoritaires), rendant ainsi le graphe beaucoup plus lisible. Et d'autre part, de limiter l'impact de la filtration sur les probabilités de succession. Dans le cas du regroupement puis du filtrage, les probabilités de succession à l'activité A sont : $P(A \rightarrow A) = 39\%$; $P(A \rightarrow B) = 32\%$; $P(A \rightarrow D) = 26\%$ et $P(A \rightarrow \text{Autre}) = 3\%$.

3.4.3 Choix des seuils

Les opérations de regroupement et de filtrage permettent de donner plus de lisibilité au graphe DFG en simplifiant sa structure tout en conservant le même ordre dans les probabilités de succession des activités. Ces opérations de traitement sont effectuées en s'appuyant sur des seuils fixés arbitrairement. Une analyse des effets d'une variation sur ces seuils permet de dégager les résultats suivants : un seuil de regroupement tenant compte de la contribution des activités à la base de données permet de produire un graphe identique ou très similaire au graphe brut. A contrario, un seuil de regroupement trop élevé engendre une perte d'information trop importante pour une bonne analyse du graphe DFG. Nous proposons de choisir le seuil de regroupement de manière visuelle à l'aide du graphique représentant la distribution des activités dans la base de données. Le seuil devrait être choisi au niveau d'une « marche » afin de séparer les activités principales des activités fortement minoritaires.

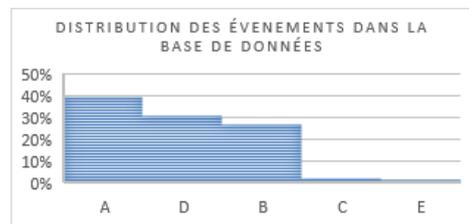


Figure 5 : Distribution des événements

Ainsi, d'après la figure 5, le seuil de regroupement devrait appartenir à l'intervalle [5% ; 25%] pour garder un niveau d'information suffisant tout en simplifiant le modèle. Tout comme

pour le seuil de regroupement, le seuil de filtrage peut être déterminé par l'analyse du graphique de distribution des transitions (fig 6). Le seuil de filtrage devrait appartenir à l'intervalle [1% ; 6%] pour garder un niveau d'information suffisant tout en simplifiant le modèle.

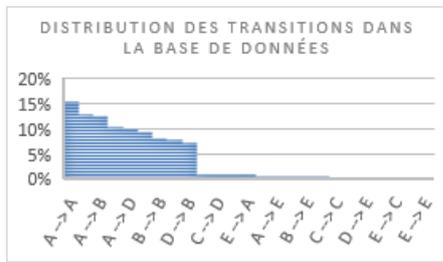


Figure 6 : Distribution des transitions

3.5 Analyse des résultats

L'application de la méthodologie permet de dégager les trois activités les plus fréquentes (A, B et D) ainsi que leurs relations directes. Suite à la réalisation des activités A, B ou D, on retrouve généralement l'activité A. On note la forte tendance à la répétition des événements fréquents. Les événements C et E sont marginaux et sont regroupés dans la catégorie « Autre ». Ce regroupement montre que, généralement, la succession à une activité « Autre » est l'activité A. Comme précisée dans les sections 3.1 et 3.2, la connaissance du domaine d'application peut être utilisée pour filtrer selon certains critères et obtenir des graphes ciblés. En effet, certains attributs ne faisant pas partie des informations essentielles au Process Mining (section 2.3) peuvent cependant avoir un grand intérêt pour l'analyse puisqu'ils permettent d'ajouter de l'information sur le contexte dans lequel les données évoluent. Ceci peut avoir un impact important sur la succession des activités au sein d'un processus. Le filtrage par attribut dans la phase de préparation des données permet d'ajouter de l'information à notre analyse et de proposer des graphes ciblant des comportements plus spécifiques. L'analyse des graphes permet d'une part de repérer les activités les plus fréquentes au sein d'un processus en limitant la représentation d'activités fortement minoritaires pouvant être perçues comme du bruit et, d'autre part, en analysant des arcs reliant les activités de quantifier les relations directes qu'elles entretiennent entre elles. L'analyse de ces graphes permet d'avoir une représentation lisible des successions directes d'activité au sein d'un processus et de les quantifier.

4 CAS D'ETUDE

En 2018, la province du Québec au Canada comptait 28 919 exploitations agricoles et 1 867 000 hectares cultivés [Institut de la statistique Québec, 2020]. Les zones agricoles sont concentrées sur les berges du Saint-Laurent au sud de la province [Gouvernement Québec, 2019]. Les données dont nous disposons décrivent les pratiques agricoles de 81 exploitations québécoises réparties sur ces territoires agricoles. Cinq types d'exploitations peuvent être définis soit les élevages de ruminants (bovins, ovins), les élevages de monogastrique (porcs, volailles), les exploitations de grandes cultures, les exploitations maraîchères ainsi que les exploitations de gazon. Ces données retracent les différents types de cultures exploitées dans 394 champs entre 2000 et 2020 au Québec. Les données des 20 dernières années ne sont pas

disponibles pour l'ensemble des champs et les données concernant certains champs pour certaines années d'exploitation sont manquantes. 81 types de cultures (exemple : foin de graminées, foin de graminée 100%, foin de graminées 60%, laitue pommée terre noire, laitue frisée transplantée en terre noire ...) sont répertoriées.

4.1 Exploration statistique des données

Comme présenté précédemment, 81 types de cultures sont répertoriées cela dit, une grande partie de ces cultures font référence à une catégorie d'espèces particulières. L'exploration des données permet également d'identifier deux attributs pouvant servir de filtre pour définir des sous-ensembles d'analyse. Ces attributs sont les types d'exploitation (voir section 4.3.1) et la situation géographique des champs (voir section 4.3.2). Nous identifions les cultures comme les activités possibles tandis que les champs représentent les différents cas étudiés dans le processus de rotation de cultures. L'année d'exploitation d'une catégorie de culture dans un champ représente le marqueur temporel nécessaire à l'utilisation des techniques de Process Mining (voir section 2.3). L'analyse des données permet d'identifier la première opération de traitement des données à réaliser, soit la catégorisation des cultures selon leurs espèces. En effet, sur 81 types de cultures identifiées, 66 représentent moins de 1% des enregistrements de la base de données. Finalement, certains champs comptent plusieurs enregistrements pour une même année, mais le type de culture exploitée reste le même pour ces enregistrements.

4.2 Préparation des données

Dans un premier temps, l'analyse se porte sur l'ensemble des champs référencés dans la base de données. La première étape de préparation consiste à regrouper les cultures selon leur « espèce ». Ainsi, les cultures de « foin graminées », « foins graminées 100% », « foin graminées 60% » ... sont regroupées sous la même appellation « Foin ». Il en est de même pour les autres cultures. Cette opération permet de réduire le nombre d'activités considérées, passant de 81 types de cultures à 36 cultures. Dans un second temps, une clé d'identification unique des champs est créée par une association du numéro de l'exploitation et du numéro du champ. Cette clé d'identification unique à chaque champ est celle utilisée pour l'identification des cas dans l'algorithme de Process Mining. De plus, les enregistrements dupliqués (plusieurs enregistrements d'une même culture dans un même champ pour une même année) sont regroupés sous un même enregistrement soit 6 552 enregistrements. Finalement, seuls les attributs d'intérêt pour l'étude (clé d'identification, type de culture et date d'exploitation) sont sélectionnés et renommés respectivement « case:concept:name », « concept:name », « time:timestamp » pour la réalisation des étapes suivantes.

4.3 Création des graphes

Le premier graphe DFG, crée via la librairie PM4PY disponible sous Python, reprenant l'ensemble des enregistrements, est présenté à la figure 7. Le grand nombre d'activités et d'interrelations le rend relativement peu compréhensible. L'enchevêtrement des arcs demande un certain effort pour pouvoir comprendre les dynamiques entre les différents types de cultures. Cette complexité est appréhendée par les différentes étapes de traitement proposées ci-après.

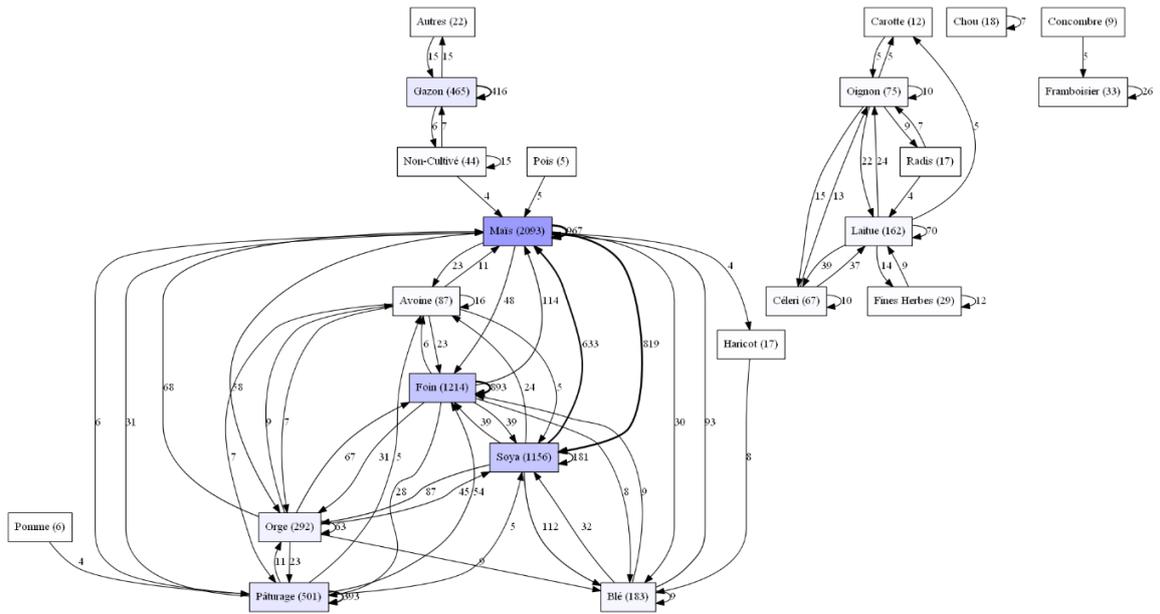


Figure 7 : DFG brut pour les champs étudiés

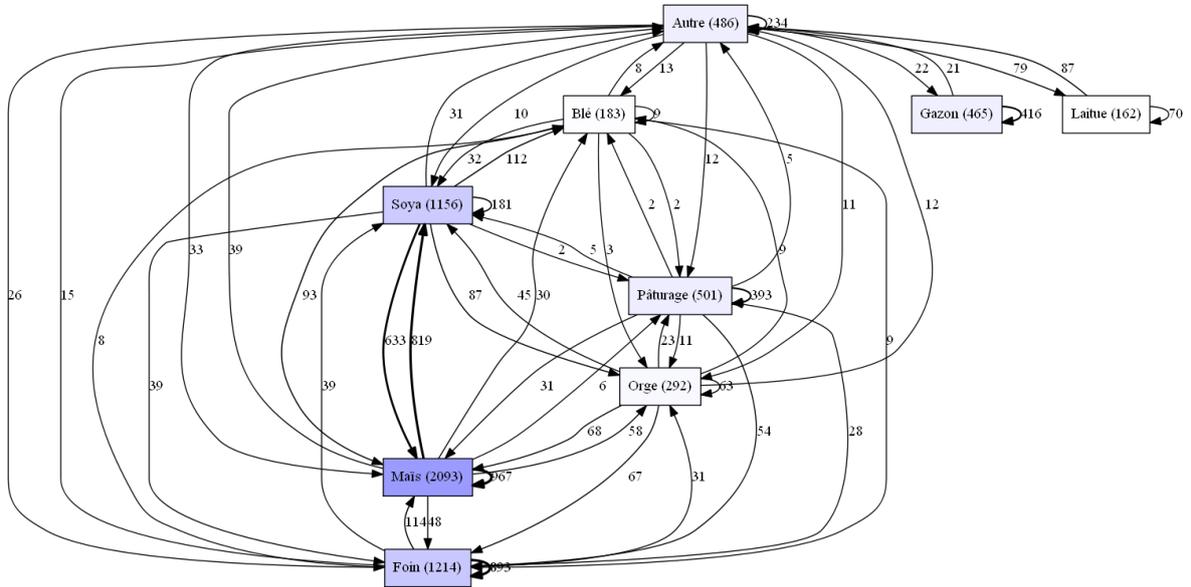


Figure 8 : DFG regroupé (seuil 2%) pour les champs étudiés

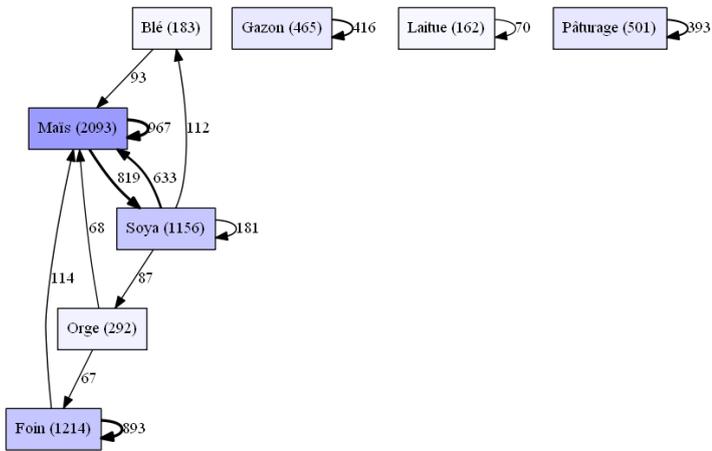


Figure 9 : DFG filtré (seuil 1%) pour les champs étudiés

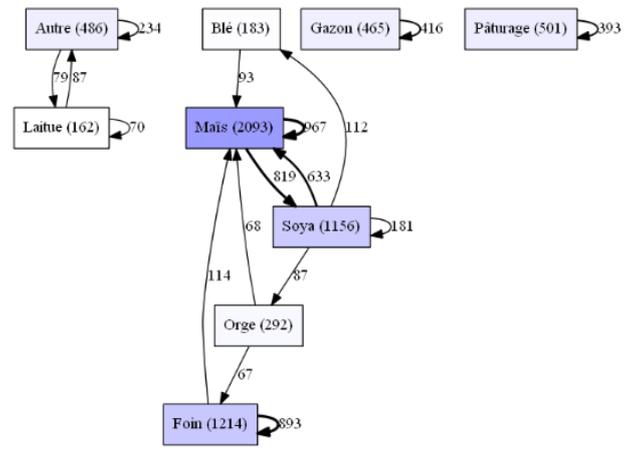


Figure 10 : DFG groupé et filtré (seuils 2% et 1%) pour les champs étudiés

4.3.1 Regroupement des cultures

Afin de regrouper les cultures, le seuil de regroupement a été choisi à la suite de l'analyse du graphique des distributions des activités. Un seuil de 2% a été choisi, soit le regroupement de toutes les cultures comptabilisant moins de 132 occurrences dans la base de données. Le graphe résultant est proposé figure 8.

4.3.2 Filtrage des arcs

Afin de filtrer les arcs, le graphique des distributions des arcs a été réalisé puis analysé. Le choix d'un seuil de 1% a été fait de sorte que seuls les arcs dont le nombre d'occurrences est supérieur à 65 sont visualisés dans la figure 9. Le filtrage permet de simplifier grandement le modèle en ne modélisant que les successions fréquentes. La forte tendance à la monoculture du maïs, du foin ainsi que l'alternance du maïs-soja sont très facilement visibles. On note la diminution du nombre d'activités, particulièrement les cultures de légumes. L'opération de groupement puis de filtrage permet de mettre en valeur un besoin d'investigation des pratiques maraîchères (Fig10).

4.4 Analyse des résultats

On note la forte tendance à l'alternance soja-maïs, ainsi que la forte tendance à la monoculture des cultures de maïs, foin, gazon et pâturage. On peut également noter la faible tendance à la monoculture du soja et du blé. On semble pouvoir observer deux boucles de trois éléments soit maïs-soja-blé et maïs-soja-orge. Cela dit, le graphe DFG ne propose qu'une vision des flux directe donc il est impossible de savoir combien de ces boucles ont réellement été observées. Il semble se dégager un comportement particulier relatif à l'exploitation maraîchère avec une rotation laitue → « Autre » relativement importante. Dans ce cas précis, l'apport des informations relatives au type d'exploitation peut être utile puisque la même méthodologie peut être utilisée sur un ensemble de données ciblé afin de dégager un comportement particulier. C'est pourquoi nous avons réalisé cette analyse par secteur agricole qui est présentée ci-après.

4.4.1 Analyse par secteur agricole

Comme mentionné dans la section 4, il existe 5 types d'exploitations distinctes. L'analyse des rotations de culture générale a dégagé le besoin de cibler les comportements relatifs à l'exploitation de légumes. C'est pourquoi nous utilisons l'information relative au type d'exploitation pour approfondir nos analyses. La base de données précédemment décrite est filtrée pour ne garder que les 30 champs issus d'exploitations de type « Maraîcher ». Seuls les 3 attributs d'intérêts sont conservés. Les colonnes regroupant respectivement les identifiants des champs, l'année d'exploitation et la culture exploitée sont renommées par « case:concept:name », « time:timestamp » et « concept:name ».

Le graphe « brute » est réalisé de manière à visualiser l'ensemble des transitions de culture ayant existé entre 2000 et 2020 dans les 30 champs étudiés. Dans ces 30 champs, 23 types de cultures sont représentés dans les 464 occurrences de cultures, d'où un nombre total de 434 transactions.

Afin de faire ressortir les comportements les plus fréquents, un regroupement des cultures représentant moins de 1% des enregistrements est effectué. Ainsi, les cultures présentes moins de 4 fois dans l'ensemble des données sont regroupées dans la catégorie « Autre ». Dans un second temps, les transitions représentant moins de 1% des transitions totales, soit inférieures à 5 occurrences, sont filtrées (Fig 11).

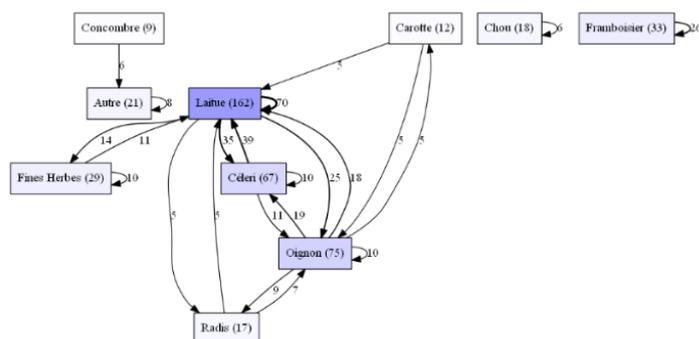


Figure 11 : DFG groupé et filtré (seuils 1%) dans « Exploitations maraîchères »

L'ensemble de ces traitements permet de mettre en évidence, par exemple, la tendance à la monoculture de framboise ce qui paraît cohérent à la vue du caractère pluriannuel de cette culture [CRAAQ, 2003]. L'alternance fréquente des cultures de laitue et de céleri peut également être remarquée (23% des transitions issues de la culture de laitue sont à destination de la culture de céleri et 65% des transitions issues de la culture de céleri sont à destination de la culture de laitue). Finalement, les recommandations proposées par [Grignon, 2015] peuvent, en partie, se retrouver dans la figure 11 avec une forte relation de la culture d'oignon à la suite d'une culture de laitue ou de céleri, la relation existante entre la culture de carotte et d'oignon, ou encore l'absence de monoculture de carotte. Cependant on note une divergence entre les pratiques réelles observées à la figure 11 et les recommandations agronomiques puisque la monoculture d'oignon peut être observée dans 16.39% des cas ce qui est déconseillé d'un point de vue agronomique (recommandation d'un minimum de 3ans entre deux cultures d'oignon) [Grignon,2015]. Le même type d'analyse peut être réalisé par rapport à la localisation spatiale des champs.

4.4.2 Analyse par zone géographique

La pédologie des sols, le relief, le climat influencent grandement les pratiques agricoles. Les écodistricts sont des délimitations géographiques du territoire canadien « caractérisées par des assemblages distinctifs de formes du relief, de matériaux de surface, de sols, de plans d'eau, de végétation et d'utilisations du territoire » [AAC, 1995]. C'est le plus faible niveau de généralisation de la classification écologique des terres [Statistique Canada, 2018]. C'est pourquoi il a été choisi comme indicateur spatial associé aux champs.

Les données disponibles sont réparties dans 6 écodistricts distincts dont les limites géolocalisées sont disponibles sur les données ouvertes du Canada [Gouvernement Canada, 2017].

À l'aide du système d'information géospatial QGIS, les différents champs géolocalisés ont été associés à écodistrict auxquels ils appartiennent. L'analyse des rotations de culture en fonction de la zone géographique permet de cibler de manière plus précise les comportements des agriculteurs, mais cette fois sans l'intervention de connaissances extérieures. L'étude des 242 champs présents dans l'écodistrict « Middle St.Lawrence plaine » est entreprise. Comme dans l'analyse par secteur agricole, seuls les 3 attributs d'intérêts sont conservés et renommés. 17 types de cultures sont présents dans ce sous-ensemble de données comptant 4 107 enregistrements (occurrences de cultures) et 3 865 transitions. Le graphe « brut » est présenté dans la figure 14.

Le regroupement et le filtrage des données au seuil de 1% est là aussi effectué. Ainsi, les cultures représentant moins de 42

enregistrements sont regroupées est les transitions inférieures à 39 occurrences sont filtrées (Fig 12).

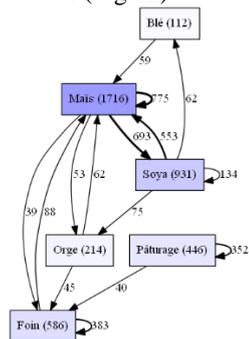


Figure 12 : DFG groupé et filtré (seuils 1%) dans "Middle St. Lawrence Plain"

On note la très grande tendance à la monoculture de maïs (50%) de pâturage (90%) et de foin (81%). La monoculture de foin est en adéquation avec les recommandations agronomiques préconisant une monoculture de foin de 4 à 6 ans [Turcotte, 2018]. La succession soja→maïs est très fortement représentée (67%) tout comme la transition maïs → soja (44%). On note également que le soja a une très faible tendance à la monoculture avec seulement 16% des successions au soja vers le soja. Finalement, deux cycles de trois cultures sont facilement identifiables soit : maïs→soja→blé ainsi que maïs→soja→orge. D'après [Kirkegaard et Ryan, 2014][Kirkegaard et al., 2008] l'insertion d'une culture de légumineuse à grain ou d'oléagineux (ici le soja qui est une oléo-protéagineuse) dans une rotation céréalière (maïs, orge et blé) permet d'augmenter les rendements des cultures céréalières grâce, entre autres, à une amélioration de l'apport en azote et à une diminution des mauvaises herbes et maladies dans les cultures. Les bénéfices de l'apport de la culture de blé dans la rotation maïs→soja ont également été démontré au Canada par [Janovicek et al., 2021]. L'intégration de légumineuse dans une rotation céréalière a également l'avantage de diversifier les revenus des exploitants [Kirkegaard et Ryan, 2014]. Ceci pose la question de l'intérêt de la monoculture de maïs pourtant largement représenté.

5 CONCLUSION

La méthodologie proposée a permis de représenter les successions de cultures les plus fréquentes, après un découpage par type d'exploitation et par zone géographique. Cette méthodologie s'appuie sur 5 phases : l'exploration statistique des données, la préparation des données, la création de graphes DFG, la simplification du modèle par regroupement et filtrage selon des seuils préalablement définis et, finalement, l'analyse des résultats. Cette méthode permet la création de graphes DFG qui peuvent être utilisés dans un modèle agricole bioéconomique. Les graphes DFG obtenus étant fortement cycliques, une analyse plus approfondie à l'aide d'algorithmes de détection de cycles devra être réalisée. Connaître les habitudes culturelles des agriculteurs permet de mieux comprendre et anticiper les schémas de rotation de cultures. Cette connaissance peut aider les agronomes à prévoir l'impact environnemental potentiel des exploitations et orienter leurs services auprès des agriculteurs pour promouvoir une agriculture plus durable et productive. Une limite importante de cette méthodologie est la perte de conformité dans les flux. L'opération de groupement et de filtrage entraîne un éloignement du modèle avec la réalité. Cette question est traitée dans la

littérature [Leemans et al., 2019] et devra être étudiée dans de prochains travaux. Cette étude est un premier pas dans la modélisation des rotations de culture à l'aide des techniques de process mining. Ensuite, le but sera de prédire les besoins du sol pour la prochaine culture pour enfin optimiser les opérations d'épandage, en amont. Il existe plusieurs perspectives d'améliorations telle l'utilisation d'ontologie qui pourrait améliorer l'opération de regroupement. Ou encore l'approfondissement de la question du choix des seuils puisqu'ils ont un impact significatif sur les graphes obtenus. L'étape de filtrage devrait faire l'objet d'une réflexion plus poussée puisqu'elle semble nécessaire à la lisibilité du modèle, mais entraîne une perte d'information pouvant être problématique. Finalement des analyses complémentaires telles qu'une comparaison des relations existant entre les cultures pour deux périodes données seraient intéressantes afin de visualiser les modifications des pratiques agricoles en matière de rotation de culture.

6 REMERCIEMENTS

Nous tenons à remercier notre partenaire industriel pour sa collaboration dans le projet ainsi que IVADO et PROMPT pour leur soutien financier.

7 REFERENCES

- Agriculture and Agroalimentaire Canada. (1995). Cadre écologique national pour le Canada. Ottawa/Hull: Groupe de travail sur la stratification écologique.
- Bachinger, J., & Zander, P. (2007). ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. *European Journal of Agronomy*, 26(2), 130-143.
- Biazzo, S. (2002), Process mapping techniques and organisational analysis: Lessons from sociotechnical system theory, *Business Process Management Journal*, Vol. 8 No. 1, pp. 42-52. <https://doi.org/10.1108/14637150210418629>
- Castellazzi, M., Wood, G., Burgess, P., Morris, J., Conrad, K., & Perry, J. (2008). A systematic representation of crop rotations. *Agricultural Systems*, 97(1-2), 26-33.
- Clark, A. (2007). Benefits of cover crops. Dans *Managing cover crops profitably*. 3rd ed (p. 9). Beltsville, MD.: Sustainable Agriculture Network.
- CRAAQ. (2003). Production de framboises biologiques. pp.4 disponible sur : <https://www.craaq.qc.ca/Publications-du-CRAAQ/production-de-framboises-biologiques/p/PABI0004>, consulté le 10/02/2021.
- Cook, J., & Wolf, A. (1995). Automating Process Discovery through Event-Data Analysis. 17th International Conference on Software Engineering, (p. 73). Seattle, Washington, USA.
- Cui, Z., Liu, Y., Huang, Z., He, H., & Wu, G.-L. (2019). Potential of artificial grasslands in crop rotation for improving farmland soil quality. *Land Degradation & Development*, 30, 2187-2196.
- Detlefsen, N., & Jensen, A. (2007). Modelling optimal crop sequences using network flows. *Agricultural Systems*, 94(2), 566-572.
- Dogliotti, S., Rossing, W., & van Ittersum, M. (2003). ROTAT, a tool for systematically generating crop rotations. *European Journal of Agronomy*, 19(2), 239-250.
- Dustdar, S., Hoffmann, T., & van der Aal, W. (2005). Mining of ad-hoc business processes with TeamLog. *Data & Knowledge Engineering*, 55(2), 129-158.

- Grignon, E., (2015). Itinéraires techniques des principales productions maraichères du Québec comme outil permettant la gestion et le transfert efficace des connaissances liées à la régie de culture entre les conseillers et les producteurs. MAPAQ, Gouvernement Québécois. Consulté le 10/02/2021. Disponible sur : https://www.mapaq.gouv.qc.ca/sitecollectiondocuments/Agroenvironnement/1589_rapport.pdf
- Gouvernement Canada. (2017, 05 03). Cadre Écologique National pour le Canada : Données SIG. (Gouvernement Canada) Consulté le 11 21, 2020, sur http://sis.agr.gc.ca/pages/nsdb/ecostrat/gis_data.html
- Gouvernement Québec. (2019, 01 31). Zone agricole du Québec. (Données Québec) Consulté le 11 24, 2020, sur <https://www.donneesquebec.ca/recherche/dataset/zone-agricole-du-quebec>
- Institut de la statistique Québec. (2020). AGRICULTURE. Dans Le Québec chiffres en main (p. 34). Québec: Bibliothèque et Archives nationales du Québec.
- Janovicek, K., Hooker, D., Weersink, A., Vyn, R., Deen, B. (2021), Corn and soybean yields and returns are greater in rotations with wheat. *Agronomy Journal*. DOI:10.1002/agj2.20605
- Pretty, J., et al. (2010). The top 100 questions of importance to the future of global agriculture. *International Journal of Agricultural Sustainability*, 8(4), 219–236.
- Kirkegaard, J.A., Ryan, M.H. (2014). Magnitude and mechanisms of persistent crop sequence effects on wheat. *Field Crops Research*, Vol.164, pp.154-165
- Kirkegaard, J.A., Christen, O., Krupinsky, J., Layzell, D., (2008). Break crop benefits in temperate wheat production. *Field Crops Research*, 107, pp.185-195
- Klein Haneveld, W., & Stegeman, A. (2005). Crop succession requirements in agricultural production planning. *European Journal of Operational Research*, 166(2), 406-429.
- Kopitke, P., Menzies, N., Wang, P., McKenna, B., & Lombica, E. (2019). Soil and the intensification of agriculture for global food security. *Environment International*, 132.
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience And Remote Sensing Letters*, 14(5).
- Le Ber, F., Benoît, M., Schott, C., Mari, J.-F., & Mignolet, C. (2006). Studying crop sequences with CarrotAge, a HMM-based data mining software. *Ecological Modelling*, 191(1), 170-185.
- Leemans, S., Fahland, D., & Van der Aalst, W. (2018). Scalable process discovery and conformance checking. *Software & Systems Modeling*, 17, 599–631.
- Leemans, S., Poppe, E., & Wynn, M. (2019). Directly Follows-Based Process Mining: Exploration & a Case Study. *International Conference on Process Mining (ICPM)*. Aachen, Germany.
- Levavasseur, F., Bouty, C., Barbottin, A., Verret, V., & Martin, P. (2015). Characterization of crop rotations variability by combining modelling and local farm interviews. 5th International Symposium for Farming Systems Design. 7-10 September 2015, Montpellier, France
- Liakos, K., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *sensors*, 18(8).
- Ma, Y. (2019). Seed coating with beneficial microorganisms for precision agriculture. *Biotechnology Advances*, 37(7).
- Nieves, D., Ramírez-Quintana, M., Monserrat, C., Ferri, C., & Hernández-Orallo, J. (2020). Learning alternative ways of performing a task. *Expert Systems With Applications*, 148.
- Osman, J., Inglada, J., & Dejoux, J.-F. (2015). Assessment of a Markov logic model of crop rotations for early crop mapping. *Computers and Electronics in Agriculture*, 113, 234-243.
- Porter, P. (2009). Crop Rotations in Organic Production Systems. Dans *In Organic Farming: The Ecological System* (pp. 49-67). Charles Francis (Ed.).
- Pretorius, J., & Matthee, M. (2006). The Impact of Spatial Data on the Knowledge Discovery Process. *Proceedings of the Conference on Information Technology in Tertiary Education*. Pretoria, South Africa.
- Reeves, D. (1994). Principles of Crop Rotation. Dans *Cover Crops and Rotations* (p. 127). Boca Raton, FL: J.L. Hatfield and B.A. Stewart (Ed.).
- Salmon-Monviola, J., Durand, P., Ferchaud, F., Oehler, F., & Sorel, L. (2012). Modelling spatial dynamics of cropping systems to assess agricultural practices at the catchment scale. *Computers and Electronics in Agriculture*, 81, 1-13.
- Statistique Canada. (2018, 01 23). Introduction à la Classification écologique des terres (CET) 2017. (Gouvernement Canada) Consulté le 11 21, 2020, sur <https://www.statcan.gc.ca/fra/sujets/norme/environnement/cet/2017-1>
- The Fraunhofer Institute for Applied Information Technology. (2020). Process Discovery. (PM4PY) Consulté le 11 25, 2020, sur <https://pm4py.fit.fraunhofer.de/documentation>
- Turcotte, F. (2018). Prairies et pâturages pour chevaux : ce qu'il faut savoir. MAPAQ. Gouvernement Québec. Consulté le 11/02/2021. Disponible sur : https://www.mapaq.gouv.qc.ca/fr/Regions/monteregie/articles/production/Pages/Prairies_paturages_chevaux_ce_quil_faut_savoir.aspx
- United Nations. (2017). World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100. (Department of Economic and Social Affairs) Consulté le 11 24, 2020, sur <https://www.un.org/development/desa/en/news/population/world-population-prospects-2017.html>
- USDA. (s.d.). Crop Rotation Practice Standard. (Agricultural Marketing Service) Consulté le 11 22, 2020, sur <https://www.ams.usda.gov/grades-standards/crop-rotation-practice-standard>
- van der Aalst, W. (2011). Process Discovery: An Introduction. Dans *Process Mining : Discovery, Conformance and Enhancement of Business Processes* (pp. 125-156). Berlin, Heidelberg: Springer.
- van der Aalst, W. (2011). Introduction. Dans *Process Mining : Discovery, Conformance and Enhancement of Business Processes* (pp. 1-25). Berlin, Heidelberg: Springer.
- van der Aalst, W. (2016). Inductive Miner Based on Event Log Splitting. Dans *Process Mining : Data Science in action second edition* (p. 223). Eindhoven, The Netherlands: Springer.
- Yamasu, R., Bandaru, V., & Pnvr, K. (2020). pre-season crop type mapping using deep neural networks. *Computers and electronics in agriculture*, 176.
- Zhang, C., Liping, D., Lin, L., & Guo, L. (2019). Machine-learned prediction of annual crop planting in the US Corn Belt based on historical crop planting maps. *Computers and Electronics in Agriculture*, 166.

CIGI 2021

C3PO : Une ontologie développée pour le pilotage de la production fromagère au Québec

LUCAS GUILHEN¹, ROBERT PELLERIN¹, CHRISTOPHE DANJOU¹

¹ Polytechnique Montréal
C.P. 6079, SUCC. CENTRE-VILLE,
MONTREAL (QUEBEC) CANADA H3C 3A7
lucas.guilhen@polymtl.ca
robert.pellerin@polymtl.ca
christophe.danjou@polymtl.ca

Résumé – À l’image de tous les autres secteurs, les PME manufacturières dans l’industrie fromagère sont soumises à une réalité économique difficile. Elles doivent à la fois améliorer leur productivité, assurer une qualité de produit exigeante et maintenir des coûts d’exploitation réduits. De nouvelles technologies numériques, rassemblées sous le giron de l’industrie 4.0, ont le potentiel d’aider ses organisations à atteindre ces objectifs. Ces outils, auparavant utilisés par les grandes entreprises et pratiquement inexistantes pour le secteur fromager, doivent être adaptés au contexte particulier des petits fromagers. Face à ce constat, le Centre d’Expertise Fromagère du Québec (CEFQ) a souhaité développer des outils modernes pour faire face à cette nouvelle réalité. De ce fait, nous avons proposé le développement d’une ontologie d’application permettant la représentation des concepts clés liés à la transformation fromagère, ainsi que leurs liens. Elle permet ainsi de saisir le fonctionnement d’une fromagerie et constitue de ce fait un premier pas vers le développement futur d’outils modernes permettant de tirer parti des leviers de l’industrie 4.0. En effet, cette ontologie pourra être utilisée afin de mettre en œuvre diverses techniques de valorisation de données en fromageries et permettre le développement d’un outil de pilotage de la transformation fromagère.

Abstract – Like all other sectors, manufacturing SMEs in the cheese industry are subject to a difficult economic reality. They must simultaneously improve their productivity, ensure a demanding product quality, and maintain low operating costs. New digital technologies, brought together under the umbrella of Industry 4.0, have the potential to achieve these objectives. These tools, previously used by the large companies for which they were developed, must therefore be adapted for the cheese sector in which they are practically non-existent. Faced with this situation, the Centre d’Expertise Fromagère du Québec (CEFQ) wanted to develop modern tools to address this new reality. Therefore, we proposed the development of an application ontology to represent the key concepts related to cheese processing, as well as their links. It allows to understand the functioning of a cheese factory and is therefore a first step towards the future development of modern tools to take advantage of the levers of Industry 4.0. Indeed, this ontology can be used to implement various data enhancement techniques in cheese dairies and will allow the development of a tool for managing cheese processing.

Mots clés – Ontologie, Agro-alimentaire, Fromageries, Cartographie, Processus

Keywords – Ontology, Agri-food, Dairies, Mapping, Process

1 INTRODUCTION

Le Québec et l’Ontario sont les provinces produisant de très loin le plus de lait au Canada. En 2018, ces deux provinces regroupent respectivement 48.3% et 33.4% des producteurs laitiers canadiens (AGÉCO, 2018). Il est donc normal de retrouver dans ces provinces une très grande utilisation du lait. A l’image de tous les autres secteurs, les PME manufacturières dans l’industrie fromagère sont soumises à une réalité économique difficile. Historiquement, le Québec et le Canada étaient très fermés aux importations de produits laitiers étrangers. La récente modification des accords commerciaux internationaux impose plus de pression sur les petits artisans fromagers pour améliorer leur productivité tout en assurant une

qualité de produit répondant aux exigences de la clientèle et de la régulation agroalimentaire, et cela en maintenant des coûts d’exploitation réduits.

L’introduction rapide des nouvelles technologies numériques adaptées au secteur manufacturier a encore plus accentué ces pressions. Ces nouvelles technologies permettent de regrouper l’ensemble des ressources des entreprises, facilitant ainsi la prise de décision basée sur des données pertinentes en temps réel. Ce nouveau contexte industriel, né au début des années 2010, porte le nom d’Industrie 4.0 et le secteur fromager a tout intérêt à prendre part à cette nouvelle ère industrielle. On note en particulier les nouvelles pratiques de gestion des opérations et de contrôle de la production au sein des PME qui adoptent et

mettent à profit des systèmes de gestion intégrée (ERP), des systèmes d'exécution manufacturière (MES) ainsi que de multiples outils favorisant le suivi des activités de transformation en temps réel.

Le Centre d'Expertise Fromagère du Québec (CEFQ) a lancé un projet de grande envergure afin de doter les fromagers artisans du Québec d'outils modernes leur permettant de faire face à cette nouvelle réalité. Le nombre et la variété de ces dernières se reflètent dans les données qu'elles créent. De plus, les liens entre ces données et leur potentiel de valorisation demeurent incertains ou inconnus. Pour répondre à ce besoin, **nous proposons le développement d'une ontologie applicative du pilotage des procédés de transformation pour les petits fromagers du Québec.** Concrètement, ce modèle opérationnel comprend la définition de processus standards intégrés et le développement de modèles de valorisation des données de pilotage. La connaissance des données utilisées ou générées lors des activités de transformation est cruciale au développement de ses deux éléments du programme de recherche.

Cet article est structuré de la façon suivante. Nous détaillerons dans un premier temps une revue de littérature mettant en avant des solutions envisagées au problème d'identification des données et de leurs liens. Dans le chapitre 3, nous introduirons la méthodologie de recherche préconisée. Cette méthodologie s'appuie fortement sur l'analyse de cas réels. Le chapitre 4 présentera les résultats obtenus et le modèle proposé. Dans le chapitre 5, nous testerons la robustesse du modèle développé avant de discuter des résultats. Nous concluons dans le chapitre 6 avec une présentation des principales contributions de ce projet et des opportunités de recherche subséquentes à ces travaux.

2 REVUE DE LITTÉRATURE

2.1 Introduction

Dans le cas du projet du CEFQ, il est nécessaire d'identifier les données pertinentes dans la production fromagère. Pour y parvenir, il est capital de passer par une étape d'analyse des processus existants afin d'identifier les améliorations potentielles, comme on le fait par exemple dans des projets de réingénierie des processus ou d'amélioration continue. Nous avons donc effectué une revue des projets ayant réalisé des activités d'identification de données dans l'industrie fromagère ou laitière. Dans une deuxième partie, nous nous penchons sur les différents travaux adressant l'identification des relations entre ces données. Enfin, nous concluons en effectuant une revue critique de la littérature afin de montrer les manques des solutions proposées.

2.2 Identification des données dans l'industrie fromagère

L'identification des données pertinentes dans un contexte de production est nécessaire dans de nombreux projets, car elle permet d'avoir une vision précise du problème considéré. Elle peut être réalisée notamment via des techniques de cartographie de processus.

De nombreux langages et modèles ont été développés, permettant de couvrir tous les aspects d'un système (fonctionnel, comportemental, informationnel...) et de doter les ingénieurs chercheurs d'outils de représentations des processus et de représenter les flux de données. Ainsi, les travaux de De Cicco et al. (2007) ont produit une modélisation du processus de production de fromage de chèvre, de la production du lait à la transformation fromagère, en utilisant le formalisme DEVS (Discrete Event System Specification—

Spécification de système à événements discrets). Coelho et al. (2017) ont présenté une cartographie des processus de production et de distribution d'un fromage d'appellation protégée dans le but d'améliorer leur efficacité globale. Ce travail a été réalisé dans le but de soutenir le développement durable de ce fromage. D'autres projets ont vu le jour afin de réduire ou d'éliminer les sources de déchets. La méthodologie Lean Six Sigma (LSS) a été appliquée chez un producteur de produits laitiers en Norvège (Powell et al. (2017)). De façon similaire, l'utilisation de la technique VSM (Value Stream Mapping—flux informationnel et matériel, une carte de « l'état actuel ») par Melvin and Baglee (2008) a permis d'identifier les sources de déchets et de trouver des solutions potentielles à leur réduction dans l'industrie laitière britannique. Lors de projets d'amélioration de la qualité, des activités d'analyse et d'identifications des procédés sont mises en place. C'est le cas du rapport produit par Costa Dias et al. (2012) sur l'implantation des GMP (Good Manufacturing Practices – Bonnes manières de fabrication) dans une usine de fabrication de mozzarella qui a montré une nette réduction du ratio de non-conformités/conformités avant et après le projet. Ces outils ont été utilisés dans différents travaux de recherche afin d'identifier les données nécessaires. Par exemple, une étude des processus en place a été réalisée dans un projet de développement d'un système de contrôle informatique d'une boulangerie (Blagodatsky et al., 2019). Un travail similaire a été réalisé par Chen and Chen (2009) dans un projet de développement d'un système de traçabilité pour le riz organique en Indonésie.

2.3 Identification des relations entre les données de l'industrie fromagère

Il est aussi capital d'identifier les relations entre ces données afin de comprendre leur utilité lors de prises de décisions ou leur potentiel de valorisation. Au cours des dernières années, de nombreux travaux ont été menés afin de doter cette industrie d'instruments de valorisation des données. Dans les travaux de Aikenhead et al. (2015), une cartographie des processus et du système en place sur un site canadien de production laitière a été réalisée. Les diagrammes de boucle causale ont permis de montrer les répercussions entre toutes les opérations du site et d'identifier l'inefficacité de l'utilisation des ressources à travers les processus, ainsi que certains risques. De son côté, Fairweather (2010) a utilisé des cartes causales pour construire des cartes des environnements socioéconomiques des fermes laitières en Nouvelle-Zélande.

Le problème de la complexité et de la diversité des données est un problème récurrent dans l'industrie manufacturière. Pour répondre à ce problème, les chercheurs ont proposé diverses solutions permettant l'intégration et le lien des données. Une solution intéressante à ce problème est l'utilisation d'une ontologie. L'ontologie désigne un modèle de données permettant de représenter un ensemble de concepts dans un domaine, ainsi que les relations liant ces concepts. Son emploi permet d'exploiter les objets du domaine concerné, avec raison. Un travail intéressant a été réalisé par l'équipe de Pénicaud et al. (2019). Leur travail a été de développer une ontologie, basée sur PO2 (Process and Observation Ontology), spécialisée sur la production de fromage à pâte cuite. Le but de cette ontologie est de créer un vocabulaire commun, afin d'exploiter une base de données présentant des données de sources et de natures différentes pour les lier et résoudre différentes problématiques. Leurs travaux se basent sur ceux de Ibanescu et al. (2018). Ils ont également développé une ontologie PO2 DG spécialisée dans la production de lait caillé.

Leur plus grande inspiration a été l'ontologie développée par Dibia et al. (2016). Leur ontologie porte le nom de [MS]2O pour Multi-Step and Multi-Scale Ontology - une ontologie multi étapes et multi échelles concernant les procédés de transformation des organismes biologiques.

De leur côté, Verhoosel and Spek (2016) ont développé une ontologie pour l'exploitation laitière. Cette ontologie, nommée CDO – Common Dairy Ontology, a pour but d'uniformiser les données recueillies sur plusieurs années à l'aide de différents capteurs installés dans une exploitation laitière. Il est également intéressant de citer les travaux de la FAO qui développe et maintient à jour un dictionnaire, nommé AGROVOC.

2.4 Revue critique

Dans le cas de l'industrie fromagère au Québec actuellement, il n'y a aucun travail qui a été réalisé afin d'identifier les données pertinentes de l'industrie fromagère ou de cartographie des processus en place dans les fromageries. En effet, il n'existe pas de travaux qui ont permis de décrire à la fois la production laitière et la transformation fromagère dans sa globalité, qui prennent en compte les aspects de gestion de la qualité et de la production, tout en assurant la traçabilité et la tenue des inventaires d'une entreprise. Les outils proposés pour identifier les données proposent des vues partielles de l'industrie fromagère, car ils identifient seulement une partie des données et leurs relations en fonction de leurs objectifs de recherche. De ce constat, nous proposons le développement d'un nouvel outil permettant d'identifier et de modéliser la globalité des données de l'industrie fromagère au Québec, comprenant également leurs relations. Cet outil permettra d'apporter une structure exploitable dans l'optique de valoriser les données issues de la transformation fromagère.

3 METHODOLOGIE

Afin de répondre aux besoins du CEFQ et en considérant les résultats de la revue de littérature que nous avons menée, nous pouvons définir trois objectifs spécifiques de recherche :

- Objectif 1 : identifier les données existantes dans l'industrie fromagère québécoise;
- Objectif 2 : identifier les liens entre les données issues du premier objectif; et
- Objectif 3 : développer et valider un modèle permettant d'exploiter ces données et leurs relations.

Pour répondre à ces objectifs, nous proposons une méthodologie mettant en œuvre plusieurs activités, notamment une analyse de l'état actuel des fromageries grâce à une cartographie de leurs processus d'affaires. De cette cartographie, nous avons développé une ontologie décrivant le pilotage de la production fromagère au Québec. Les ontologies, dans le domaine de l'ingénierie, sont particulièrement appréciées pour leur capacité de partage et de diffusion des connaissances, leurs aspects évolutifs et participatifs, ainsi que pour toutes leurs potentielles utilisations (base de données, outils d'aides à la décision, systèmes d'information ...)

Nous nous basons sur la méthodologie DRM — Design Research Methodology afin de cadrer notre démarche globale. Cette méthodologie prévoit plusieurs phases : une phase de clarification de la recherche, une phase d'étude descriptive de l'existant, une phase de développement ontologique et enfin une phase d'expérimentation de l'ontologie développée. Afin de développer notre ontologie, nous nous sommes aidé de la méthode NeOn (Network of Ontologies), développée par

Suárez-Figueroa (2010). Ce choix s'est effectué sur la grande généralité de cette méthodologie, en permettant le développement d'ontologies à partir de sources non ontologiques comme une cartographie des processus par exemple, et prévoit également l'utilisation possible d'ontologies existantes.

La première activité concerne l'identification des concepts clés potentiels. Un concept clé est une représentation abstraite d'un objet ou d'un ensemble d'objets, d'une idée générale, ayant des caractères communs et un comportement commun. Cette identification a été effectuée à partir des données issues de la cartographie de processus, des solutions existantes ainsi que de la revue de littérature. Les étapes suivantes concernent la sélection des concepts clés et des relations à conserver dans l'ontologie développée. Enfin, une activité d'implantation permet de transcrire les modèles semi-informatiques dans le langage d'implantation ontologique défini, grâce au logiciel Protégé. C'est un logiciel gratuit dont le code source est libre d'accès. Il a été développé par des chercheurs de l'université Stanford pour permettre la création d'ontologie et est très populaire dans le domaine du Web sémantique. Nous avons ainsi obtenu une ontologie peuplée avec les données récoltées dans la phase d'acquisition des connaissances. Finalement, nous avons effectué une activité de validation. Cette dernière s'est appuyée sur une étude ACV (Analyse de Cycle de Vie) réalisée par une équipe de recherche externe dans un cadre de travail similaire. Une comparaison entre les données de cette étude et notre ontologie a permis de valider sa pertinence.

4 RESULTATS

4.1 Cartographie de processus

Grâce à toutes les notes et données collectées lors des visites de 4 fromageries, nous avons pu dresser une cartographie des processus observés en fromagerie. Le défi de cette étape est le choix de la méthode de cartographie la plus appropriée au processus non linéaire des fromageries. Nous avons choisi la formalisation ANSI, qui reste très simple et facilement partageable aux fromagers qui sont novices dans ce domaine. Étant donné l'ampleur et l'envergure du projet et des données collectées lors des visites, il était vital d'avoir une structure claire permettant de se retrouver et de naviguer facilement à travers toute la cartographie. Nous avons structuré notre cartographie sur trois niveaux de détail, de plus en plus spécifiques. :

- Niveau 1 — BPML : ce niveau est la *Business Process Master List*, soit la liste principale des processus d'affaires. Dans ce niveau, on répertorie de manière très agrégée tous les processus cartographiés.
- Niveau 2 — Cartographie ANSI : c'est le niveau de modélisation. C'est à cette étape que nous avons utilisé le langage de modélisation ANSI pour décrire les processus observés dans les fromageries.
- Niveau 3 — Fiche-activité : pour chaque activité présente dans la modélisation (niveau 2) d'un processus, on détaille un fichier avec toutes les informations de l'activité (intrants, extrants, équipement utilisés, données générées, KPI, éventuels problèmes rencontrés, ...).

Au total, la cartographie comporte un fichier BPML comportant la liste des processus observés, 29 fichiers de processus modélisés et 205 fiches-activités.

4.2 Identification des concepts potentiels

Afin d'identifier les concepts clés potentiels, nous nous sommes appuyés sur les modélisations de processus ainsi que

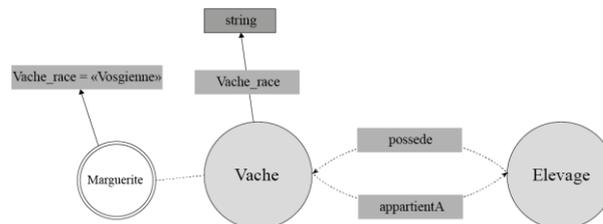
sur les fiches-activités. Les différentes activités ainsi que les acteurs apparaissant dans la modélisation peuvent correspondre à des concepts clés et nous avons donc recensé tout ce qui nous semblait pertinent pour le pilotage de la production fromagère. De la même manière, nous avons parcouru les fiches-activité pour identifier toutes les données pouvant correspondre à des concepts clés. Le partenariat entre Polytechnique Montréal et le CEFQ a permis l'utilisation de nombreuses ressources du CEFQ, en particulier leur documentation. Nous avons eu accès à de nombreux documents « types », comme des registres ou des procédures. Ces documents possèdent un nombre important d'informations liées aux processus d'affaires des fromageries. De la même manière que nous avons analysé nos fiches-activité, nous avons également parcouru la documentation du CEFQ afin de recenser les données pouvant être des concepts-clés. Au total, nous avons créé une liste de plus de 200 concepts clés, issus de l'analyse empirique de l'existant ainsi que de la littérature. Un travail similaire a été réalisé sur les relations entre les concepts. Nous avons utilisé les mêmes sources de données. Ces listes comportent un grand nombre de concepts, plus ou moins pertinents pour le modèle que l'on souhaite développer. Il convient donc de trier et sélectionner les concepts que l'on souhaite retenir pour l'ontologie à développer.

4.3 Sélection des concepts clés et relations

Afin de procéder à cette sélection, nous avons établi une méthode, qui permet l'examen individuel de chaque terme afin de définir d'une part son applicabilité et sa pertinence dans ce projet, ainsi que son rôle et sa place dans l'ontologie que l'on souhaite développer. Un terme est un mot identifié lors de nos différentes recherches et qui semble avoir une importance dans le domaine considéré. En effet, un terme est applicable et pertinent dans ce projet s'il est applicable au milieu de la transformation fromagère d'une part, et d'autre part s'il répond à des besoins de suivi de production, de gestion de la qualité, ou de logistique d'une fromagerie. Nous nous concentrons sur ces aspects, car il s'agit des objectifs souhaités par le CEFQ. Si un terme est applicable et pertinent pour le projet, il faut ensuite élucider d'éventuels conflits avec d'autres termes puis décider de son rôle et de sa place. Les différents rôles sont les suivants : Concept, Instance de concept ou Propriété de concept. Un concept est une notion clé d'une ontologie, il s'agit de représenter une idée. Une instance de concept est une version concrète d'un concept, à travers la mise en œuvre de l'ontologie sur un cas. Une propriété est une caractéristique d'un concept. Ces distinctions sont illustrées dans la figure 1. Prenons par exemple le terme *Fromage*. Il est applicable au contexte de la transformation fromagère, son lien est assez évident. Il est pertinent pour les objectifs du CEFQ, étant donné que c'est un concept central dans cette étude. Il n'y a pas de conflit sémantique avec d'autres termes dans la liste, et on peut assez aisément fixer son rôle comme concept dans l'ontologie. Nous exerçons le même raisonnement avec le terme *Article*, qui est un terme assez générique, applicable dans la transformation fromagère donc, et pertinent pour les objectifs du CEFQ. En effet, il permettra de décrire et rassembler un grand nombre d'objets présents en fromagerie. Il existe un éventuel conflit avec le terme *Produit*. En réalité, l'idée véhiculée par *Produit* est incluse dans *Article*, et nous pourrions décider de fixer *Produit* comme un sous concept d'*Article*, car le terme d'*Article* comprend d'autres objets, comme les matières premières ou les auxiliaires de production par exemple. D'ailleurs, à l'étape de hiérarchisation, nous pouvons détailler l'ordre hiérarchique suivant : *Fromage* est un

sous concept de *Produit*, lui-même sous concept d'*Article* (*Article* > *Produit* > *Fromage*). Les termes *Produit*, *Article* et *Fromage* sont validés et ordonnés. Il est à noter que dans cet exercice de modélisation, une grande part de réflexion et de choix sont liés à l'interprétation du modélisateur, de ses expériences et de ses connaissances du domaine. Un travail collaboratif peut permettre d'ajouter de nouvelles visions et de nouveaux liens entre les concepts, mais il peut aussi freiner l'expression et l'efficacité de la modélisation.

Figure 4.1 Illustration des relations au sein d'une ontologie



4.4 Implémentation de l'ontologie

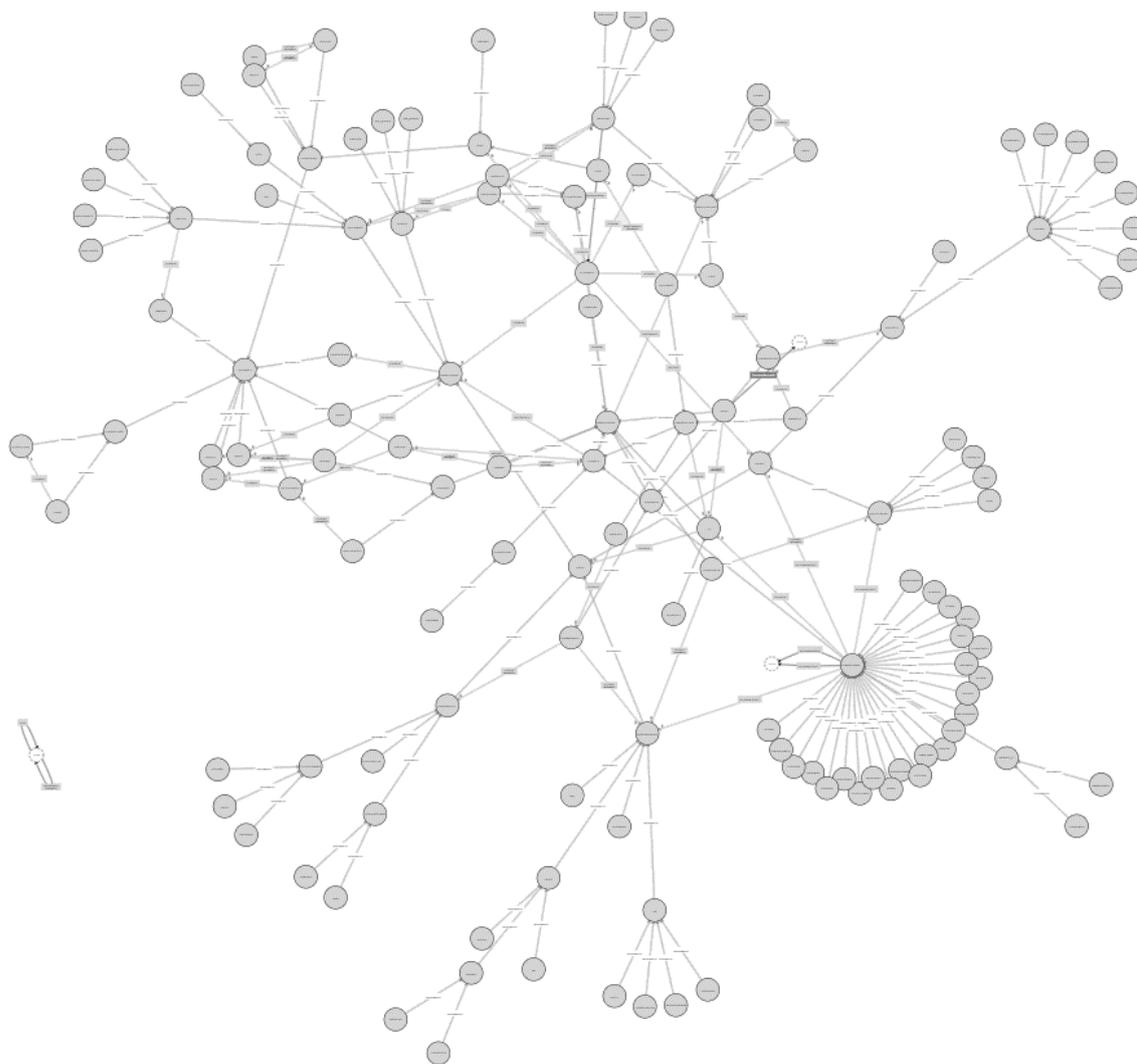
Afin de construire à proprement parler l'ontologie, nous utilisons le logiciel Protégé. Selon le site web de Stanford, Protégé «est un environnement d'édition d'ontologie riche en fonctionnalités, avec une prise en charge complète du langage d'ontologie Web OWL 2 et des connexions directes en mémoire aux raisonneurs logiques de description comme HermiT et Pellet. ». Ce logiciel permet de créer puis peupler une ontologie via ses constituants : concepts (appelés *Classes*), liens (appelés *Object properties*), propriétés (appelés *Data properties*), instances (appelés *Individuals*). Nous avons dans un premier temps implémenté tous les concepts retenus lors de la sélection décrite plus haut, en prenant soin de documenter chacun afin de permettre la bonne compréhension de l'ontologie par tous les utilisateurs. Protégé permet l'ajout d'annotations pour chaque objet de l'ontologie. Il est également possible à cette étape de préciser si des concepts sont équivalents ou disjoints à d'autres, ce qui permettra aux raisonneurs logiques de déceler d'éventuelles erreurs de conception de l'ontologie.

Après avoir implémenté les concepts, nous pouvons définir puis assigner les liens. L'éditeur de liens permet, à l'image des concepts, de spécifier l'équivalence ou le caractère inverse entre des liens. Il est aussi possible de fixer certaines spécifications à des liens : fonctionnalité, transitivité, symétrie ou asymétrie, réflexivité. Enfin, Protégé permet de préciser le domaine et la portée d'une relation, qui correspondent respectivement au domaine de définition d'une fonction (en anglais « domain »), et à son ensemble d'arrivée (en anglais, « range »). Protégé permet également d'assigner les quantificateurs suivants aux relations : some, exactly, minimum, maximum, only. Nous avons ensuite créé toutes les propriétés liées à chaque concept. Protégé possède un grand nombre de types de donnée, qui correspondent aux types définis par le consortium W3C.

Il est possible de se rendre compte à cette étape de l'impertinence de certains concepts qui sont sous-utilisés, peu reliés à d'autres concepts, ou complètement inutilisés. Ils pourront tout de même avoir des instances lors du peuplement de l'ontologie, mais l'absence de relations avec d'autres concepts marque leur relative faible importance dans ce contexte.

Une vue globale de l'ontologie est donnée en figure 2. L'ontologie est explorable à l'adresse suivante : <https://bit.ly/2ZABqL2>. Dans la section suivante, nous décrivons la méthode de validation que nous avons adoptée.

Figure 2 Vue globale de l'ontologie, représentation simplifiée avec concepts et liens uniquement



5 VALIDATION

Afin de valider la pertinence d'une ontologie, il est nécessaire de tester sa cohérence logique, structurelle, ainsi que sa cohérence sémantique. En effet, on considère qu'une ontologie est adaptée si le choix de termes employés est adapté et si sa structure est logique et ne présente d'incohérences.

5.1 Validation structurelle

Pour parvenir à la validation structurelle de l'ontologie, nous avons fait appel à des outils et descripteurs logiques qui utilisent les règles implantées et définies dans Protégé pour déceler les incohérences logiques. Il existe un nombre important de raisonneurs logiques qui ont été développés dans ce but. Hermit et Pellet étant les raisonneurs les plus reconnus dans le domaine, nous avons choisi de les utiliser tout au long du développement de l'ontologie. Des plug-ins permettant leur utilisation dans Protégé nous ont permis de vérifier au fur et à mesure de la construction de l'ontologie la cohérence structurelle de celle-ci. Nous avons aussi utilisé OOPS! à la fin du développement de l'ontologie afin de vérifier si nous avons commis des erreurs. Ces différents outils n'ont pas révélé d'erreurs majeures et nous a permis d'identifier des éventuelles corrections mineures à apporter à l'ontologie.

5.2 Validation sémantique

La validation sémantique d'une ontologie est la plus compliquée à effectuer, car elle nécessite une vision et une compréhension experte du domaine considéré. Tout comme la modélisation comporte une part de subjectivité et d'interprétation, il est fort possible que la validation sémantique en comporte également.

Afin de procéder à cette vérification, nous avons décidé de nous baser sur une étude de cas d'analyse de cycle de vie (ACV) de la production d'un fromage Cheddar au Québec et son impact environnemental. L'étude d'ACV utilisée comprend les étapes allant de la production du lait cru jusqu'au conditionnement du fromage produit et son départ du transformateur fromager. Elle exclut les étapes de distribution et de consommation du fromage. Ce n'est pas dérangeant pour notre validation puisque ce sont des étapes du cycle de vie du produit qui ne sont pas ou que légèrement représentées dans notre ontologie. Nous avons choisi d'utiliser des données issues d'une ACV suivant l'assertion suivante : si l'ontologie que nous avons développée permet la représentation des données d'un ACV (dans un cadre de travail similaire et avec les mêmes hypothèses), alors nous pouvons considérer que notre ontologie est suffisamment fine et adaptée à la représentation du domaine considéré.

D'une manière similaire à l'identification et à la sélection des concepts de notre ontologie, nous avons parcouru la description et les diagrammes présents dans la thèse de Benoit (2018) pour extraire les concepts tests. Ainsi, les mots essentiels utilisés pour décrire l'étude de cas ont été recensés et consignés pour obtenir une liste détaillée de concepts tests. Cette liste comprend 94 concepts tests et 16 liens tests. Nous avons pu ensuite comparer et associer les différents concepts et liens avec ceux de notre ontologie, et obtenir des scores de représentation. Nous considérons qu'un concept de notre ontologie est capable de représenter avec adéquation une conception test s'il n'y a pas d'ambiguïté entre les deux termes et que leurs propriétés sont équivalentes. Le même raisonnement est applicable à la comparaison des liens. Les scores de représentation sont résumés au tableau 1. Un score

de représentation de 100 % indique que la totalité des données tests est représentable par des données de notre ontologie.

Tableau 1 Scores de représentation

Il peut y avoir plusieurs cas possibles :

- Adéquation parfaite : le terme employé dans les deux cas est le même et il ne peut y avoir aucune ambiguïté ;

Représentation des concepts	Représentation des liens
Nombre de concepts tests : 94	Nombre de liens tests : 16
95 %	67 %

- Adéquation sémantique : les termes employés diffèrent légèrement ou sont synonymes. Dans ce cas, nous avons considéré qu'il n'y avait pas d'ambiguïté sémantique si le sens véhiculé par les deux termes était le même. Cette décision relève en grande partie d'une bonne connaissance du domaine et d'un choix subjectif de modélisation et représentation ;
- Lien inverse manquant : le lien test correspond à une relation inverse d'un lien présent dans l'ontologie, mais qui n'a pas été implanté ;
- Lien manquant : le lien test ne correspond à aucun lien de l'ontologie ;
- Le concept test correspond à une propriété d'un concept de l'ontologie. Prenons par exemple le concept test Dimension. Il n'y a pas de concept équivalent dans l'ontologie développée, cependant il s'agit d'une propriété du concept Équipement. Dans ce cas, nous pouvons donc considérer que cette donnée est représentée dans l'ontologie ; et
- Le concept test correspond en réalité à un résultat de calcul ou à un indicateur lié à un calcul issu de plusieurs propriétés de concepts. Prenons par exemple le concept test Dépense énergétique. Il n'y a pas de concept équivalent, cependant il pourrait être calculé grâce aux propriétés de consommation des Équipements par exemple. Dans ce cas, nous pouvons donc considérer que cette donnée est représentée dans l'ontologie.

Dans certains cas, il est pertinent de compléter l'ontologie avec les nouvelles informations afin de couvrir les données manquantes. Là aussi, nous pouvons remarquer que la validation vient apporter une valeur ajoutée au développement de l'ontologie en permettant d'augmenter le nombre de concepts représentés.

6 CONCLUSION

Après un travail d'identification des données, grâce à une cartographie détaillée des procédés en place, nous avons développé un modèle de données standard, sous la forme d'une ontologie. Cet outil permet la représentation et la compréhension du pilotage d'une fromagerie, en mettant à disposition tous les concepts clés pertinents, ainsi que leurs relations. Ce premier travail est la base solide qui permettra le développement d'un outil applicatif destiné au CEFQ et à ses membres, et qui servira à la fois à stocker les données issues des fromageries, mais aussi et surtout à les valoriser par l'exploitation des propriétés des concepts et de leurs liens. Ainsi, il sera possible au niveau du CEFQ, de dégager des tendances de production fromagère. À l'échelle d'une fromagerie, il sera possible de d'identifier les différents facteurs influençant son rendement fromager par exemple.

Au cours de ce projet, nous avons pu découvrir de nombreuses problématiques liées à la création d'une ontologie. Il est notamment important de disposer de connaissances solides sur le domaine et le modélisateur a tout intérêt à se familiariser le plus possible au domaine en question. Le développement collaboratif peut influencer l'avancement du projet, en apportant de nombreux avantages au moment du développement, dans le choix de la terminologie par exemple, mais il est parfois plus efficace pour une personne de travailler seul afin d'obtenir une modélisation cohérente.

Il est également capital de bien définir le cadre et les limites de l'ontologie, car il est facile de déborder sur des aspects non pertinents. Les limites peuvent parfois être difficiles à cerner, et le développement de l'ontologie peut rapidement s'alourdir et prendre du retard. De manière similaire, il est important de définir le niveau de détail qu'on souhaite couvrir, car là aussi, il est facile de se lancer dans un niveau de détails trop important qui ne fait qu'alourdir l'ontologie.

Enfin, il serait intéressant de discuter de l'universalité de l'ontologie produite. Sera-t-il possible de représenter toutes les productions possibles (actuelles et futures)? À l'heure actuelle, cet objectif semble atteint. À l'avenir il serait pertinent d'apporter de nouvelles contributions à cette ontologie, par exemple par l'ajout de concepts concernant la production et la transformation d'autres produits dérivés du lait (yaourt, crème, lait en poudre, caséine...), ou de concepts concernant d'autres pans de l'industrie laitière (gestion du personnel, gestion financière, gestion de la maintenance...). Ce travail d'extension pourra par exemple être effectué par l'interconnexion avec d'autres ontologies, ce qui permet la méthodologie employée.

En somme, nous avons exprimé le besoin de développer un outil de représentation du domaine de la transformation fromagère, qui s'inscrit dans le courant de la révolution numérique en cours. Nous avons ensuite répondu aux objectifs énoncés, en développant une ontologie spécialisée sur le pilotage de la transformation fromagère. Ce nouveau modèle, favorisant l'interopérabilité entre les systèmes, permettra la réalisation et le développement de projets innovants au sein des fromageries québécoises, leur permettant ainsi rentrer dans l'ère du 4.0.

7 REFERENCES

AGÉCO, G. (2018). *Statistiques globales*.
<http://www.groupeageco.ca/fsl/>

Aikenhead, G., Farahbakhsh, K., Halbe, J., & Adamowski, J. (2015). Application of process mapping and causal loop diagramming to enhance engagement in pollution prevention in small to medium size enterprises: case study of a dairy processing facility. *Journal of Cleaner Production*, 102, 275-284.
<https://doi.org/10.1016/j.jclepro.2015.04.069>

Benoit, S. (2018). *Évaluation de l'éco-efficacité des procédés de transformation des produits laitiers : développement d'un outil de simulation* [Université Laval].

Blagodatsky, G. A., Vologdin, S. V., Gorokhov, M. M., & Dokuchaev, D. E. (2019). Information system UML-model development by OMG RUP technology for Food industry enterprises. *IOP Conference Series*:

Materials Science and Engineering, 537.
<https://doi.org/10.1088/1757-899x/537/4/042037>

Chen, H. H., & Chen, S. C. (2009). Constructing the process models for the traceability of rice production and distribution. WMSCI 2009 - The 13th World Multi-Conference on Systemics, Cybernetics and Informatics, Jointly with the 15th International Conference on Information Systems Analysis and Synthesis, ISAS 2009 - Proc.,

Coelho, D., Carrola, T., & Couvinhas, A. (2017). Improvement of Certified Artisan Cheese Production through Systemic Analysis—Serra da Estrela PDO. *Sustainability*, 9(3).
<https://doi.org/10.3390/su9030468>

Costa Dias, M. A., Sant'Ana, A. S., Cruz, A. G., Faria, J. d. A. F., Fernandes de Oliveira, C. A., & Bona, E. (2012). On the implementation of good manufacturing practices in a small processing unit of mozzarella cheese in Brazil. *Food Control*, 24(1-2), 199-205.
<https://doi.org/10.1016/j.foodcont.2011.09.028>

De Cicco, A., De Gentili, E., & Santucci, J. F. (2007). Modelling and simulation applied to the cheese tradition. Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation, ICMA 2007,

Dibie, J., Dervaux, S., Doriot, E., Ibanescu, L., & Pénicaut, C. (2016). *[MS]2O - A multi-scale and multi-step ontology for transformation processes: Application to micro-organisms* [Conference Paper]. Lecture Notes in Computer Science
https://link.springer.com/chapter/10.1007%2F978-3-319-40985-6_13

Fairweather, J. (2010). Farmer models of socio-ecologic systems: Application of causal mapping across multiple locations. *Ecological Modelling*, 221(3), 555-562. <https://doi.org/10.1016/j.ecolmodel.2009.10.026>

Ibanescu, L., Allard, T., Dervaux, S., Dibie, J., Guichard, E., Pénicaut, C., & Raad, J. (2018). A use case of data integration in food production. *CEUR Workshop Proceedings*,

Melvin, A., & Baglee, D. (2008). Value stream mapping: a dairy industry prospective. *2008 IEEE International Engineering Management Conference (IEMC-Europe 2008)* 2008 IEEE International Engineering Management Conference, Piscataway, NJ, USA.

Pénicaut, C., Ibanescu, L., Allard, T., Fonseca, F., Dervaux, S., Perret, B., Guillemain, H., Buchin, S., Salles, C., Dibie, J., & Guichard, E. (2019). Relating transformation process, eco-design, composition and sensory quality in cheeses using PO2 ontology. *International Dairy Journal*, 92, 1-10.
<https://doi.org/10.1016/j.idairyj.2019.01.003>

Powell, D., Lundebj, S., Chabada, L., & Dreyer, H. (2017). Lean Six Sigma and environmental sustainability: the case of a Norwegian dairy

producer. *International Journal of Lean Six Sigma*,
8(1), 53-64. <https://doi.org/10.1108/ijlss-06-2015-0024>

Suárez-Figueroa, M. C. (2010). *NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse* [Universidad Politécnica de Madrid].

Verhoosel, J. P. C., & Spek, J. (2016). Applying ontologies in the dairy farming domain for big data analysis. *CEUR Workshop Proceedings*,

Estimation du poids et du volume des fruits et légumes via les technologies d'imagerie et de vision par ordinateur : état de l'art

AHLAM BOUDINA¹, MAHA BEN ALI¹, ROBERT PELLERIN¹, CHRISTOPHE DANJOU¹

¹Laboratoire Poly-Industries 4.0, Département de mathématiques et de génie industriel, Polytechnique Montréal
2500 Chemin de Polytechnique, Montréal, Canada

ahlam.boudina@polymtl.ca, maha.ben-ali@polymtl.ca, robert.pellerin@polymtl.ca, christophe.danjou@polymtl.ca

Résumé – Le volume et le poids sont des attributs importants de l'inspection externe de la qualité des fruits et légumes. Le besoin d'automatiser l'estimation de ces derniers d'une manière fiable devient imminent dans un contexte d'industrie 4.0 agroalimentaire où la rapidité est synonyme de rentabilité. Plusieurs technologies de vision par ordinateur tentent de capturer la réalité, mais très peu d'entre elles sont vraiment adaptées aux industries. De la vision 2D à la vision 3D, en passant par la 3D reconstruite, les algorithmes développés sont différents en fonction de chaque fruit et légume, des paramètres externes qui peuvent résumer le poids et/ou le volume et des méthodes employées dans l'estimation finale de ces deux attributs physiques. Seule une bonne combinaison des trois volets permet de donner une estimation fiable. L'enjeu devient important compte tenu de certains produits réputés irréguliers et non- axisymétriques. Cet article d'état de l'art présente les principales méthodes analytiques pour l'estimation du poids/volume des fruits et légumes, ainsi que les technologies employées pour arriver à cette fin.

Abstract – Volume and mass are the most important attributes for an external quality inspection of fruits and vegetables. To make these attributes automatically estimated with the highest accuracy, manufacturers need to implement a food industry 4.0 process to gain in time and thus, in profitability. In such a context, many computer vision technologies attempt to capture reality, but very few of them are suitable due to industrial constraints. From 2D imaging to 3D and reconstructed 3D imaging technologies, the algorithms developed are different depending on each fruit and vegetable specifications, on external parameters that can summarize the mass and /or volume variation, and finally on the methods employed to predict these two physical attributes. Only a good combination of the three components can give a reliable estimation. This task becomes challenging considering some irregular and non-axisymmetric product shapes. Therefore, this review discusses the main solutions for automatic estimation of the mass/volume of fruits and vegetables, as well as the technologies employed to achieve this objective.

Mots clés – Industrie 4.0, agroalimentaire, vision par ordinateur, poids, volume.

Keywords – Industry 4.0, agrifood, computer vision, mass, volume.

1 INTRODUCTION

L'inspection externe de la matière première agricole est une phase cruciale du processus de transformation agroalimentaire. Elle intervient pour assurer une production optimale et éviter aux industriels des pertes conséquentes (déchets, temps) pendant la transformation (Demartini et al., 2018; Luque, Peralta, de las Heras, & Córdoba, 2017). C'est le cas pour la transformation des fruits et légumes qui présentent une hétérogénéité, au sein d'un même lot. D'une part, plusieurs paramètres externes de la matière première (p.ex. la couleur, la texture, la forme, la taille, le volume et le poids des fruits et légumes) doivent être évalués et comparés aux standards préétablis de la qualité (S. Jana, Parekh, & Sarkar, 2019; Zhang et al., 2014). D'autre part, les exigences de rentabilité font de cette tâche un réel enjeu : évaluer ces paramètres externes d'une manière rapide, fiable et automatisée devient inévitable afin d'ajuster en temps réel les installations de l'usine.

Dans un tel contexte, les solutions intelligentes reposant sur les technologies de l'industrie 4.0 peuvent remplacer les méthodes traditionnelles souvent réputées par leur subjectivité et lenteur (Chen, Zhang, Zhao, & Ouyang, 2013). En particulier, l'émergence de technologies de vision par ordinateur et de traitement d'images a permis le développement de plusieurs

techniques d'évaluation externe des fruits et légumes. Il s'agit de méthodes non destructives qui suscitent l'intérêt grâce à leur rapidité, leur précision et leur objectivité (Luque et al., 2017). En général, la fiabilité de ces méthodes est étroitement liée à la forme de l'objet inspecté. Moins de régularité implique plus de complexité dans le processus d'évaluation. Par conséquent, le choix de la technologie et des méthodes d'estimation à employer devient décisif (Kumar & Parkavi, 2020). Dans ces conditions, les possibilités d'intégration de telles solutions technologiques dans un milieu industriel deviennent plus complexes partant du fait que plusieurs choix doivent se faire simultanément, tout en prenant en considération les spécificités opérationnelles de chaque usine.

Dans cet article, nous présentons une revue de littérature en relation avec l'estimation du poids et du volume des fruits et légumes dans un contexte d'automatisation industrielle. D'abord, nous décrivons la relation qui existe entre le volume et le poids des fruits et légumes, ainsi que des exemples de paramètres externes permettant de les estimer. Ensuite, un état de l'art sur les technologies de vision par ordinateur employées et sur les méthodes d'estimation développées à ce sujet est présenté. Enfin, une critique sur la faisabilité des méthodes en respectant les principes d'une usine intelligente est exposée. Une attention particulière est portée aux travaux traitant des fruits et légumes non-axisymétriques.

2 METHODOLOGIE DE RECHERCHE

Une revue exploratoire a été exécutée en utilisant la base de données Compendex (étape 1 dans la Figure 1), ce qui a permis de distinguer deux différentes vagues de littérature selon la forme du produit étudié (régulière/axisymétrique ou irrégulières/non-axisymétrique). Notons que la majorité des publications traitent les produits de formes régulières et axisymétriques (S. Jana et al., 2019).

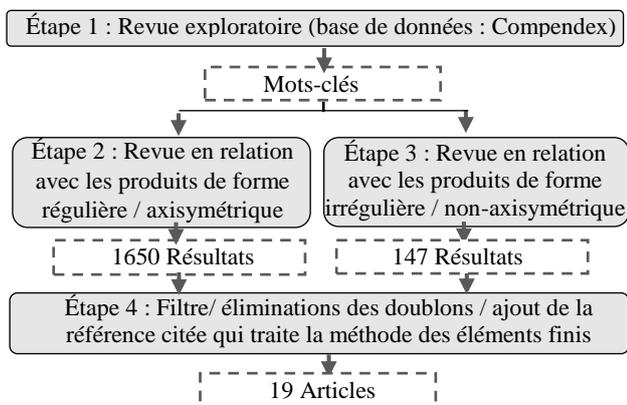


Figure 1 Méthodologie de recherche

Partant de ce fait, deux étapes supplémentaires ont été exécutées. Dans l'étape 2, nous nous sommes concentrés sur les publications traitant les fruits et les légumes de formes régulières et axisymétriques. En particulier, nous avons considéré les études les plus récentes. Dans l'étape 3, nous nous sommes concentrés sur les publications traitant les fruits et les légumes de formes irrégulières et non-axisymétriques. Dans l'étape 4 plusieurs articles ont été éliminés sur la base de la redondance dans les méthodes ou dans les technologies employées. Nous avons enfin retenu une liste de publications qui tente de couvrir toutes les méthodes et technologies, en particulier celles qui apportent des nouveautés dans une ou plusieurs étapes des processus. Cette liste est présentée dans le Tableau 1

3 ESTIMATION DU POIDS ET DU VOLUME DES FRUITS ET LEGUMES

Dans cette section, nous évoquons l'importante relation qui existe entre le volume et le poids des fruits et légumes. Ensuite, nous présenterons des exemples de paramètres externes employés dans l'estimation du poids/volume, vu que la sélection des paramètres explicatifs joue sur la fiabilité de la méthode proposée et influence le choix de la technologie et la méthode d'estimation du volume/poids.

3.1 Relation entre le poids et le volume

La relation entre ces deux composantes évaluées lors d'une inspection externe des fruits et légumes a été abordée dans la majorité des publications retenues. Cette relation est très significative pour certains fruits et légumes. En général, les auteurs estiment d'abord le volume (certains s'arrêtent à cette étape), puis confirment l'hypothèse de relation avec le poids. C'est le cas par exemple de l'étude conduite par Huynh, Tran, and Dao (2020), où la densité moyenne déduite à partir de l'échantillon permettait d'expliquer le poids en fonction du volume estimé. Les coefficients de détermination résumant la relation en question pour le concombre et la carotte ont été respectivement 0,9199 et 0,9289. L'étude souligne aussi la présence de corrélations très significatives issues de divers travaux ultérieurs. Ces derniers estimaient le poids selon le même principe (soit la densité multipliée par le volume) et

renvoyaient un coefficient de corrélation de 0,9978 pour la mangue, 0,9045 pour la mandarine, 0,9592 pour le citron vert, 0,9641 pour le citron, et 0,9232 pour l'orange.

Tableau 1 Publications retenues et produits/paramètres étudiés.

Auteurs	Produits/paramètres
Produits irréguliers	
Haque et al. (2020)	Patate douce : Longueur courbée et droite, largeur maximale, longueur de la queue et du corps, courbures de 31 ST.
Villordon, Gregorie, and LaBonte (2020)	Patate douce : Voxels
Long et al. (2018)	Pomme de terre : Pixels, échelle de gis de chaque pixel.
S. Jana et al. (2019)	Pomme de terre : Contours, le rayon de chaque ST et sa longueur.
Su et al. (2018)	Pomme de terre : Pixels, profondeur de chaque pixel.
Lopez-Juarez, Rios-Cabrera, Hsieh, and Howarth (2018)	Pomme de terre : Pixels contenus dans les 3 images
Produits réguliers	
Xie, Wang, and Yang (2019)	Carotte : Surface supérieure (Pixels).
Ornek (2018)	Carotte : Longueur, pour chaque ST : longueur et largeur ou rayon.
Huynh et al. (2020)	Carotte et concombre : Contours, pour chaque ST : longueur, largeur et profondeur.
Soleimanipour and Chegini (2019)	Concombre : Contours, courbure, circonférence, longueur, surface et volume en voxels
Al-kaf, Chia, and mubin (2020)	Ananas : Circonférence et diamètre du haut, milieu et bas, longueur.
Nyalala et al. (2019)	Tomate : Surface supérieure, longueur et largeur, excentricité, distance radiale, Volume en voxels et superficie 3D.
Cui, Yang, Daesik, and Cho (2017)	Tomate : Contours, pour chaque ST : longueur et largeur, courbure.
Mon and ZarAung (2020)	Mangue : Pour chaque ST : profondeur selon l'intensité de la lumière, longueur et largeur.
Goñi, Purlis, and Salvadori (2007)	Pomme : Contours, coordonnées 3D de N nœuds de maillage
Vivek Venkatesh, Iqbal, Gopal, and Ganesan (2015)	Pomme/citron : Contours de chaque produit, longueur et largeur
Format général	
Siswantoro, Prabuwno, Abdullah, and Idrus (2014)	Général : Contours supérieurs et inférieurs du produit, coordonnées 3D de N points aléatoires
Jadhav, Singh, and Abhyankar (2019)	Général : Voxels
Susovan Jana, Parekh, and Sarkar (2020)	Divers : Contours, rayon de chaque ST et sa longueur

Haque et al. (2020), Susovan Jana et al. (2020), et Mon and ZarAung (2020) ont aussi souligné la même notion de densité que celle de Huynh et al. (2020). Vivek Venkatesh et al.

(2015) et Su et al. (2018) ont utilisé des régressions linéaires pour modéliser le rapport existant entre le poids et le volume, tandis que Goñi et al. (2007) et Nyalala et al. (2019) ont exprimé cette relation via des régressions non-linéaires. Sans quantification, Ornek (2018) et Soleimanipour and Chegini (2019) ont mis en évidence l'importante relation entre le volume et le poids. Nous retenons donc que le volume est généralement estimé avant de déduire le poids des fruits et légumes.

3.2 Quelques exemples de paramètres externes utilisés pour l'estimation du volume et du poids

Il est important de choisir la meilleure combinaison des paramètres externes qui expliquent le volume et le poids d'un fruit ou d'un légume, vu que cette sélection influence la fiabilité de la méthode proposée. Selon Zhang et al. (2014), les paramètres les plus importants pour une inspection externe sont ceux qui expliquent la forme (p.ex. la compacité, la convexité, l'élongation, la rondeur, la longueur et la largeur ainsi que leurs variantes, la description des contours), en plus des moments centrés et les descripteurs de Fourier. Ces deux derniers sont des approximations mathématiques en deux dimensions des courbes modélisant les contours d'un objet donné. L'identification de tels paramètres externes pour estimer le volume et le poids est une tâche difficile à réaliser par un robot, particulièrement lorsque le produit étudié présente une irrégularité ou un défaut externe (Zhang et al., 2014). En examinant les paramètres présentés dans le Tableau 1, nous remarquons qu'il n'y a pas de standards pour faire la sélection de ces paramètres, néanmoins certains sont mieux connus que d'autres pour résumer la variabilité du volume/poids. Nyalala et al. (2019) soulignent que pour les fruits la surface de chaque objet sur un angle donné (appelée surface projetée), son périmètre, son excentricité, sa longueur, sa largeur et sa distance radiale constituent les six paramètres les plus utilisés. Toutefois, par souci de fiabilité, certains fruits (ou légumes) nécessiteraient l'utilisation de plus de paramètres, qu'ils soient tirés de plans 2D ou encore de 3D. C'est le cas d'ailleurs pour Soleimanipour and Chegini (2019), S. Jana et al. (2019), Cui et al. (2017), Huynh et al. (2020) et Goñi et al. (2007) qui ont inclus la description des contours dans leurs estimations du volume/poids (d'une manière directe ou indirecte). Selon Zhang et al. (2014), la description des contours contribue à mieux expliquer le volume/poids. D'autres auteurs ont pris en considération des paramètres 3D. Notons pour exemples Villordon et al. (2020), Haque et al. (2020), Nyalala et al. (2019), Mon and ZarAung (2020), Soleimanipour and Chegini (2019), Su et al. (2018), Jadhav et al. (2019), Lopez-Juarez et al. (2018), Huynh et al. (2020) et Goñi et al. (2007). Contrairement à toutes ces études, Xie et al. (2019) n'ont employé que la surface supérieure pour prédire le poids de la carotte.

Certains de ces paramètres sont plus difficiles à identifier dans un milieu industriel qui impose plus de contraintes reliées au temps et au budget alloués pour l'inspection externe. Afin de mieux comprendre ce défi, la partie suivante abordera les différentes technologies employées dans la littérature pour la détection automatique des paramètres externes.

4 TECHNOLOGIES D'IMAGERIE ET DE VISION PAR ORDINATEUR UTILISEES POUR ESTIMER LES PARAMETRES EXTERNES DES FRUITS ET LEGUMES

Les technologies d'imagerie et de vision par ordinateur ont rapidement évolué au cours des dernières années en suivant le développement des caméras et des processeurs (Chen et al.,

2013). Dans cette section, nous nous concentrerons uniquement sur les technologies utilisées pour estimer les paramètres externes des fruits et légumes. Le Tableau 2 résume les systèmes de vision et les technologies d'imagerie adoptés par les publications retenues dans le Tableau 1.

4.1 Systèmes de vision

Tel que présenté par la Figure 2, un système de vision (SV) est généralement composé d'une ou de plusieurs caméras, d'une source d'éclairage, d'une station de tests, d'un échantillon et d'un ordinateur doté d'un logiciel spécialisé (en traitement d'images, en analyse et modélisation, etc.) connecté à la station de tests (Abdullah, 2007; Chen et al., 2013; Du et al., 2020; Smith, Zhang, Hansen, Hales, & Smith, 2018; Zhang et al., 2014).

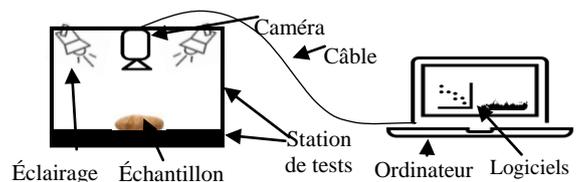


Figure 2 Système de vision standard

La qualité des images produites par un SV et qui serviront à l'estimation des paramètres externes des fruits et légumes dépend en grande partie de l'éclairage et du type de caméras utilisées. En effet, l'illumination adéquate des surfaces étudiées évite d'avoir un temps de traitement supplémentaire causé par un bruit (résultat d'une ombre et/ou d'un éblouissement) (Blasco, Aleixos, Cubero, & Lorente, 2012). Plusieurs auteurs ont utilisé un SV standard, tandis que d'autres ont apporté d'autres éléments au SV de base (stations de base rotatives, éclairage de la caméra, tissus noirs enveloppant les murs intérieurs, plusieurs caméras, etc.).

4.2 Vision par ordinateur 2D

Appelée également « système de vision traditionnel », la technologie 2D imite la vision de l'humain en capturant des images RGB qui résument les données par couleur : Red pour rouge, Green pour le vert et Blue pour le bleu (Zhang et al., 2014). Ce système est souvent équipé avec des caméras de type CCD (Charge Coupled Device) ou CMOS (Complementary Metal Oxide Semiconductor) les deux connues sous le nom de caméras ou capteurs RGB 2D. Cette technologie permet de convertir les images au format numérique.

Il est souvent très commun d'apporter des changements minimes aux SV standards pour la technologie 2D qui permet de générer une seule image prise sur un angle donné. Ainsi, pour estimer plus de paramètres, notamment ceux complexes, il est indispensable d'avoir recours à d'autres types de technologies.

4.3 Vision par ordinateur 3D passive

La vision par ordinateur 3D passive (ou encore la reconstruction 3D : R-3D) tente d'estimer les paramètres d'un objet inspecté, généralement détectables sur un plan 3D, à l'aide de plusieurs images sur un plan 2D (Moreda, Ortiz-Cañavate, García-Ramos, & Ruiz-Altisent, 2009). La littérature distingue plusieurs techniques de reconstruction, parmi lesquelles Siswantoro, Prabuwno, and Abdulah (2013) ont souligné les plus importantes, soit la rétro-ingénierie, le lissage et l'intersection du volume. La rétro-ingénierie consiste à remodeler l'objet inspecté à partir de multiples sections transversales fictives qui le composent. Le lissage est une

Tableau 2 Synthèse des Technologies de vision par ordinateur et des méthodes employées dans l'estimation du poids/volume

Auteurs	Technique	Nb images	SV	Éclairage	Caméra	RGB IR/NIR	Méthode	Échantillon	Nb Tests	Classes	CPU en secondes	Performance ++ : 1-MAPE + : à défaut (la statistique est mentionnée)	Envir. Industriel
Haque et al. (2020)	R-3D	2	×	✓	Intégrée dans la trieuse Exeter	✓ ✓	Vol-G	125 79	×	2	NS	NS	×
Villordon et al. (2020)	3D	1	×	×	Occipital Inc.	×	Vox/Pix	210	×	×	NS	Vol : R ² = 0,9	(+) ×
Long et al. (2018)	3D	2	✓	×	RGB-D avec modèle NS	✓ ✓	Vox/Pix	120	×	2	NS	Vol 78% : Sans déformations 91% Avec déformations 66%	(++) ×
S. Jana et al. (2019)	R-3D	6~36	✓	✓	4x Nikon-S3600 digitale	✓ ×	Vox/Pix	700	100	×	5 à 25	Vol : 92,57%	(++) ×
Su et al. (2018)	3D	2	✓	✓	Primesense Camine 1.09	✓ ✓	Vox/Pix	110	70	2	NS	Poids 95,6% : Sans déformations 95% Avec déformations 96,6%	(++) ×
Lopez-Juarez et al. (2018)	R-3D	3	×	✓	CCD Basler firewire colour	✓ ×	Vox/Pix	50	×	×	NS	NS (classification)	✓
Xie et al. (2019)	2D	1	✓	✓	U2-YW500	✓ ×	ML	147	×	×	NS	R ² = 0,9022	(+) ×
Ornek (2018)	×	×	×	×	×	×	ML	464	100	×	NS	Vol : Modèle PUNN 97%	(++) ×
Huynh et al. (2020)	2D	1	✓	✓	CANON 60D	✓ ×	Vol-G	281	70	×	0,032 à 0,05	Carotte Vol : 96,6% Concombre Vol : 96,8%	(++) ×
Soleimanipour and Chegini (2019)	R-3D	1	✓	✓	fps IP camera	✓ ×	Vol-G	150	×	×	NS	Vol : 96,5% basé sur l'erreur relative totale	(+) ×
Al-kaf et al. (2020)	2D	1	✓	✓	NS/ Caméra CDD	✓ ×	ML	4	×	3	NS	NS	×
Nyalala et al. (2019)	2D/R-3D	9	✓	×	Kinect depth cameea	✓ ✓	ML	300	50	×	NS	Vol : 95,5% Poids : 97,2%	(++) ×
Cui et al. (2017)	2D	1	✓	✓	Téléphone intelligent	✓ ×	Vol-G	40	10	×	NS	R ² = 0,819	(+) ×
Mon and ZarAung (2020)	R-3D	1	✓	✓	Canon Power Shot SX110	✓ ×	Vol-G	150	20	×	0,45	Vol : 97,8% Poids : 94,6%	(++) ×
Goñi et al. (2007)	R-3D	NS	✓	×	IP Camera Model 550710	✓ ×	E-finis	12	×	×	NS	Vol : 98,18%	(++) ×
Vivek Venkatesh et al. (2015)	2D	5	✓	✓	JAI make CV – M7 + CL CCD	✓ ×	Vol-G	120	60	×	<0,047	Vol : 96,09% - Poids : 95,93%	(++) ×
Siswanto et al. (2014)	R-3D	5	×	✓	5 x Logitech web cameras	✓ ×	Monte C	150	×	×	9,65	Vol : 99,07% (≈ sur tous les fruits) (max de 3% d'erreur sur 98% des cas)	(++) ×
Jadhav et al. (2019)	R-3D	6~36	✓	✓	4x Nikon-S3600 digitale	✓ ×	Vox/Pix	700	100	×	5 à 25	Vol : 99,4%	(++) ×
Susovan Jana et al. (2020)	2D	1	✓	✓	Téléphone intelligent	✓ ×	Vol-G	77	×	×	0,76 à 0,94	Vol : 91,4% (PDT 92,5%-Citron 88,8% -tomate 89%) Poids : 91,7% (PDT 93%-Citron 89% -tomate 88,6%)	(++) ×

méthode qui fait une approximation continue du contour d'un objet, puis crée un objet reconstruit à partir de l'assemblage de courbes fictives autour du contour. Quant à l'intersection du volume, il s'agit d'un processus qui permet de rassembler plusieurs images en extrapolant les points à coordonnées 2D en des points à coordonnées 3D. Siswanto et al. (2013) ont également cité d'autres méthodes moins utilisées, c'est le cas par exemple de la théorie du *space carving* (soit un algorithme qui corrige les reconstructions 3D classiques pour donner à la fin de meilleures modèles 3D), et la méthode de fil de fer (soit une représentation d'un objet en sommets et arêtes, de façon à constituer un modèle en fils croisés) largement employée en impression 3D.

L'étude de Haque et al. (2020) propose de construire une image 3D de la patate douce à l'aide de deux images 2D obtenues par rotation manuelle à 90°. À cette fin, la méthode des sections transversales a été employée : les contours de chaque patate douce sur deux angles ont été alignés et orientés de façon à créer pour chacune des deux images un axe centroïde superposable. Ceci a permis d'avoir N ellipses reconstruites avec dimensions connues. Par la suite, plusieurs paramètres physiques ont été capturés à partir du modèle reconstruit grâce aux deux images tirées de la caméra intégrée dans une trieuse industrielle. Notons que, malgré l'introduction de la trieuse industrielle, la rotation manuelle de la patate douce fait que le SV utilisé (muni d'une source d'éclairage) ne respecte pas les conditions réelles de l'industrie.

En utilisant également la méthode des sections transversales, Mon and ZarAung (2020) ont reconstruit en 3D la mangue à partir d'une image 2D prise sur l'angle de face (soit un angle qui permet de déterminer la largeur). Les auteurs ont toutefois estimé mathématiquement la profondeur de chacune des N ellipses (c-à-d, les sections transversales) composant l'objet inspecté, et ce sur deux étapes. D'abord, la relation entre la largeur maximale et la profondeur maximale correspondante sur 3 points de chaque fruit a été établie grâce à une régression linéaire simple. Puis, à l'aide de l'intensité de la lumière projetée sur les N-3 points restants, les autres profondeurs ont été déduites en partant des 3 profondeurs maximales estimées dans la première étape. Ainsi, la mangue a été reconstruite à l'aide d'une seule image 2D et en utilisant uniquement l'éclairage de la caméra.

Dans un autre ordre d'idées, Soleimanipour and Chegini (2019) ont utilisé une seule image 2D obtenue avec un SV standard équipé d'une caméra RGB pour reconstruire différents concombres avec la technique du lissage. D'abord, les contours de chaque objet inspecté sont détectés sur une seule image 2D RGB. Ensuite, quelques points sont marqués sur les contours avec des distances équivalentes. Un sous-ensemble de pixels sur les contours a été interpolé à l'aide d'une courbe *B-spline*, de sorte que la courbure des deux extrêmes de l'objet soit maximale (les extrémités sont détectables en utilisant ce qu'on appelle la théorie des courbes algébriques). Enfin, les courbures *B-splines* sont assemblées pour créer un modèle reconstruit. La même méthode est utilisée par Gofii et al. (2007) pour reconstruire une pomme et un morceau de viande. En revanche, les courbures *B-splines* sont faites autour des sections transversales au lieu du contour de l'objet. De plus, un algorithme de Delaunay est utilisé pour créer un maillage sur la surface de l'objet à partir de plusieurs images prises horizontalement. Siswanto et al. (2014), de leur côté, ont choisi d'utiliser un SV standard avec cinq caméras, chacune donnant une image RGB. Le calibrage des caméras a été fait de sorte à reproduire les paramètres extrinsèques et intrinsèques des images 2D et à appliquer la

méthode de l'intersection du volume. Par la suite, une boîte enveloppante fictive avec des dimensions 3D connues est créée autour de chaque objet pour aider à estimer le volume. La même méthode de reconstruction est utilisée par Jadhav et al. (2019), mais avec un nombre plus important d'images (de 6 à 36 images). Les échantillons étaient placés sur une base tournante qui permettait de capturer plusieurs angles de l'objet en rotation. Les caméras RGB ont été calibrées dans le but de déterminer les limites du volume réel de l'objet et de créer différentes projections 2D. À partir de chaque image, il était possible de cerner la silhouette du fruit et de cartographier chaque point du plan réel sur un plan 3D reconstruit (processus souvent appelé « estimation des homographies inter-images »). La méthode de l'intersection du volume a été également utilisée par Lopez-Juarez et al. (2018) pour reconstruire la pomme de terre en 3D. À cette fin, une station d'inspection externe a été conçue. Leur SV industriel est muni de trois caméras et de plusieurs sources d'éclairage. L'estimation du poids est possible en prenant en photos chaque pomme de terre qui tombe du convoyeur entrant vers le convoyeur sortant (soit 3 photos sur 3 angles différents).

4.4 Vision par ordinateur 3D active

La vision par ordinateur 3D active (souvent appelée vision 3D) se distingue par l'utilisation de deux types de caméras. Le premier type utilise le principe d'un télémètre laser, et le deuxième se base sur le temps de vol de la lumière TOF (soit *Time-of-Flight*) (Gongal, Amatya, Karkee, Zhang, & Lewis, 2015). Ces caméras permettent d'envoyer des faisceaux laser pulsés, ou simplement des lumières, sur les surfaces des objets inspectés. L'intensité des lumières reflétées dépend de la distance avec les surfaces inspectées. Partant de ce principe, il est possible de déduire les différentes hauteurs à partir des lois de triangulation, ainsi, de réaliser une construction géométrique en 3D de l'objet (Gongal et al., 2015; Moreda et al., 2009; Villordon et al., 2020). En particulier, Villordon et al. (2020) ont utilisé un système de vision 3D non standard et rotatif pour construire le format de la patate douce. L'idée est de détecter automatiquement sa surface, son volume, sa longueur et sa largeur. Pour ce faire, les auteurs proposent de suspendre la patate douce à l'aide d'un fil et de capturer en 360° son format 3D à l'aide d'un scanner connecté à un iPad Air 2. Contrairement à Villordon et al. (2020), Su et al. (2018) et Long et al. (2018) ont utilisé des SV standards. Dans les deux études, deux images ont été prises avec des caméras à détection en profondeur (*3D depth*). Les deux images représentent en profondeur respectivement la partie inférieure et la partie supérieure de la pomme de terre et permettent ainsi de construire son modèle 3D. Nyalala et al. (2019) ont utilisé une approche différente pour estimer le poids de la tomate et ont considéré une combinaison de variables extraites d'un SV standard avec une rotation manuelle de l'objet. Le SV est doté d'une caméra Kinect profonde qui produit neuf images sur les deux plans 2D et 3D, le but étant d'avoir les paramètres de la tomate sur deux plans différents.

4.5 Traitement d'images

L'étape de traitement d'images est une étape commune dans toutes les publications retenues (voir Tableau 2), sauf pour l'étude de Ornek (2018) où le volume de la carotte a été estimé sans faire de vision par ordinateur.

Le traitement d'images pour l'estimation du volume et du poids des fruits et légumes comporte généralement trois phases. La première phase consiste à prétraiter les images obtenues à partir des SV (les nettoyer des bruits, convertir les espaces RGB, orienter les axes, etc.). La deuxième phase est

plus poussée que la première puisqu'elle comporte l'évaluation des images prétraitées, le calibrage des distances, l'extraction des paramètres souhaités, etc. La troisième phase consiste à l'utilisation de l'algorithme développé qui estime le volume et le poids de l'objet inspecté en fonction des paramètres obtenus à la deuxième phase, ou encore développer, voir personnaliser, les algorithmes intégrés aux technologies 2D, 3D passive et 3D active.

5 METHODES D'ESTIMATION DU POIDS ET/OU DU VOLUME

Nous présenterons dans cette section une synthèse des méthodes utilisées pour l'estimation du volume et du poids des fruits et légumes. À notre connaissance, il n'existe pas dans la littérature une classification claire pour les méthodes basées sur la vision par ordinateur. Dans ce qui suit, nous proposons une classification des méthodes développées pour estimer le poids/volume en cinq principales catégories : le calcul du volume géométrique et des sections transversales, l'apprentissage machine, le comptage de pixels et de voxels, la méthode Monte Carlo et enfin la méthode des éléments finis.

5.1 Le calcul du volume géométrique et des sections transversales

Il s'agit d'assimiler le fruit/légume à une forme géométrique, ou à défaut, d'assimiler les sections transversales qui le composent à une ou à plusieurs formes géométriques. Le volume est ensuite calculé selon la formule mathématique adéquate à chaque forme. Pour les sections transversales, la somme des volumes est appliquée pour estimer le volume total de l'objet. Dans ce sens, Vivek Venkatesh et al. (2015) ont d'abord classé les pommes, les citrons (verts et jaunes) et les oranges selon l'aspect de chaque unité. Ensuite, selon la forme déduite (circulaire, elliptique ou parabolique), ils calculent le volume en fonction des paramètres détectés lors du traitement d'images. De leur côté, Haque et al. (2020) ont reconstruit la patate douce à partir de 20 ellipses, puis divisé la représentation 3D en 31 sections transversales. La surface de chaque section a été par la suite calculée, ce qui a permis de déduire le volume de la patate douce. De façon similaire, Mon and ZarAung (2020) ont imaginé la mangue divisée en 50 sections transversales. Le volume est par la suite déduit en sommant les volumes des sections cylindriques et elliptiques. Soleimanipour and Chegini (2019) ont aussi estimé le volume de chaque section transversale qui compose le concombre après avoir fait une approximation ellipsoïde du fruit reconstruit par la méthode du lissage. S. Jana et al. (2019) et Susovan Jana et al. (2020) ont calculé la somme des volumes des mi-sections transversales (moitiés supérieure et inférieure) en rapprochant leurs formats à des demi-cercles. Cui et al. (2017) ont divisé la tomate en plusieurs sections circulaires. Huynh et al. (2020) ont procédé différemment. Le corps de chaque objet a été modélisé par un ensemble de plusieurs troncs de pyramides, et les deux extrémités ont été représentées comme des cônes elliptiques. Ensuite, la formule adéquate pour le calcul du volume de chaque composante a été appliquée. Les Figures 3 et 4 illustrent, de façon non-exhaustive, les différentes manières de faire correspondre les sections transversales à des formes géométriques.

5.2 L'apprentissage machine

Dans un contexte d'estimation du poids et/ou du volume des fruits et légumes, l'apprentissage machine est largement utilisé pour prédire les paramètres physiques des objets et étudier les relations entre eux. Dans ce qui suit, nous présentons uniquement les publications employant l'apprentissage machine pour la modélisation du poids et/ou du volume.

Al-kaf et al. (2020) ont proposé deux modèles mathématiques linéaires pour le calcul du poids de l'ananas. Le premier modèle tient compte de la longueur de l'ananas, tandis que le deuxième modèle tient compte des diamètres du haut, du milieu et du bas du fruit. Les coefficients de chaque modèle sont estimés à l'aide d'un réseau de neurones à une seule couche. Xie et al. (2019) ont proposé d'estimer le poids de la carotte via une régression polynomiale quadratique en fonction des pixels de la surface supérieure de cette dernière. Ornek (2018) ont aussi modélisé le poids de la carotte en fonction des paramètres de chaque section transversale qui la composent, et ce grâce à trois types de réseaux de neurones artificiels (*Back-propagation*, *Large Margin Nearest Neighbor LMNN* et *Product Unit Neural Network PUNN*). Nyalala et al. (2019) ont d'abord procédé à la modélisation du volume en considérant les paramètres 2D de la tomate, puis les paramètres 3D. Les deux ensembles de paramètres ont été par la suite combinés pour prédire le volume de la carotte. Les auteurs ont retenu deux modèles, soit le réseau de neurones artificiels Bayésien (*Bayesian-ANN*) et le *Support Vector Machine* avec *Radial Basis Function Kernel* (RBF-SVM) qui emploient les données 2D et 3D combinées.

5.3 Le comptage de pixels et de voxels

Cette méthode vient à la suite d'une construction 3D ou reconstruction 3D d'un objet. Une fois que le modèle numérique a été développé, chaque voxel occupant l'espace ou pixel en rapport à la profondeur (équivalent à l'échelle de gris) compose une unité volumétrique. L'ensemble des voxels (pixels \times profondeur) permettront d'obtenir le volume total de l'objet. Il est important dans un premier lieu d'établir la relation entre un pixel/voxel et un mm^3 . Le calibrage des caméras (ou encore le scan d'un objet témoin avec volume et dimensions connus) est une étape préalable à la définition du volume.

Dans ce sens, Villordon et al. (2020) ont utilisé le logiciel *Meshmixer* pour estimer le volume de chacun des modèles 3D de la patate douce en les comparant à un objet témoin. Long et al. (2018); Su et al. (2018) ont multiplié les pixels de la surface de la pomme de terre par le niveau de gris correspondant à chacun pour estimer le volume. Quant à Jadhav et al. (2019), ils ont converti la résolution des voxels de chaque modèle de fruit en mm^3 en se basant sur les paramètres de la caméra. Lopez-Juarez et al. (2018) ont estimé le poids selon le nombre moyen de pixels contenu dans les trois images de la pomme de terre.

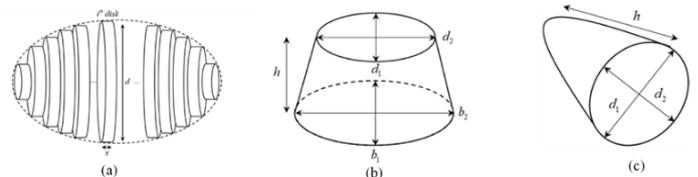


Figure 3 Simulation de quelques sections transversales (Huynh, Tran et al. 2020). (a) un objet découpé en disques, (b) une partie centrale d'un objet, (c) une extrémité d'un objet

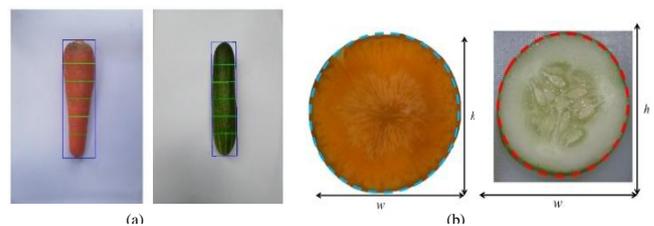


Figure 4 Approximation géométrique (Huynh et al. 2020) (a) Découpage d'un concombre et d'une carotte. (b) Approximation elliptique des sections transversales.

5.4 La méthode de Monte Carlo

La méthode consiste d'abord à tracer la silhouette d'un objet donné, puis à déterminer le ratio en volume qu'occupe cette silhouette par rapport au volume d'une forme témoin.

Siswanto et al. (2014) ont employé cette méthode pour estimer le volume des fruits et légumes à formats irréguliers. Ils ont d'abord créé une boîte fictive à volume connu V_0 enveloppant la reconstruction 3D de l'objet en question. Puis, ils ont généré N points 3D aléatoires à partir de la boîte et ils les ont projetés sur chacune des 5 images qui composent la reconstruction. Les N_0 points projetés contenus à l'intérieur de la silhouette sont considérés comme intersection. Ainsi, cette intersection donne lieu au ratio N_0/N qui permet de déterminer le volume de l'objet irrégulier V en rapport à V_0 et ce selon la formule (1).

$$V = V_0 \times N_0/N \quad (1)$$

5.5 La méthode des éléments finis

Avec cette méthode, on suppose qu'une variable générale (u) est égale à 1 sur tous les nœuds du modèle reconstruit (technique du lissage ou fil de fer). Ceci est équivalent à résoudre l'équation différentielle partielle ayant pour solution $u=1$, puis à calculer l'intégrale triple de (u) sur le domaine d'éléments (Ω_i). Le volume estimatif \tilde{V} de l'objet en question (Goñi et al., 2007) est obtenu par la formule (2).

$$\tilde{V} = \sum_i \iiint_{\Omega_i} u \, d\Omega_i \quad (2)$$

6 DISCUSSION ET CONCLUSION

Estimer automatiquement le poids et le volume des fruits et légumes peut se faire par différentes solutions, ces dernières doivent combiner trois volets : la détermination des variables explicatives, la conception du système/technologie de vision et le choix de la méthode estimative. En l'absence d'une solution combinatoire standard, une disparité dans la littérature a été fortement constatée. Cette dernière s'explique non seulement par les spécificités géométriques que peut présenter l'espèce étudiée mais aussi par les différentes technologies de vision (la reconstruction 3D plus précisément) et les façons d'extraire les paramètres externes (calibrage des caméras, conversion des pixels, apprentissage machine, utilisation d'objets témoins, etc.). Il est à noter que dans notre état de l'art, les champs d'expertise des auteurs interviennent en premier lieu dans cette dissemblance. En effet, la vision par ordinateur peut être associée à différentes disciplines (Kakani, Nguyen, Kumar, Kim, & Pasupuleti, 2020) : le traitement du signal, la reconnaissance des formes, le traitement d'images, l'intelligence artificielle, la photométrie, la conception optique et la géométrie, etc. La vision par ordinateur pourrait donc nécessiter des compétences pluridisciplinaires.

Excepté le cas de Lopez-Juarez et al. (2018), toutes les études présentées dans le Tableau 2 ont été réalisées dans des laboratoires expérimentaux. Ainsi, la plupart a permis d'atteindre des bons niveaux de précision (voir la colonne « Performance » dans le Tableau 2). Il est à noter que pour des fins de comparaison, nous avons standardisé plusieurs performances en calculant la moyenne de l'erreur absolue relative, quand cela était possible.

Nous avons relevé que pour la majorité des études, les échantillons étaient minutieusement sélectionnés pour avoir des unités uniformes. Or, la matière première présente généralement des difformités, en particulier dans les usines de

transformation agroalimentaire. Dans un tel contexte, les systèmes de vision 3D ou 3D-Reconstruits performant mieux que les technologies 2D. Cette performance est considérablement réduite si ces difformités sont importantes. Pour de tels cas, les technologies de vision 3D basées sur le TOF n'arrivent pas à faire une bonne approximation de la forme géométrique (Jadhav et al., 2019). De son côté, la reconstruction 3D consomme beaucoup de temps et exige des grandes capacités de calcul algorithmique (CPU et GPU) (Siswanto et al., 2014).

Parmi les méthodes d'estimation du poids et du volume des fruits et légumes, nous notons d'abord une large utilisation du calcul géométrique et du comptage des pixels/voxels. Pour les produits de formes plutôt régulières et qui ne présentent pas des déformations, la méthode du volume géométrique fournit de très bons niveaux de précision. Cependant, sa performance en cas de présence de déformations n'a pas été évaluée dans la littérature. Relativement au calcul des voxels, cette méthode garantit des résultats semblables à ceux du calcul géométrique. Mais il faut rappeler que les technologies de vision 3D sont moins performantes quand il s'agit de produits présentant des difformités. C'est pourquoi l'utilisation de cette méthode doit être reconsidérée (l'étude de Long et al. (2018) en est témoin). Dans de telles situations, des étapes supplémentaires de traitement d'image (portant principalement sur la réduction du bruit) peuvent améliorer les performances, tel que démontré par l'étude de Su et al. (2018). Vient ensuite, la méthode d'apprentissage automatique qui est moins utilisée dans l'estimation du poids et du volume des fruits et légumes, mais qui permet d'obtenir de bonnes performances pour les produits de formes régulières et sans difformités. Toutefois, ses algorithmes peuvent être adaptés selon l'objet étudié et ainsi permettre de traiter le cas des espèces hétérogènes (modélisation par classe). Enfin, la méthode Monte Carlo et celle des éléments sont les moins utilisées en dépit d'un excellent niveau de performance. Leurs complexités freinent toute tentative d'adaptation industrielle.

À première vue, les études retenues ne peuvent être généralisées pour le contexte industriel. Les milieux industriels exigent i) une rapidité dans le traitement de l'information, ii) aucune intervention manuelle pour la rotation de la matière première, et iii) un minimum de changement à apporter aux lignes de production. Dans notre littérature, le choix des technologies et des systèmes de vision était basé seulement sur l'objectif d'améliorer les estimations, et n'a pas pris en compte les enjeux d'implantation dans une usine intelligente. À cet effet, la majorité des principes clés reliés à l'industrie 4.0, notamment la virtualisation, la décentralisation, le système en temps réel et par conséquent l'orientation service, était partiellement ou totalement non respectée si nous considérons chaque solution proposée comme un système cyber-physique. Pourtant, les composantes de base y sont présentes (sous-systèmes physiques et non physiques, IIoT, méga données et analytique, infonuagique...etc.). Une éventuelle tentative d'intégration dans ce sens nécessiterait plus de ressources matérielles (SV industriels automatiques) et des efforts techniques pour corriger les erreurs résultantes afin d'améliorer l'estimation.

Somme toute, en tenant compte de l'aspect opérationnel, de la possibilité d'intégration dans une usine intelligente et de l'hétérogénéité que peut présenter la même espèce d'un fruit ou légume, l'apprentissage automatique reste un choix à considérer. Il peut d'ailleurs combiner les deux méthodes les plus utilisées dans la littérature (Vol-G et Vox/Pix) en équilibrant leurs résultats et ce, sans aucune intervention

manuelle. Cette méthode peut également modéliser les variabilités expliquées par les paramètres des ST et des pixels/voxels réunis. De plus, les algorithmes de ML ont un pouvoir d'adapter leurs calculs à chaque classe de produit détectée. Il serait cependant plus judicieux d'inclure d'abord plus de paramètres externes résumant les défauts visibles (dommages externes, trous, moisissures...etc.), puis de procéder à une préalable classification (ou segmentation) par rapport aux : formes, formats et défauts externes de chaque fruit et légume.

Cet état de l'art recense les différentes technologies de vision par ordinateur employées pour l'estimation du volume/poids des fruits et légumes. Il s'agit d'un point de départ pour la sélection de l'outil approprié selon le contexte industriel. Dans ce sens, cet article constitue une base pour une étude de cas qui sera réalisée dans une usine de transformation de patates douces (forme fortement irrégulière et non-axisymétrique) et permettra d'identifier les outils appropriés pour automatiser l'estimation de son poids/volume dans un contexte 4.0.

7 REMERCIEMENTS

Les auteurs remercient Worximity Technology Inc. et MITACS pour le financement.

8 REFERENCES

- Abdullah, M. Z. (2007). 1 - Computer vision and infrared techniques for image acquisition in the food and beverage industries. Copyright (2007), with permission from Elsevier. In D.-W. Sun (Ed.), *Computer Vision Technology in the Food and Beverage Industries* (pp. 3-26): Woodhead Publishing.
- Al-kaf, H. A. G., Chia, K. S., & mubin, F. R. B. A. (2020, 26-27 May 2020). *The Size And Weight Prediction For Intact Pineapples Using A Low Cost Vision System*. Paper presented at the 2020 Zooming Innovation in Consumer Technologies Conference (ZINC).
- Blasco, J., Aleixos, N., Cubero, S., & Lorente, D. (2012). 14 - Fruit, vegetable and nut quality evaluation and control using computer vision. In D.-W. Sun (Ed.), *Computer Vision Technology in the Food and Beverage Industries* (pp. 379-399): Woodhead Publishing.
- Chen, Q., Zhang, C., Zhao, J., & Ouyang, Q. (2013). Recent advances in emerging imaging techniques for non-destructive detection of food quality and safety. *TrAC Trends in Analytical Chemistry*, 52, 261-274.
- Cui, J., Yang, M., Daesik, S., & Cho, S.-I. (2017). *Machine vision and thermographic imaging for determining of grading of tomato on postharvest*.
- Demartini, M., Pinna, C., Tonelli, F., Terzi, S., Sansone, C., & Testa, C. (2018). Food industry digitalization: from challenges and trends to opportunities and solutions. *IFAC-PapersOnLine*, 51(11), 1371-1378.
- Du, Z., Zeng, X., Li, X., Ding, X., Cao, J., & Jiang, W. (2020). Recent advances in imaging techniques for bruise detection in fruits and vegetables. *Trends in Food Science & Technology*, 99, 133-141.
- Gongal, A., Amatya, S., Karkee, M., Zhang, Q., & Lewis, K. (2015). Sensors and systems for fruit detection and localization: A review. *Computers and Electronics in Agriculture*, 116, 8-19.
- Gofni, S. M., Purlis, E., & Salvadori, V. O. (2007). Three-dimensional reconstruction of irregular foodstuffs. *Journal of Food Engineering*, 82(4), 536-547.
- Haque, S., Lobaton, E., Nelson, N., Yencho, G. C., Pecota, K. V., Mierop, R., . . . Williams, C. M. (2020). Computer vision approach to characterize size and shape phenotypes of horticultural crops using high-throughput imagery. *bioRxiv*, 2020.2007.2024.199539.
- Huynh, T., Tran, L., & Dao, S. (2020). Real-Time Size and Mass Estimation of Slender Axi-Symmetric Fruit/Vegetable Using a Single Top View Image. *Sensors (Basel, Switzerland)*, 20(18), 5406.
- Jadhav, T., Singh, K., & Abhyankar, A. (2019). Volumetric estimation using 3D reconstruction method for grading of fruits. *Multimedia Tools and Applications*, 78(2), 1613-1634.
- Jana, S., Parekh, R., & Sarkar, B. (2019, 27-28 Sept. 2019). *Volume Estimation of Non-axisymmetric Fruits and Vegetables using Image Analysis*. Paper presented at the 2019 International Conference on Computing, Power and Communication Technologies (GUCON).
- Jana, S., Parekh, R., & Sarkar, B. (2020). A De novo approach for automatic volume and mass estimation of fruits and vegetables. *Optik*, 200, 163443.
- Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H., & Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research*, 2, 100033.
- Kumar, M., & Parkavi, A. (2020). Quality Grading of the Fruits and Vegetables Using Image Processing Techniques and Machine Learning: A Review. In (pp. 477-486).
- Long, Y., Wang, Y., Zhai, Z., Wu, L., Li, M., Sun, H., & Su, Q. (2018). Potato volume measurement based on RGB-D camera. *IFAC-PapersOnLine*, 51(17), 515-520.
- Lopez-Juarez, I., Rios-Cabrera, R., Hsieh, S. J., & Howarth, M. (2018). A hybrid non-invasive method for internal/external quality assessment of potatoes. *European Food Research and Technology*, 244(1), 161-174.
- Luque, A., Peralta, M. E., de las Heras, A., & Córdoba, A. (2017). State of the Industry 4.0 in the Andalusian food sector. *Procedia Manufacturing*, 13, 1199-1205.
- Mon, T., & ZarAung, N. (2020). Vision based volume estimation method for automatic mango grading system. *Biosystems Engineering*, 198, 338-349.
- Moreda, G. P., Ortiz-Cañavate, J., García-Ramos, F. J., & Ruiz-Altisent, M. (2009). Non-destructive technologies for fruit and vegetable size determination – A review. *Journal of Food Engineering*, 92(2), 119-136.
- Nyalala, I., Okinda, C., Nyalala, L., Makange, N., Chao, Q., Chao, L., . . . Chen, K. (2019). Tomato volume and mass estimation using computer vision and machine learning algorithms: Cherry tomato model. *Journal of Food Engineering*, 263, 288-298.
- Ornek, N. (2018). Determining the Carrot Volume via Radius and Length Using ANN. *International Journal of Intelligent Systems and Applications in Engineering*, 2, 165-169.
- Siswanto, J., Prabuwo, A. S., & Abdulah, A. (2013). Volume Measurement of Food Product with Irregular Shape Using Computer Vision and Monte Carlo Method: A Framework. *Procedia Technology*, 11, 764-770.
- Siswanto, J., Prabuwo, A. S., Abdullah, A., & Idrus, B. (2014). Monte Carlo Method with Heuristic Adjustment for Irregularly Shaped Food Product Volume Measurement. *The Scientific World Journal*, 2014, 683048.
- Smith, L. N., Zhang, W., Hansen, M. F., Hales, I. J., & Smith, M. L. (2018). Innovative 3D and 2D machine vision methods for analysis of plants and crops in the field. *Computers in Industry*, 97, 122-131.
- Soleimanipour, A., & Chegini, G. R. (2019). Three-dimensional reconstruction of cucumbers using a 2D computer vision system. *Journal of Food Measurement and Characterization*, 13(1), 571-578.
- Su, Q., Kondo, N., Li, M., Sun, H., Al Riza, D. F., & Habaragamuwa, H. (2018). Potato quality grading based on machine vision and 3D shape analysis. *Computers and Electronics in Agriculture*, 152, 261-268.
- Villordon, A., Gregorie, J., & LaBonte, D. (2020). Direct Measurement of Sweetpotato Surface Area and Volume Using a Low-cost 3D Scanner for Identification of Shape Features Related to Processing Product Recovery. *HortScience*, 55, 1-7.
- Vivek Venkatesh, G., Iqbal, S. M., Gopal, A., & Ganesan, D. (2015). Estimation of Volume and Mass of Axi-Symmetric Fruits Using Image Processing Technique. *International Journal of Food Properties*, 18(3), 608-626.
- Xie, W., Wang, F., & Yang, D. (2019). Research on Carrot Surface Defect Detection Methods Based on Machine Vision. *IFAC-PapersOnLine*, 52(30), 24-29.
- Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., & Liu, C. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International*, 62, 326-343

**Session 4.3 - Session spéciale 1 :
Traitement du langage naturel dans
l'industrie 4.0 et la chaîne logistique**
Chairman : Samir Lamouri - Abdourahim Sylla

¹Laboratoire de Génie Chimique, Université de Toulouse,
CNRS, INPT, UPS, Toulouse, France
4 allée Émile Monso, 31030 Toulouse, France
yohann.chasseray@ensiacet.fr

² Centre Génie Industriel, IMT Mines Albi, Université de
Toulouse, Albi, France
1 allée des sciences, 81000 Albi, France
anne-marie.barthe@mines-albi.fr

CIGI 2021

Extraction générique de connaissances à partir de données textuelles et mesure de la performance des systèmes d'extraction de relations dans un contexte non supervisé.

Résumé – Parmi les défis à venir dans le monde de l'industrie et dans le pilotage des systèmes industriels, l'agrégation, la synthèse et la gestion des connaissances au travers de structures ontologiques occupent une place primordiale. Beaucoup des systèmes d'extraction de connaissances actuels adoptent une approche supervisée, qui se base sur des données labellisées dont le processus d'annotation est long et fastidieux. Cet article présente une approche par les règles non supervisée, auto-alimentée et indépendante du domaine pour la population d'ontologie à partir de données textuelles. Par ailleurs, l'évaluation de tels systèmes, réalisant de l'extraction de connaissances par les méthodes de traitement automatique du langage, requiert l'utilisation d'indices de performance. Dans le cadre particulier de la population d'ontologie non supervisée, les indices habituellement utilisés pour réaliser ces évaluations présentent des limites dues notamment à l'absence de données annotées. Cet article propose donc également une méthode de mesure des performances dans un contexte où les données de référence et les données extraites ne se recouvrent intrinsèquement pas de manière optimale. Le mode d'évaluation proposé s'appuie sur l'exploitation de données faisant office de références mais qui ne sont pas spécifiquement liées aux données sur lesquelles est réalisée l'extraction, ce qui en fait sa particularité.

Abstract – Among the incoming challenges in the industrial domain and in the monitoring of industrial systems, the aggregation, synthesis and management of knowledge through ontological structures occupy an essential place. Existing knowledge extraction systems often use a supervised approach which rely on labelled data for which the annotation process is fastidious. This paper presents an unsupervised self-feeding rule-based approach for domain-independent ontology population from textual data. Moreover, the evaluation of such systems, performing knowledge extraction using natural language processing methods requires the use of performance indicators. The indicators usually used in such evaluations have limitations in the specific context of knowledge extraction for unsupervised ontology population. Thus, the definition of new evaluation methods becomes a need arising from the singularity of the harvested data, especially when these are unlabelled. Hence, this article also proposes a method for measuring performance in a context where reference data and extracted data do not overlap optimally. The proposed evaluation method is based on the exploitation of data that serve as a reference but are not specifically linked to the data used for extraction, which makes it an original evaluation method.

Mots clés - Traitement automatique du langage, Mesure de performance, Extraction de connaissances, Base de connaissances.

Keywords – Natural language processing, Performance evaluation, Ontologies, Knowledge extraction, Knowledge bases.

1 INTRODUCTION

La construction et l'automatisation de systèmes dans le cadre de la transformation vers l'industrie 4.0 suppose l'appui d'une structure solide basée sur la perception en temps réel du contexte interne et externe dans lequel évolue l'industrie en question. Par exemple, il est primordial pour le pilotage d'une chaîne de production d'avoir une connaissance fine, sinon un aperçu de la demande en aval et des ressources en amont, afin de faire face à tout sur-stock ou arrêt de la chaîne de production. Dans cette optique, des systèmes d'aide à la décision permettent non seulement de représenter l'état des systèmes de production mais également de produire une analyse prédictive quant aux comportements futurs du système

(prévision de la demande, maintenance prédictive). La tendance actuelle dans le développement de tels systèmes concerne l'intégration de plus en plus importante de connaissances. Par ce biais, il devient possible pour ces systèmes d'adopter un raisonnement contextualisé de plus en plus proche de celui de l'être humain et ainsi de proposer un meilleur accompagnement dans la prise de décision. L'apport de connaissances aux systèmes d'aide à la décision se matérialise par l'intégration de bases de connaissances qui sont le plus souvent le prolongement d'une ontologie développée pour le domaine au sein duquel interviennent lesdits systèmes. Les ontologies, initialement définies par (Gruber, 1993) comme la *spécification explicite d'une conceptualisation*

constituent un véritable support structuré pour la représentation de la connaissance au sein d'un domaine. Une base de connaissances peut quant à elle être définie comme la version instanciée d'une ontologie dans le cadre d'une application à un problème défini. Une base de connaissance comprend ainsi un ensemble d'instances dérivées des concepts de l'ontologie et un ensemble de relations entre ces instances, dérivées des relations définies au niveau de l'ontologie.

Malheureusement, et malgré l'adoption grandissante de l'approche par les ontologies, la plupart des systèmes développés fonctionnent avec une ontologie et, par extension, avec une base de connaissances propres au cas d'application. En effet, les ontologies et bases de connaissances utilisées sont majoritairement développées en suivant une approche dirigée par les problèmes rendant leur ré-exploitation limitée, voire impossible. Pourtant, des ontologies considérées comme génériques vis-à-vis du domaine de l'industrie 4.0 existent (Cheng et al., 2016; Giustozzi et al., 2018). Cependant, les bases de connaissances qui en découlent, lorsqu'elles existent, ne contiennent pas ou peu de connaissance, rendant leur utilisation également très limitée. Parallèlement, les données générées par l'humain, dans des formats plus ou moins structurés, renferment des éléments de connaissance d'une grande richesse. La multiplication grandissante de ces données et l'avènement de techniques d'analyse des données non-structurées sont autant d'opportunités pour l'enrichissement automatique des ontologies caractéristiques d'un domaine.

Les travaux présentés dans cet article proposent une approche générique pour la population d'ontologie en base de connaissance à partir de données hétérogènes. Cette approche, indépendante du domaine concerné par l'ontologie, se sert des concepts déjà présents dans cette dernière comme d'un support pour la recherche et l'extraction de connaissances au sein de données non structurées.

Cet article aborde, par ailleurs, différents aspects de la construction de bases de connaissances à partir d'ontologies du domaine pré-existantes et est organisé comme suit. La section 2 présente un état de l'art des méthodes utilisées pour la population d'ontologie et plus précisément pour l'extraction de relations. La section 3 introduit un cadre méthodologique générique pour la population d'ontologie indépendamment du domaine et s'intéresse de plus près à la spécification de ce cadre méthodologique pour l'extraction de connaissances à partir de données textuelles. La section 4 présente plus en détail une méthode d'extraction de relations d'hyponymie (Concept-Instance) basée sur l'utilisation de schémas d'extraction. Enfin, la section 5 constitue une prise de recul vis-à-vis de ce cadre en proposant des méthodes de mesure de la performance d'un tel système. Ces méthodes seront discutées dans la section 6.

2 ÉTAT DE L'ART

2.1 Population d'ontologie

Les problématiques de population d'ontologie ne sont pas nouvelles et plusieurs études se sont déjà intéressées à l'exploitation de données brutes pour la création d'ontologies. Beaucoup d'entre elles traitent le problème sous un angle spécifique soit relativement à la diversité des données et des formats de données (Martinez-Rodriguez et al., 2018), soit relativement au domaine d'application (Ferrara et al., 2014; Kaushik & Chatterjee, 2018; Oramas et al., 2016).

D'autres systèmes (Louge et al., 2018; Paukkeri et al., 2012) ne se rattachent pas à une ontologie existante, mais proposent

d'extraire directement depuis les données textuelles les concepts permettant d'en générer une. Ce genre d'approche permet dans certains cas de s'affranchir de la création antérieure de l'ontologie, mais ne permet pas de répliquer l'opération avec une ontologie qui serait déjà liée au système d'aide à la décision, par exemple.

D'autres outils pour la population d'ontologie font usage de techniques de machine learning et de deep learning (Faria et al., 2014). Ces outils présentent de bonnes performances même lorsqu'ils sont appliqués à des domaines disjoints. Mais, adoptant une approche supervisée, ces derniers peuvent toutefois se retrouver limités pour des domaines où la disponibilité de données annotées, utilisées à des fins d'apprentissage, est limitée. (Ayadi et al., 2019) proposent une méthode de deep learning non supervisée basée sur l'algorithme de deep-learning Word2Vec (Mikolov et al., 2013). Si cette méthode permet de s'affranchir des jeux de données labellisés pour son fonctionnement, son évaluation se fait toutefois sur des jeux de données annotés manuellement.

2.2 Extraction de relations

La détection et l'extraction de relations, au même titre que la détection d'entités, font partie des sujets les plus présents au sein des problématiques de traitement automatique du langage. On distingue dans la littérature, deux types d'approches pour la détection de relations.

La première est une approche par règles, dans laquelle il s'agit de définir un jeu de schémas d'extraction caractéristiques des relations telles qu'elles apparaissent dans les données. Ainsi, des études telles que (Agichtein & Gravano, 2000; Kim et al., 2009; Niladri Chatterjee & Neha Kaushik, 2017; Snow et al., 2005) s'intéressent à la définition et la déduction de ces schémas d'extraction.

Cette approche se révèle souvent efficace car elle cible directement les schémas représentatifs de ce qui doit être extrait et limite ainsi le taux d'erreur du système. En revanche le taux de connaissance extrait par ces schémas d'extraction reste relativement faible du fait de ce ciblage car des relations apparaissant avec un schéma légèrement différent de celui recherché ne peuvent être détectées. Par ailleurs, la mise en place des règles d'extraction peut s'avérer coûteuse et nécessite le plus souvent une expertise tant dans le domaine ciblé que relativement à la source de données étudiée. Par exemple, pour définir des schémas génériques représentatifs de relations au sein de données textuelles, une connaissance de l'expression syntaxique de ce type de relation dans la langue étudiée est nécessaire.

La seconde approche est statistique, et se base sur des indicateurs caractéristiques de la donnée étudiée. Ces mesures peuvent prendre plusieurs formes. L'une d'entre elles est le TF-IDF (Sparck Jones, 2004) qui est une mesure utilisée pour regrouper des termes au sein d'un corpus de documents dans un même champ sémantique (Paukkeri et al., 2012). Certaines études se réfèrent, pour la détection de relations spécifiques, au taux d'apparition des couples de termes au sein d'un texte (de Boer et al., 2007). Sur la base de ces analyses de co-occurrences, des algorithmes de clustering sont appliqués afin d'extraire les relations qui présentent un intérêt vis-à-vis du domaine métier (Rajpathak, 2013). Également, (Nguyen et al., 2017) utilisent une fonction objectif s'appuyant sur les fréquences d'apparition d'un terme avec son contexte pour construire une représentation statistique des relations d'hyponymie. Enfin, avec l'apparition des réseaux de neurones utilisant les mécanismes d'attention (Vaswani et al., 2017), la détection de relations par des mécanismes

d'apprentissage précis de la dimension sémantique d'un texte est également une méthode qui peut être adoptée (Geng et al., 2020; Luo et al., 2020). Celle-ci se révèle néanmoins gourmande en termes de volume de données, nécessaires à l'entraînement des modèles.

2.3 Apport de la méthode proposée

Dans les approches présentées dans la section précédente, deux verrous ont été retenus, qui concernent (1) la difficulté à coupler genericité et précision lors de la population d'ontologies et (2) l'incapacité à évaluer la performance des systèmes sur des données qui n'ont pas été labellisées manuellement au préalable.

Pour s'attaquer à ces deux verrous de l'extraction de connaissances, cet article propose donc une méthode de mesure de la performance des systèmes d'extraction pour des données non annotées, applicable au sein d'un système d'extraction générique, réalisant la population d'ontologie en s'appuyant sur la connaissance contenue dans cette dernière.

3 UN FRAMEWORK GÉNÉRIQUE POUR LA POPULATION D'ONTOLOGIE

3.1 Présentation globale du framework.

Pour réaliser l'extraction de données indépendamment de la source de données et du domaine auquel l'ontologie ciblée fait référence, un cadre méthodologique a été défini, reprenant les principes généraux de l'ingénierie dirigée par les modèles.

Ce cadre méthodologique s'articule autour d'un méta-modèle générique pour la représentation des données hétérogènes. Ce méta-modèle, présenté dans (Chasseray et al., 2021), répartit les données extraites au sein de six classes qui sont les classes : *Objet Ontologique*, *Concept*, *Instance*, *Relation*, *Donnée Extraite* et *Contexte*. Dans ce méta-modèle, la classe *Concept* permet de récupérer au sein des données analysées les éléments représentant des classes de l'ontologie à peupler. La classe *Instance* permet de récupérer dans les données les instances relatives à ces concepts, qui deviendront, par la suite, des individus de l'ontologie. Les classes *Concept* et *Instance*, héritées toutes les deux de la classe *Objet Ontologique* sont liées par la classe *Relation* qui qualifie la nature de la relation identifiée entre deux occurrences de la classe *Objet Ontologique*. Les classes *Donnée extraite* et *Contexte*, permettent quant à elles d'embarquer de l'information supplémentaire sur les concepts et instances extraits. Cela permet de capturer d'une part la donnée brute à partir de laquelle l'élément est extrait, et d'autre part des éléments de contexte qui enrichissent sémantiquement l'élément extrait.

Ces six classes permettent ainsi de structurer assez d'informations pour réaliser par la suite des transformations vers une ontologie cible. Le framework général dans lequel s'intègre ce méta-modèle a été présenté dans (Chasseray, 2020) dont est issue la figure 1. Ce dernier est constitué de deux chaînes de traitement qui, si elles peuvent fonctionner indépendamment l'une de l'autre, sont complémentaires car elles permettent de combiner à la fois l'approche sémantique et l'approche par règles. Ce framework fonctionne ainsi de manière itérative, le processus étant initié par l'approche par règles que rejoint et complète par la suite l'approche sémantique. Les deux boucles de rétroaction se nourrissent ainsi des instances extraites lors de l'étape initiale pour, d'un côté, déduire de manière automatique de nouvelles règles d'extraction, et de l'autre, dériver des matchings sémantiques.

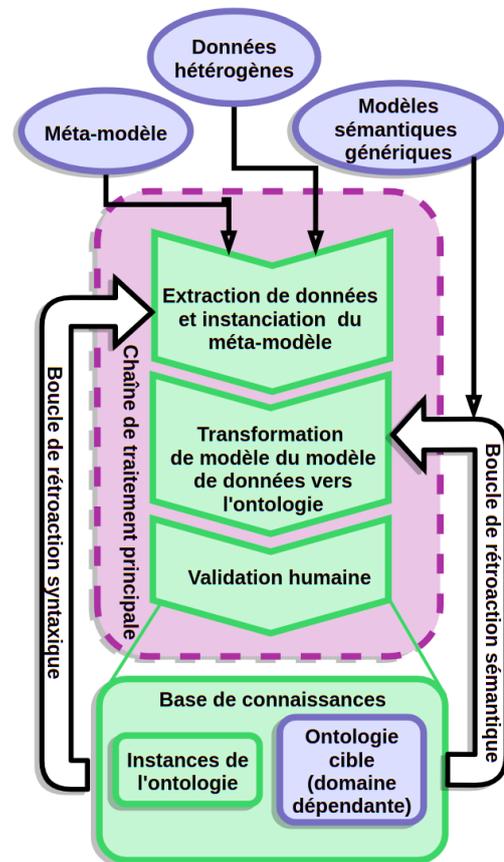


Figure 1. Framework générique pour la population d'ontologie (issu de (Chasseray, 2020)).

La section 4 de cet article se focalise sur la boucle de rétroaction qui s'appuie sur les règles. Ainsi, le processus initial d'extraction et le processus de déduction de nouvelles règles d'extraction, utilisant respectivement les schémas d'extraction proposés par (Hearst, 1992) et le principe de bootstrapping introduit par (Herbelot & Copestake, 2006) et (Pennacchiotti & Pantel, 2006) sont présentés en section 4.

3.2 Spécification du framework pour le traitement de données textuelles.

Le framework générique présenté dans la section 3.1 peut être spécifié pour le traitement de données textuelles grâce à l'utilisation des techniques de traitement automatique du langage. Dans cette section, on présente succinctement les méthodes de traitement automatique du langage qui sont utilisées afin de déduire un modèle de données et nourrir les boucles de rétroaction du framework lors de cette spécification. Ainsi sur la figure 2, on peut distinguer une chaîne de traitement commune qui donne lieu par la suite à trois chaînes de traitement spécifiques. La chaîne de traitement commune contient des opérations de traitement classique de tokenisation et de tagging. Sur la base de ce pré-traitement, les chaînes de traitement différenciées (1), (2), et (3) sont appliquées dans un but qui leur est propre. La chaîne de traitement notée (1) sur la figure 2 a pour objectif d'extraire des couples d'entités liées par une relation. Cette relation peut être de différentes natures (hyponymie, hyperonymie, relation définie par l'ontologie). Ces couples seront par la suite utilisés pour instancier le modèle de données et enfin être rattachés à l'ontologie. Cette chaîne de traitement contient trois étapes. La première permet d'identifier et de labelliser des concepts qui apparaissent dans le texte sur la base des concepts contenus dans l'ontologie. La deuxième

étape consiste à dresser l'arbre des dépendances syntaxiques qui existent entre les tokens du texte. Sur la base des concepts identifiés et de cet arbre des dépendances syntaxiques, des schémas d'extraction sont appliqués au cours de la troisième étape du traitement. Cette chaîne de traitement intervient dans l'étape d'initialisation du système évoquée dans les sections précédentes.

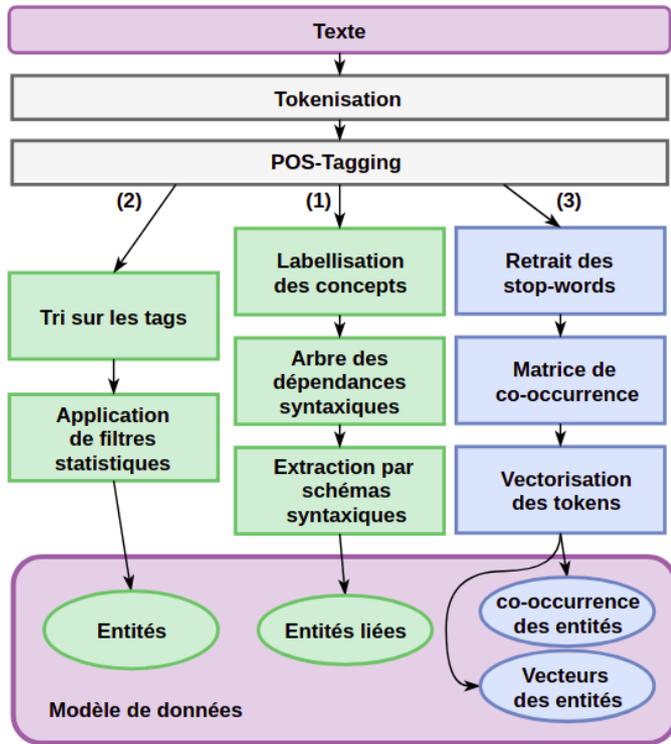


Figure 2. Chaînes de traitement NLP pour la création d'un modèle de données à partir de données textuelles

La chaîne d'extraction notée (2) sur la figure 2 permet d'extraire d'autres entités, cette fois non liées entre elles. L'extraction se fait selon deux filtres. Le premier filtre porte sur les POS-tags représentatifs des termes pouvant s'apparenter à des instances. Un deuxième filtre, statistique, sélectionne les entités les plus susceptibles d'être des instances liées à des concepts du domaine, définis par l'ontologie. Bien que la recherche concerne des instances du domaine, cette chaîne d'extraction n'est pas guidée par l'ontologie. Ainsi, les entités qui en sont issues ne seront pas directement reliées à l'ontologie mais devront subir une étape de matching sémantique et viendront donc alimenter la boucle de rétroaction sémantique.

Enfin, la troisième chaîne de traitement spécifique, notée (3) sur la figure 2, a un objectif légèrement différent des deux premières puisqu'elle est destinée à extraire l'ensemble des éléments de contexte (co-occurrences, vecteurs) qui permettent de caractériser les entités précédemment extraites. Pour ce faire, le traitement s'effectue en trois étapes. La première étape consiste à éliminer les mots vides (stop-words) afin de se concentrer sur les termes porteurs de sens. La seconde étape permet de construire la matrice de co-occurrence qui peut être utilisée pour lier les termes qui co-occurrent avec une entité extraite ou un concept de l'ontologie par exemple. Enfin, la troisième étape utilise un modèle du langage pour obtenir une représentation vectorielle des entités extraites. Ces éléments de contexte ont pour objectif d'être réutilisés par la suite pour

identifier des instances parmi les entités non liées dans la boucle de rétroaction sémantique.

Dans cet article, il a été choisi de détailler les méthodes d'extraction et de déduction utilisées pour le fonctionnement de la boucle de rétroaction basée sur les règles. Cette boucle s'appuie sur la chaîne de traitement (1). Une application de l'extraction par schémas syntaxiques sur des données textuelles est ainsi présentée en section 4.1. Cette dernière permet notamment de détailler la manière de procéder à l'extraction à partir de concepts extraits d'une ontologie.

4 APPROCHE PAR SCHÉMAS D'EXTRACTION (APPROCHE PAR LES RÈGLES)

Dans de nombreux cas d'application, il n'existe pas ou très peu de données déjà disponibles qui permettraient d'adopter une approche supervisée pour l'extraction de connaissances. Ainsi, un système de population d'ontologie générique doit être en mesure de fonctionner dans un contexte non supervisé, c'est-à-dire sans connaissance a priori sur les entités qui doivent être extraites. De plus, ce système doit se révéler adaptable à différents domaines. L'approche par règles avec l'utilisation de schémas d'extraction génériques prend alors tout son sens.

Dans cette section, la définition de schémas d'extraction générique pour l'extraction à partir de données textuelles est explicitée et le fonctionnement de la boucle de rétroaction basée sur les règles, incluant la déduction de nouveaux schémas est illustré.

4.1 Initialisation par schémas d'extraction génériques

Comme évoqué dans la section 3, la boucle de rétroaction basée sur les règles, et par extension le système global, contient une étape d'initialisation pour l'extraction des premières instances. Cette étape est primordiale car elle conditionne l'apprentissage futur de nouveaux schémas d'extraction et doit également garantir la performance de la boucle de rétroaction sémantique.

La figure 3 illustre un schéma d'extraction syntaxique se basant sur les schémas de (Hearst, 1992). Ce schéma représente la relation d'hyponymie qui peut apparaître dans le texte entre un concept et son instance. Le schéma ainsi construit est en réalité la superposition de trois séquences d'extraction.

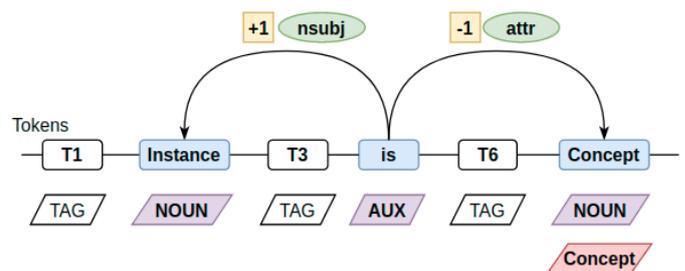


Figure 3. Exemple de schéma générique d'extraction pour les relations d'hyponymie.

La première séquence est relative aux POS-tags et décrit l'enchaînement de tags recherché. Dans l'exemple illustré par la figure 3, la séquence recherchée est la suivante : $(Concept) \rightarrow (AUX) \rightarrow (NOUN) OR (PROPN) OR (PROP)$.

La deuxième séquence est relative aux tags affectés aux dépendances dans l'arbre des dépendances syntaxiques et décrit donc l'enchaînement des dépendances syntaxiques recherché. Dans cet exemple, la séquence recherchée est la suivante : $(attr) \rightarrow (nsubj)$.

Une troisième séquence permet de préciser le sens de navigation dans l'arbre des dépendances syntaxiques, c'est-à-dire d'indiquer si le schéma remonte (-1) ou descend (+1) l'arbre des dépendances syntaxiques. Il est important de préciser que l'ordre de parcours du schéma d'extraction ne respecte pas, a priori, l'ordre de lecture du texte étudié. Dans l'exemple proposé, la séquence utilisée pour préciser la navigation est la suivante : (-1)->(+1).

Au-delà de l'utilisation des dépendances syntaxiques pour plus de généralité, cette manière de définir des schémas d'extraction contient une particularité. En effet, si la séquence de tags utilisée est majoritairement constituée de POS-tags classiques, un tag supplémentaire est utilisé pour indiquer que le token représente un concept de l'ontologie. C'est ce tag qui déclenche la recherche d'une relation à partir de ce concept en suivant le schéma générique prédéfini. Le procédé de labellisation des concepts sur la base des classes de l'ontologie permettant d'assurer l'identification de ces concepts est effectué en amont, dans la chaîne de traitement automatique du langage (voir figure 2).

Si le schéma peut être respecté par les données du texte, alors une ou plusieurs instances peuvent être détectées. Dans l'exemple donné, un concept et une instance seront systématiquement détectés aux deux extrémités du schéma d'extraction, donnant lieu à la création de deux objets ontologiques liées par une relation d'hyponymie dans le modèle de données.

4.2 Mécanisme de déduction de schémas syntaxiques.

L'extraction par schémas génériques ou spécifiques est une méthode qui existe déjà depuis plusieurs années. La figure 4 donne une illustration globale du fonctionnement de la boucle de rétroaction basée sur les règles au sein du framework d'extraction. Cette figure présente le principe de bootstrapping mis en avant par (Pennacchiotti & Pantel, 2006).

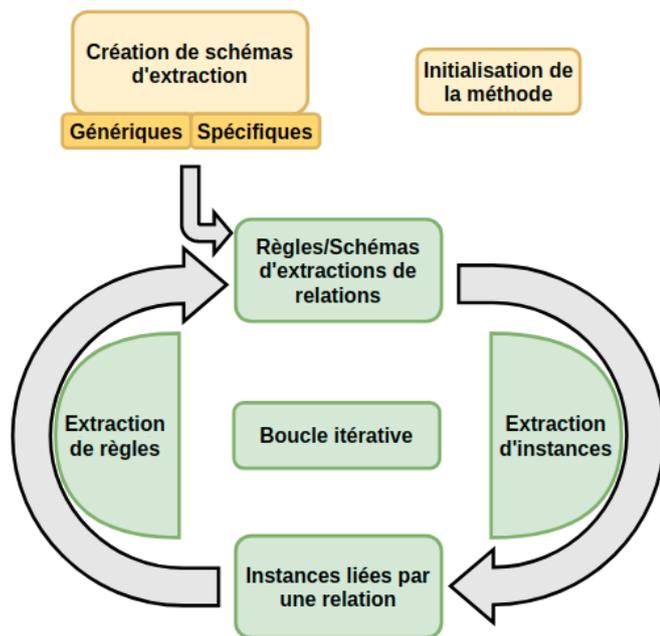


Figure 4. Illustration de l'approche bootstrapping pour la déduction de schémas d'extraction de relations.

Cette méthode consiste en l'auto-alimentation des étapes de déduction des schémas et d'application des schémas d'extraction. La déduction de schémas se fait en inversant le processus d'extraction. Ainsi, lorsque des instances connues apparaissent liées à leur concept dans les données, le principe

de bootstrapping consiste à déduire un schéma qui lie ces deux instances. Si ce schéma est représentatif de la relation exprimée entre deux instances ou entre une instance et son concept associé alors celui-ci peut être ajouté à l'ensemble des schémas. Il est ensuite utilisé pour détecter de nouvelles relations dans les données. Cette méthode suppose donc bel et bien l'existence préalable de couples concept-instance identifiés, d'où l'importance de l'étape d'initialisation et de l'utilisation de schémas génériques prédéfinis par l'humain.

Dans le cadre du traitement de données textuelles, le schéma présenté dans cet article est un schéma syntaxique, appliqué sur l'arbre des dépendances syntaxiques préalablement construit. Toutefois, cette approche peut être étendue à d'autres types de schémas, applicables à d'autres types de données (structure, enchaînement de balises).

5 MÉTHODE DE MESURE DE LA PERFORMANCE DU SYSTÈME D'EXTRACTION

Le manque de données annotées comme frein à l'adoption d'une approche supervisée a été évoqué en introduction de la section 4. Cette problématique affecte également les démarches qui touchent à l'évaluation d'un système d'extraction. En effet, dans des tâches de traitement automatique de données, la performance d'une méthode est de manière générale évaluée par application de cette dernière sur un jeu de données de référence (communément identifié sous le terme de jeu de test). C'est le cas par exemple dans la plupart des tâches de machine learning qui font appel à des méthodes d'évaluation telles que le calcul du F1-Score et le tracé des courbes ROC (Fawcett, 2006). Malheureusement, dans un contexte où les données annotées ne sont pas répandues, et avec des méthodes non supervisées telles que celle présentée dans ce papier, ces méthodes sont limitées car :

- Elles ne permettent pas d'évaluer la performance d'un système lorsque celui-ci est appliqué à de nouveaux jeux de données.
- Elles ne tiennent pas compte des connaissances nouvelles, éventuellement extraites par le système mais ignorées au cours de l'annotation.

L'apport de la proposition faite dans cet article, est de traiter ces deux aspects en définissant une mesure de performance ne s'appliquant pas strictement aux extractions faites sur un jeu de test donné et défini à l'avance.

On supposera donc que les données que l'on traite ne sont pas, ou très rarement, annotées. Deux solutions peuvent alors être envisagées. La première consiste à labelliser manuellement les données une fois l'extraction réalisée. Cette méthode demande un investissement fort d'experts métiers spécialisés dans le domaine de l'ontologie, mais permet en contrepartie, de déterminer la précision du système de manière rigoureuse. La deuxième s'appuie sur d'éventuelles données de référence existantes indépendamment du processus de population. Cette deuxième méthode présente l'avantage de ne pas faire intervenir l'expert du domaine pour effectuer une validation manuelle. Cependant, si elle est rapide à mettre en place, elle ne permet pas de déterminer avec rigueur la précision du modèle. Cette section se concentre sur la deuxième méthode, appliquée dans le cadre de relations concept-instance et qui demande une redéfinition des mesures de performance. Pour aborder ce problème, il convient de redéfinir les notions de *concept*, d'*instance* et de *couple*. Un concept est ici la représentation dans les données (référence ou extraites) d'une des classes initialement contenues dans l'ontologie.

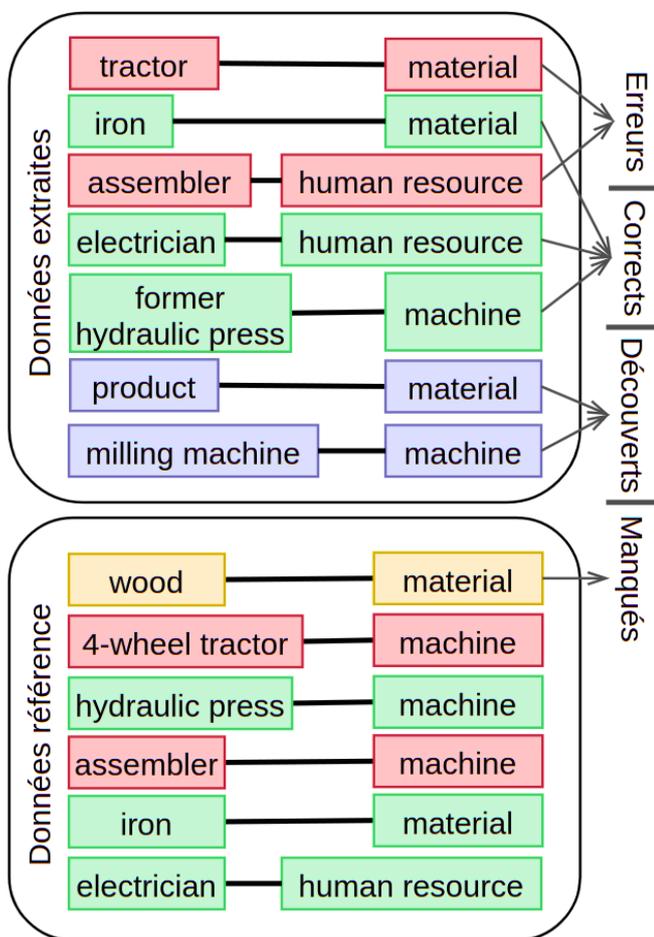


Figure 5. Représentation des jeux de données extrait et référence exemples.

Une instance (reliée à un concept) est la représentation dans les données de l'instance d'une classe de l'ontologie, qui a été extraite ou qui fait partie des données de référence. Enfin un couple est la réunion d'un concept et d'une instance (qu'il soit issu de l'extraction ou des données de référence) reliés par une relation. Pour illustrer à titre d'exemple la mesure, celle-ci sera appliquée aux jeux de couples extraites et référence simulés présentés dans la figure 5. Il est important de préciser que ces jeux de couples sont utilisés à titre d'exemple, et ne sont pas liés à un jeu de données en particulier. En revanche, des mesures de performance utilisant la méthode ont été réalisées sur un jeu de données réel. Ces résultats sont présentés en section 5.4.

5.1 Réconciliation des données extraites et référence par le ROUGE Score.

L'évaluation du système doit donc pouvoir se faire à partir de données de référence qui ne sont pas nécessairement liées au texte étudié, mais qui constituent un bon exemple de la base de connaissance qui est visée à travers la population de l'ontologie du domaine ciblé.

La problématique induite par ce contexte, est la non correspondance exacte des données extraites aux données de référence, ce qui ne signifie pas pour autant une erreur de la part du système d'extraction. En effet, dans la mesure où sont comparées des chaînes de caractères (représentatives des instances extraites), il peut arriver que certaines d'entre elles diffèrent sans pour autant que leur sens ne s'en voit modifié. Afin d'éviter les défauts d'appréciation dus à cette spécificité, il est possible de définir une distance entre les chaînes de caractères et d'y appliquer un seuil.

De nombreuses métriques existent pour faire le calcul de la distance entre deux chaînes de caractères. Parmi elles, la distance de Levenshtein (Levenshtein, 1966) ou encore la distance de Hamming (Hamming, 1950) font référence. Cependant, ces distances sont généralement utilisées sur des chaînes de caractères soit déjà fortement similaires, soit très longues, ce qui n'est pas nécessairement le cas ici. De plus, ces mesures peuvent se révéler très inefficaces pour la comparaison de chaînes de caractères proches d'un point de vue purement basé sur les caractères mais ayant une sémantique totalement différente. Par exemple, les chaînes de caractères *iron* et *wood* ont une distance de Levenshtein relativement faible (6) mais représentent cependant des instances différentes.

L'évaluation du ROUGE Score (Lin, 2004) entre deux chaînes de caractères a donc été retenue pour caractériser l'égalité des instances qu'elles représentent. Le ROUGE Score est un indice de similarité simple qui adopte une méthode de mesure semblable à la distance de Levenshtein à la différence près que ce dernier est adapté à l'échelle des termes (et non pas à l'échelle du caractère). De cette manière, les chaînes de caractères *iron* et *wood* apparaissent bien comme plus éloignées (ROUGE Score = 0) que les chaînes de caractères *hydraulic press* et *former hydraulic press* par exemple (ROUGE Score = 0.79), ce qui n'est pas le cas si les calculs de distance sont effectués à l'échelle du caractère comme dans le cadre de mesures de distance classiques ($Levenshtein(wood, iron) = 6$, $Levenshtein(hydraulic\ press, former\ hydraulic\ press) = 7$). Couplée à un seuil d'acceptation, la valeur du ROUGE Score fait donc office de validateur de l'égalité entre deux instances représentées par deux chaînes de caractères.

5.2 Construction des ensembles Corrects, Manqués, Découverts, et Erreurs

La méthode de comparaison entre les données extraites et les données de référence établie, il est possible de catégoriser les instances extraites selon que celles-ci ont été affectées ou non au concept leur correspondant dans les données de référence.

Comme les données de référence ne sont pas nécessairement liées au texte utilisé pour l'extraction, il peut cependant arriver que certaines instances des données de référence ne soient pas retrouvées dans les données extraites et inversement, que certaines données extraites ne soient pas représentées dans les données de référence. Ainsi quatre ensembles sont définis dans cette section afin de traduire cette spécificité :

- **Corrects (C) :** L'ensemble *Corrects* contient les couples dont les instances sont considérées identiques dans le jeu de données référence et dans le jeu de données extrait et pour lesquelles le concept associé est conforme au concept du jeu de référence.
- **Erreurs (E) :** L'ensemble *Erreurs* contient les couples extraits dont l'instance apparaît bien dans le jeu annoté mais est associée à un concept différent de celui du jeu annoté, soit par erreur, soit par sur-classification.
- **Manqués (M) :** L'ensemble *Manqués* contient les couples du jeu de données de référence qui n'apparaissent pas dans le jeu de données extrait, parce que l'instance n'a pas été remontée du tout dans les données extraites.
- **Découverts (D) :** L'ensemble *Découverts* contient les couples identifiés par le système mais dont l'instance n'apparaît pas dans le jeu de données annoté.

Le critère d'égalité entre deux instances est déclaré lorsque la distance entre ces deux instances passe un seuil défini par l'évaluateur. La distance utilisée s'appuie quant à elle sur le ROUGE Score. Pour l'exemple, la valeur seuil utilisée est fixée à 0.6. Les ensembles E et C contiennent des couples extraits dont les instances sont égales au sens de la distance définie. L'ensemble D contient des couples extraits dont l'instance ne présente de similarité suffisante avec aucune des instances de référence pour être classée dans l'un des ensembles C ou E .

L'ensemble M , à l'inverse, contient des couples de référence qui n'apparaissent pas dans les données extraites et ne peuvent être classés ni dans C , ni dans E , ni dans D . Le fait de ne pas inclure dans M les couples dont l'instance apparaît, dans les couples extraits, associée à un concept différent est un choix fait pour distinguer les couples manqués des couples mal définis (erreur sur le concept). Cette décision induit un biais car elle ne prend pas en compte les cas particuliers où une instance a été reliée par le système à seulement une partie des concepts présents dans les couples impliquant l'instance dans les données de référence. Par souci de clarté, ce cas de figure n'est pas présenté dans l'exemple.

5.3 Estimation de la précision et du rappel

Les ensembles M et D présentés dans la section 5.2 sont singuliers car ils concernent des couples dont les instances ne sont présentes que dans un seul des deux jeux de données.

Des couples peuvent ainsi apparaître dans l'ensemble M soit parce que l'instance associée au concept n'apparaît pas dans les données et n'a donc pas pu être extraite, soit parce que celle-ci n'a pas été extraite, malgré sa présence dans le texte. Dans le premier cas, le couple ne peut pas être estimé comme un réel manquement du système. Dans le deuxième, il s'agit bel et bien d'une instance qui n'a pas été extraite.

De la même manière, il est compliqué d'estimer la valeur des couples contenus dans l'ensemble D qui peuvent être soit de nouveaux couples pertinents, qui n'étaient pas listés dans les données de référence, soit un ensemble d'erreurs de la part du système (faux positifs).

En l'absence d'annotation manuelle, il devient alors difficile d'établir exactement les performances du système sans formuler d'hypothèses sur la nature des couples contenus dans les ensembles M et D . Il est toutefois possible d'en donner une estimation en supposant que les couples de l'ensemble M sont bel et bien des couples manqués et que les couples de l'ensemble D , comme ils ne font pas partie du jeu de référence, ne constituent pas une connaissance suffisante pour être considérés comme des vrais positifs. En suivant ces hypothèses, on peut définir génériquement les sous-ensembles Ens_{c_i} et \overline{Ens}_{c_i} :

$$Ens_{c_i} = \{cpl = (con, ins) \in Ens \mid con = c_i\}$$

$$\overline{Ens}_{c_i} = \{cpl = (con, ins) \notin Ens \mid con = c_i\}$$

et ainsi redéfinir la matrice de confusion. Les vrais positifs sont alors estimés à partir de l'ensemble C_{c_i} , les faux positifs sont estimés à partir des ensembles E_{c_i} et D_{c_i} , les vrais négatifs à partir des ensembles \overline{E}_{c_i} , \overline{D}_{c_i} et \overline{C}_{c_i} , et les faux négatifs à partir de l'ensemble M_{c_i} :

$$VP_{c_i} = \sum_{cpl \in C_{c_i}} sim_{cpl}$$

$$FP_{c_i} = \sum_{cpl \in E_{c_i} \cup D_{c_i}} sim_{cpl}$$

$$VN_{c_i} = \sum_{cpl \in \overline{E}_{c_i} \cup \overline{C}_{c_i} \cup \overline{M}_{c_i} \cup \overline{D}_{c_i}} sim_{cpl}$$

$$FN_{c_i} = \sum_{cpl \in M_{c_i}} sim_{cpl}$$

Où sim_{cpl} vaut 1 pour les couples des ensembles D et M et correspond à l'indice de similarité calculé à partir du ROUGE Score entre un couple extrait et son homologue au sein des données de référence pour les couples des ensembles C et E .

En effet, pour les couples appartenant aux ensembles C et E , le critère d'égalité a été déclaré à partir d'un calcul de similarité. Pour prendre en compte cette similarité, plutôt que d'ajouter une unité aux groupes de faux positifs ou vrais positifs, c'est la valeur du ROUGE Score qui est utilisée. Cette mesure étant comprise entre 0.6 (valeur seuil) et 1, elle constitue un marqueur du degré de similarité entre les deux couples. Ainsi, un couple considéré comme appartenant à l'ensemble C car il a juste dépassé le seuil d'admission fixé aura moins de poids qu'un couple pour lequel la valeur du ROUGE Score vaut 1 car ce dernier est parfaitement identique au couple de référence. Pour l'exemple présenté, on obtient par cette méthode les valeurs détaillées dans le tableau 1 pour chacun des concepts de la figure 5.

Tableau 1. Calcul des matrices de confusion et des performances après définition de E , C , M et D .

	human resource	machine	material
VP	1	0.79	1
FP	1	1	1.66
VN	5.45	5.66	3.79
FN	0	0	1
Précision	0.50	0.56	0.38
Rappel	1	1.00	0.74
F1-Score	0.67	0.72	0.50

Le F1-Score global peut quant à lui être calculé en pondérant les F1-Scores de chaque concept par la participation (en termes de similarité), de chacun des couples dans lesquels il intervient. Dans cet exemple simple, on obtient un F1-Score global égal à 0.6.

5.4 Application de la méthode à un jeu de données traitant de la biochimie.

La méthode illustrée dans les sections précédentes a été appliquée en utilisant un jeu de données référence lié au domaine de la biochimie (abstracts annotés) et en réalisant une extraction sur des articles de recherche associés à ce jeu de données référence (Shardlow et al., 2018). Il est important de noter ici que les données de référence et les données sur lesquelles est appliquée l'extraction sont distinctes, mettant en avant l'intérêt de la méthode proposée.

Le tableau 2 indique le volume (somme des similarités) des

ensembles C , E , M et D pour un seuil d'égalité fixé à 0.5. On obtient pour ce jeu de données une valeur de précision de 0.65. Il s'agit néanmoins, du fait de la méthode, et pour les raisons évoquées dans la section 5.3, d'une évaluation pessimiste de la précision. La valeur du rappel est quant à elle très faible (environ 5%) dans la mesure où de nombreux couples sont considérés comme manqués par l'extraction. Encore une fois, cette estimation stricte est discutée dans la section 5.3.

Tableau 2 : Volume (somme des similarités) des différents ensembles, après extraction sur des données liées au domaine de la biochimie (Shardlow et al., 2018).

<i>Corrects (C)</i>	142.27
<i>Erreurs (E)</i>	10.83
<i>Manqués (M)</i>	2503
<i>Découverts (D)</i>	78

5.5 Discussion sur la méthode d'évaluation des performances.

Il est important de souligner que la définition des vrais positifs, faux positifs, vrais négatifs et faux négatifs adopte une vision pessimiste de l'extraction pour les raisons évoquées au début de la section 5.3. Le pendant optimiste consisterait à considérer que :

- Les couples contenus dans M ne peuvent pas être assimilés à des faux négatifs dans la mesure où l'on suppose que ceux-ci n'apparaissent pas dans les données étudiées.
- Les couples contenus dans D ne sont plus assimilés à des faux positifs mais à des vrais positifs.

Néanmoins, cette vision conduit à un rappel faussement maximal (égal à 1), et à une précision dopée, qui ne seront que très rarement représentatifs des performances réelles de l'extraction. Une meilleure estimation de la précision peut toutefois être envisagée en ne traitant pas le cas des couples de l'ensemble D et en se limitant aux ensembles C et E pour définir vrais positifs et faux positifs.

Par ailleurs, on omet dans la définition de M les couples non détectés mais dont l'instance a été détectée dans un autre couple ce qui éloigne la mesure du rappel de la réalité. Une autre définition de M , incluant ces couples, reviendrait à considérer dans certains cas deux fois un même résultat du système, d'abord comme une erreur du point de vue des données extraites, puis comme un manquement du point de vue des données de référence.

Enfin, une vérification automatique de la présence des instances des couples de M dans le texte sur lequel l'extraction a été réalisée est une solution qui permet de mieux estimer la part de réels faux négatifs dans l'ensemble M . Ainsi seuls les couples dont les instances sont présentes dans le texte traité peuvent être considérées dans la définition des faux négatifs.

6 CONCLUSION

Nous avons présenté dans cet article des méthodes d'évaluation de la performance pour le cas particulier de l'extraction de relations d'hyponymie entre concept et instance au sein de données textuelles non annotées. La méthode proposée permet d'estimer la performance d'un tel modèle d'extraction à partir de données de référence qui ne sont pas nécessairement liées ou construites à partir de la source de

données dont sont extraites les relations à évaluer.

Cette définition de nouvelles méthodes d'évaluation des performances s'inscrit dans une démarche plus globale de population d'ontologie dont a été présentée la stratégie générale, et détaillée l'une des méthodes permettant de mener à bien l'extraction de connaissances à partir de données textuelles.

Néanmoins, la méthode d'évaluation définie repose en grande partie sur la mesure d'égalité qui a également ses propres limites. Le ROUGE Score ne prend pas par exemple en compte la dimension sémantique entre deux instances. Cette méthode de mesure des performances doit ainsi faire l'objet de tests plus poussés, sur des jeux de données représentatifs afin d'éprouver sa pertinence. L'étude d'autres mesures de distances, couplées au ROUGE-Score, pour définir l'égalité entre deux couples est également envisagée.

7 RÉFÉRENCES

- Agichtein, E., & Gravano, L. (2000). *Snowball*: Extracting relations from large plain-text collections. *Proceedings of the fifth ACM conference on Digital libraries*, 85-94. <https://doi.org/10.1145/336597.336644>
- Ayadi, A., Samet, A., de Beuvron, F. de B., & Zanni-Merk, C. (2019). Ontology population with deep learning-based NLP: A case study on the Biomolecular Network Ontology. *Procedia Computer Science*, 159, 572-581. <https://doi.org/10.1016/j.procs.2019.09.212>
- Chasseray, Y. (2020). Un meta-modèle pour la population d'ontologie indépendamment du domaine. *Forum Jeunes Chercheuses Jeunes Chercheurs*, 17-20.
- Chasseray, Y., Barthe-Delanoë, A.-M., Négny, S., & Le Lann, J.-M. (2021). A generic metamodel for data extraction and generic ontology population. *Journal of Information Science*, 0165551521989641. <https://doi.org/10.1177/0165551521989641>
- Cheng, H., Zeng, P., Xue, L., Shi, Z., Wang, P., & Yu, H. (2016). Manufacturing Ontology Development Based on Industry 4.0 Demonstration Production Line. *2016 Third International Conference on Trustworthy Systems and Their Applications (TSA)*, 42-47. <https://doi.org/10.1109/TSA.2016.17>
- de Boer, V., van Someren, M., & Wielinga, B. J. (2007). A redundancy-based method for the extraction of relation instances from the Web. *International Journal of Human-Computer Studies*, 65(9), 816-831. <https://doi.org/10.1016/j.ijhcs.2007.05.002>
- Faria, C., Serra, I., & Girardi, R. (2014). A domain-independent process for automatic ontology population from text. *Science of Computer Programming*, 95, 26-43. <https://doi.org/10.1016/j.scico.2013.12.005>
- Fawcett, T. (2006). Introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Ferrara, E., De Meo, P., Fiumara, G., & Baumgartner, R. (2014). Web data extraction, applications and techniques: A survey. *Knowledge-Based Systems*, 70, 301-323. <https://doi.org/10.1016/j.knosys.2014.07.007>
- Geng, Z., Chen, G., Han, Y., Lu, G., & Li, F. (2020). Semantic relation extraction using sequential and tree-structured LSTM with attention. *Information*

- Sciences, 509, 183-192.
<https://doi.org/10.1016/j.ins.2019.09.006>
- Giustozzi, F., Saunier, J., & Zanni-Merk, C. (2018). Context Modeling for Industry 4.0 : An Ontology-Based Proposal. *Procedia Computer Science*, 126, 675-684.
<https://doi.org/10.1016/j.procs.2018.08.001>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220.
<https://doi.org/10.1006/knac.1993.1008>
- Hamming, R. W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 29(2), 147-160.
<https://doi.org/10.1002/j.1538-7305.1950.tb00463.x>
- Hearst, M. A. (1992). Automatic Acquisition of Hyponyms from Large Text Corpora. *COLING*.
- Herbelot, A., & Copestake, A. (2006). *Acquiring ontological relationships from Wikipedia using RMRS*.
- Kaushik, N., & Chatterjee, N. (2018). Automatic relationship extraction from agricultural text for ontology construction. *Information Processing in Agriculture*, 5(1), 60-73.
<https://doi.org/10.1016/j.inpa.2017.11.003>
- Kim, S.-J., Lee, Y.-H., & Lee, J.-H. (2009). *Method of Extracting Is-A and Part-Of Relations Using Pattern Pairs in Mass Corpus*. 9.
- Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, 10, 707.
<http://adsabs.harvard.edu/abs/1966SPHD...10..707L>
- Lin, C.-Y. (2004). *ROUGE: A Package for Automatic Evaluation of Summaries*. 8.
- Louge, T., Karray, M.-H., & Archimède, B. (2018). Investigating a Method for Automatic Construction and Population of Ontologies for Services : Performances and Limitations. *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, 1-6.
<https://doi.org/10.1109/AICCSA.2018.8612844>
- Luo, L., Yang, Z., Cao, M., Wang, L., Zhang, Y., & Lin, H. (2020). A neural network-based joint learning approach for biomedical entity and relation extraction from biomedical literature. *Journal of Biomedical Informatics*, 103, 103384.
<https://doi.org/10.1016/j.jbi.2020.103384>
- Martinez-Rodriguez, J. L., Lopez-Arevalo, I., & Rios-Alvarado, A. B. (2018). OpenIE-based approach for Knowledge Graph construction from text. *Expert Systems with Applications*, 113, 339-355. <https://doi.org/10.1016/j.eswa.2018.07.017>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality*. 9.
- Nguyen, K. A., Köper, M., Walde, S. S. im, & Vu, N. T. (2017). Hierarchical Embeddings for Hypernym Detection and Directionality. *arXiv:1707.07273 [cs]*. <http://arxiv.org/abs/1707.07273>
- Niladri Chatterjee & Neha Kaushik. (2017). RENT : Regular Expression and NLP-Based Term Extraction Scheme for Agricultural Domain. In S. C. Satapathy, V. Bhateja, & A. Joshi (Éds.), *Proceedings of the International Conference on Data Engineering and Communication Technology* (p. 511-522). Springer Singapore.
- Oramas, S., Espinosa-Anke, L., Sordo, M., Saggion, H., & Serra, X. (2016). Information extraction for knowledge base construction in the music domain. *Data & Knowledge Engineering*, 106, 70-83.
<https://doi.org/10.1016/j.datak.2016.06.001>
- Paukkeri, M.-S., García-Plaza, A. P., Fresno, V., Unanue, R. M., & Honkela, T. (2012). Learning a taxonomy from a set of text documents. *Applied Soft Computing*, 12(3), 1138-1148.
<https://doi.org/10.1016/j.asoc.2011.11.009>
- Pennacchiotti, M., & Pantel, P. (2006). *A Bootstrapping Algorithm for Automatically Harvesting Semantic Relations*.
- Rajpathak, D. G. (2013). An ontology based text mining system for knowledge discovery from the diagnosis data in the automotive domain. *Computers in Industry*, 64(5), 565-580.
<https://doi.org/10.1016/j.compind.2013.03.001>
- Shardlow, M. J., Nguyen, N., Owen, G., O'Donovan, C., Leach, A., McNaught, J., Turner, S., & Ananiadou, S. (2018). A New Corpus to Support Text Mining for the Curation of Metabolites in the ChEBI Database. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 280-285.
<http://www.lrec-conf.org/proceedings/lrec2018/summaries/229.html>
- Snow, R., Jurafsky, D., & Ng, A. Y. (2005). *Learning Syntactic Patterns for Automatic Hypernym Discovery*. 8.
- Sparck Jones, K. (2004). A Statistical Interpretation of Term Specificity in Retrieval. *Journal of Documentation*, 60, 493-502.
<https://doi.org/10.1108/00220410410560573>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *arXiv:1706.03762 [cs]*. <http://arxiv.org/abs/1706.03762>

L'analyse de sentiments pour la prévision de la demande en temps de crise : le cas de la pandémie de COVID-19

ANGIE NGUYEN¹, SAMIR LAMOURI¹, Robert PELLERIN²

¹ LAMIH CNRS, Arts et Métiers ParisTech
151 boulevard de l'Hôpital, Paris, France
{angie.nguyen ; samir.lamouri} @ensam.eu

² CIRRELT, IVADO, Polytechnique Montréal
500 chemin de Polytechnique, Montréal, Canada
robert.pellerin@polymtl.ca

Résumé – Un événement non souhaité tel qu'une catastrophe naturelle ou une épidémie s'accompagne généralement de fortes perturbations des chaînes logistiques et d'une grande volatilité des marchés. Par ailleurs, la récente pandémie du COVID-19 a montré que les systèmes d'intelligence artificielle et les modèles d'apprentissage automatique existants, performants en temps régulier, n'ont pas été capables de fournir des prévisions fiables de la demande fortement volatile. L'objectif de cet article est de déterminer si l'analyse de sentiments des médias d'information peut permettre de fournir des prévisions fiables de la demande en période de crise. Un cas d'application de la consommation de médicament en France pendant la crise du COVID-19 est traité en trois étapes consistant à : (1) collecter et préparer les données ; (2) développer un modèle d'analyse de sentiments basé sur de l'apprentissage automatique supervisé ; et (3) analyser la corrélation entre les sentiments véhiculés par les médias et la variation de consommation de médicament. Les résultats soulignent une forte corrélation positive entre les sentiments dans les actualités et la consommation. Ils suggèrent donc que les sentiments véhiculés par les médias ont un fort pouvoir prédictif sur les dynamiques de consommation en période de crise.

Abstract – Crisis situations resulting from natural disasters or epidemic outbreaks are usually accompanied by supply chain disruption and highly volatile markets. Besides, the recent COVID-19 pandemic has shown that existing artificial intelligence systems and data analytics models, which normally provide valuable support in demand forecasting, have not been able to manage demand volatility. This study contributes addressing this issue and aims to determine whether sentiments conveyed by news media influence consumer behavior. It provides a case study conducted in three steps: (1) data were collected and prepared; (2) a sentiment analysis model was developed; and (3) a statistical analysis was performed to analyze the correlation between sentiments in news and drug consumption during the COVID-19 crisis. Findings highlighted a strong positive correlation between sentiments in news and consumption variability. They therefore suggest that sentiments in news have strong predictive power for demand forecasting in unplanned situations.

Mots clés – analyse de sentiments, prévision de la demande, crise, traitement automatique du langage naturel, média.

Keywords – sentiment analysis, demand forecasting, crisis management, natural language processing, news media.

1 INTRODUCTION

L'occurrence d'un événement non souhaité tel qu'une catastrophe naturelle ou une épidémie s'accompagne généralement de fortes perturbations dans les chaînes logistiques. Dans le secteur de la santé par exemple, de telles situations entraînent une augmentation soudaine et forte de la demande en soins, équipements médicaux et produits pharmaceutiques, tandis que le fonctionnement des chaînes d'approvisionnement se trouve affecté (Ivanov et Dolgui 2020). De plus, des phénomènes d'achats massifs de produits d'usage courant (e.g. alimentaires) sont souvent observés lors de situations de crise comme lors de l'annonce de la pandémie de COVID-19. Ces perturbations ont des conséquences multiples. D'abord, les pénuries de certains produits de santé (e.g. médicaments, masques) ou de soins médicaux affectent directement la gestion de la crise et la santé publique. D'autre part, la forte volatilité des marchés peut également entraîner des surstocks : en juin 2020, 250000 œufs ont ainsi été détruits (JLL 2020). La prévision de la demande devient alors particulièrement difficile, mais nécessaire.

Au cours de la dernière décennie, les recherches et les applications de l'intelligence artificielle (IA) et de la science, des données ont montré des résultats particulièrement

prometteurs pour améliorer les chaînes logistiques. Cependant, la récente crise du COVID-19 a mis en lumière les limites des algorithmes d'apprentissage automatique, qui, d'ordinaire performants, n'ont pas été capables de « comprendre nos comportements étranges » (Heaven 2020). Par exemple, les historiques de demande et de recherches ou encore les nombres de clics publicitaires n'ont pas permis de fournir des informations fiables sur la demande devenue fortement volatile (Heaven 2020). En outre, les recherches se sont jusqu'ici peu penchées sur les applications de l'IA en période de crise et sur la robustesse des algorithmes existants. Ainsi Ivanov et Dolgui (2020) soulignent le besoin de nouvelles applications de l'IA et de la science des données de fournir de l'aide à la décision en temps de crise.

Par ailleurs, des niveaux d'audience record des médias d'information tels que le journal *The New York Times* ont été observés pendant la crise du COVID-19 (JP Morgan 2020). En effet, lors d'événements non planifiés, les médias d'information représentent la principale source d'information actualisée au public. Or les publications des médias d'information sont une source de données encore peu exploitée dans les applications d'IA et de science des données en logistique. En outre Weill et al. (2020) soulignent l'influence

des médias sur la consommation des médicaments en France pendant la pandémie.

Dans ce contexte, l'objectif de cet article est de déterminer si l'application des méthodes de traitement automatique du langage naturel (TALN), et en particulier l'analyse de sentiments, sur les médias d'information, peut fournir une information utile pour la prévision de la demande en temps de crise. L'article propose une méthodologie pour (1) collecter et préparer les données; (2) mettre en œuvre un modèle d'analyse de sentiments basé sur de l'apprentissage automatique supervisé; et (3) analyser la corrélation entre les sentiments véhiculés dans les médias et la variation de la consommation. Ce schéma algorithmique est appliqué au cas de la consommation de la Chloroquine/l'Hydroxychloroquine (Cq/HCq) en France pendant la crise du COVID-19.

La suite de cet article s'organise de la manière suivante : la section 2 présente un état de l'art sur l'analyse de sentiments et l'IA pour l'aide à la décision en temps de crise ; la section 3 décrit le schéma algorithmique suivi dans ce cas d'application ; la section 4 présente les résultats de l'analyse statistique ; la section 5 discute certains de ces résultats ; et enfin, la section 6 conclut et dégage quelques perspectives de recherche.

2 ÉTAT DE L'ART ET CONTRIBUTIONS

La science des données désigne le sous domaine des mathématiques, des statistiques et de l'informatique visant à gérer et valoriser les données, notamment celles générées en grands volume, vélocité, et variété, aussi nommées *big data* (Gandomi et Haider 2014). En particulier, les systèmes d'intelligence artificielle (e.g. *chatbots*, reconnaissance d'images) utilisent largement les méthodes de fouille de données et des algorithmes d'apprentissage automatique. Cependant, d'après Gandomi et Haider (2014), la plupart des applications existantes s'intéressent aux données structurées, qui représentent seulement 5 % des *big data*. La valorisation des données non structurées telles que les textes ou les vidéos est encore largement lacunaire.

Un exemple de valorisation des données textuelles concerne l'analyse de sentiments pour l'aide à la décision. L'analyse de sentiments, ou fouille d'opinion, est un sous-domaine du TALN visant à identifier des sentiments, émotions, et connotations contenues dans un texte (ou une donnée vocale), la plupart du temps par l'extraction d'une polarité (c.-à-d. positif, négatif, ou neutre) (Sun, Luo, and Chen 2017). La plupart des contributions existant dans la littérature se concentrent sur les secteurs de l'énergie et de la finance. Par exemple Grundmann, Felden, et Pospiech (2016) utilisent une approche hybride basée sur les règles et sur des modèles d'apprentissage automatique pour identifier les tendances de prix du gaz naturel à partir des publications de médias spécialisés. De même, plusieurs articles ont exploré différentes techniques pour identifier les sentiments dans les médias spécialisés dans la finance pour fournir de l'information sur les marchés boursiers (Chan et Chong 2017 ; Fang et al. 2017 ; Moreno-Ortiz et Fernández-Cruz 2015). Kraaijeveld et De Smedt (2020) ont mis en œuvre un test de Granger pour quantifier l'influence des sentiments identifiables sur Twitter sur les prix de cryptomonnaies. Dans un autre domaine, Gaspar et al. (2016) identifient de manière plus fine les types de sentiments (e.g. espoir, peur, confiance, etc.) sur les réseaux sociaux pendant des événements difficiles ou traumatisants. Cependant, peu d'applications se sont intéressées à coupler ces informations de sentiments avec des données structurées (e.g. revenus) à l'aide de modèles d'analyse prédictive. Parmi les contributions trouvées, Cakra (2015) et Gupta et Banerjee

(2019) ont couplé un modèle d'analyse de sentiments avec un modèle d'analyse prédictive de séries temporelles pour prédire les valeurs boursières à partir de réseaux sociaux et médias spécialisés. Asur and Huberman (2010) ont utilisé les données de sentiments sur Twitter pour prédire les revenus du box-office.

Dans les systèmes de production et les chaînes logistiques, la plupart des contributions appliquant les méthodes d'analyse de sentiments ont utilisé les données d'avis clients. Par exemple Sun et al. (2019) utilisent les ontologies et la théorie des sous-ensembles flous pour construire un modèle d'analyse de sentiments à haut niveau de granularité sur des avis en ligne. Fan, Che, et Chen (2017) proposent un modèle permettant d'ajuster le modèle économique de Bass. Dans le secteur de la santé, une étude récente a comparé plusieurs architectures d'apprentissage profond telles que des réseaux de neurones récurrents LSTM (*Long short-term memory*) et les transformers bidirectionnels profonds (*Bidirectional Encoder Representations from Transformers*, BERT) pour de l'analyse de sentiment sur les avis des patients concernant leurs traitements médicaux (Colón-Ruiz and Segura-Bedmar 2020). Toutefois, l'influence des sentiments véhiculés par les médias sur les comportements de consommation n'a pas été explorée.

Un autre axe de recherche auquel s'intéresse cet article est celui de l'intelligence artificielle et de l'application de la science des données en gestion de crise. La crise du COVID-19 survenue en 2020 a mis en évidence les failles existantes dans les chaînes et réseaux logistiques de plus en plus complexes (Ivanov and Dolgui 2020), qui se sont par exemple traduites par des ruptures de stock en produits de santé (e.g. masques). Dans ce cadre Ivanov (2020) et Ivanov et Dolgui (2020) ont souligné le besoin de nouvelles contributions de l'intelligence artificielle et de la science des données pour prodiguer une aide à la décision dans la gestion des chaînes logistiques en temps de crise. D'autre part, la pandémie a aussi mis en lumière les limites des systèmes actuels d'intelligence artificielle et des modèles d'apprentissage automatique, qui n'ont pas su comprendre nos « comportements étranges » pendant la crise (Heaven 2020). D'après Heaven (2020), cela provient du fait que les données en entrée des algorithmes pendant la crise diffèrent trop fortement (e.g. marchés fortement volatiles) des données sur lesquelles ils ont été entraînés. Au-delà de cette difficulté, il semble que les algorithmes d'apprentissage automatique utilisés en temps régulier ne sont pas conçus pour évoluer dans un environnement de crise.

Dans ce contexte, les objectifs de cet article sont les suivants : (i) contribuer aux applications de TALN pour l'aide à la décision et de modèles prédictifs couplant des données textuelles, non structurées, à des données structurées de consommation ; (ii) analyser le potentiel pouvoir prédictif de l'analyse de sentiment des médias d'information pour la prévision des variations de consommation ; et (iii) initier la recherche et les applications d'IA et de science des données pour la gestion de crise.

3 METHODOLOGIE

Le schéma algorithmique mis en œuvre dans ce cas d'application est détaillé dans la Figure 1 et s'articule en trois étapes : (1) collecter et préparer les données de média et de consommation de médicament ; (2) définir et appliquer un modèle de TALN sur les données de média et agréger les résultats en séries temporelles ; (3) analyser la corrélation entre publications des médias et variations de la consommation de médicament.

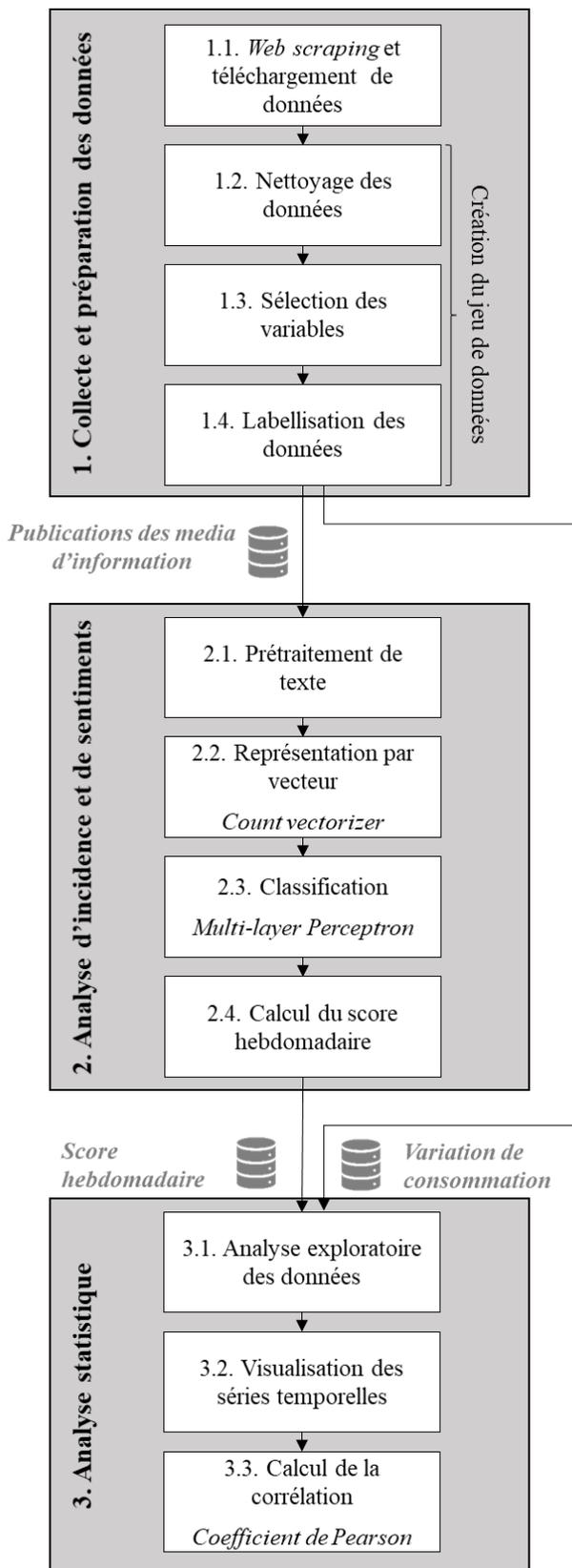


Figure 1. Schéma algorithmique

3.1 Collecte et préparation des données

L'objectif de cette étape est de collecter et préparer les données de publications des médias d'information et de consommation de Cq/HCq en France.

Dans un premier temps, un jeu de données permettant de mettre en œuvre un modèle d'analyse de sentiments des médias basés sur de l'apprentissage automatique supervisé est construit. A cette fin, les outils de *web scraping* sont utilisés

pour collecter les titres des publications sur trois médicaments liés au COVID-19 (paracétamol, ibuprofène, et Cq/HCq) (Weill et al. 2020) par les médias d'information les plus utilisés en France : 20Minutes, Le Monde, France Info, Le Figaro, HuffPost, BFMTV, L'Express, Le Point, Médiapart, TF1, L'Obs, et Libération (Gaudiau 2018). Ainsi, toutes les publications contenant les termes « paracétamol », « ibuprofène », « chloroquine », ou « hydroxychloroquine » entre janvier 2020 et novembre 2020, et relayées par ces médias sur le réseau social Twitter, ont été recueillies. Ces données ont ensuite été nettoyées (e.g. suppression des doublons) et labellisées de la manière suivante : positif (+1) si la publication contient un message explicitement ou implicitement en faveur de la consommation du médicament ; négatif (-1) si elle véhicule un message défavorable au traitement ; neutre (0) si le message véhiculé n'est pas orienté ou s'il ne permet pas de conclure clairement quant à son influence sur la consommation du médicament. Un jeu de données labellisées de 1245 instances a ainsi été construit.

Dans un deuxième temps, les données de consommation hebdomadaire de Cq/HCq entre les semaines 10-2020 et 20-2020 en France sont téléchargées. Ces données sont publiées par le groupement d'intérêt scientifique EPI-PHARE (Weill et al. 2020). En particulier, la variable étudiée dans cette recherche est le pourcentage d'écart entre la consommation réelle de Cq/HCq durant cette période et la consommation normale de ce traitement. Dans la suite de cet article, elle sera appelée *variation de la consommation de Cq/HCq*. Cette variable permet de refléter les changements de comportements de consommation entre une situation de crise et une situation régulière.

3.2 Analyse de sentiments et incidence des médias

La seconde étape du schéma algorithmique vise à définir et appliquer les méthodes de TALN et d'apprentissage automatique supervisé pour traiter les publications des médias et agréger les résultats en séries temporelles. En particulier, deux variables sont définies et comparées : d'une part, l'*incidence* représente le nombre hebdomadaire de publications sur le médicament ; d'autre part, le *sentiment* désigne un score hebdomadaire représentant la polarité (c.-à-d. positif, négatif, neutre) des médias. La comparaison entre ces deux variables permet de mesurer la valeur ajoutée potentielle de l'analyse de sentiments dans la prévision de la consommation. En effet, la variable *incidence* prend en compte les publications des médias sans faire appel aux méthodes de TALN puisqu'elle ne considère que le nombre de publications en ligne (données semi-structurées). En revanche, la construction de la variable *sentiment* sollicite des méthodes d'analyse de sentiments.

Le modèle d'analyse de sentiments mis en œuvre dans ce cas d'application s'organise en deux tâches permettant de (i) représenter un texte (c.-à-d. titre d'une publication) par un vecteur d'occurrences (*count vectorizer*) ; et de (ii) classifier automatiquement chaque vecteur dans une catégorie (c.-à-d. positif, négatif, ou neutre) à l'aide d'un réseau de neurones Perceptron multicouche (*multi-layer Perceptron*, MLP). Ces résultats sont ensuite agrégés en une série temporelle, notée *sentiment*, de la manière suivante :

$$\begin{aligned} \text{sentiment}_{\text{semaine}} = & (+1) \cdot \text{nombre de positifs}_{\text{semaine}} \\ & + (-1) \cdot \text{nombre de négatifs}_{\text{semaine}} \\ & + (0) \cdot \text{nombre de neutres}_{\text{semaine}} \end{aligned}$$

3.3 Analyse statistique

L'objectif de cette dernière étape est d'effectuer une analyse statistique de la relation entre les médias et la variation de la consommation de Cq/HCq pendant la crise du COVID-19 en France. Tout d'abord, une analyse exploratoire des données ainsi qu'une visualisation des séries temporelles permet de caractériser les publications médiatiques concernant la Cq/HCq ayant eu lieu entre les semaines 10-2020 et 20-2020 en France et leur potentiel lien avec la variation de consommation du médicament. Enfin, la corrélation linéaire entre ces variables est quantifiée à l'aide du coefficient de Pearson.

4 RESULTATS

Cette section présente les résultats de l'analyse statistique décrite en 3.3.

4.1 Analyse exploratoire des données

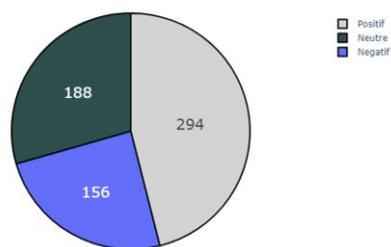


Figure 2. Distribution des données

La Figure 2 montre la distribution des publications positives, négatives, et neutres entre les semaines 10-2020 et 20-2020. Elle souligne la prépondérance des publications positives par rapport aux publications négatives durant la période étudiée. De plus, il est à noter que le nombre de publications neutres est artificiellement élevé, car certaines publications ne permettaient pas de déterminer clairement une potentielle influence positive ou négative sur la consommation de Cq/HCq. Celles-ci, qui comprennent notamment les publications au contenu politique, sont labellisées neutres. Le Tableau 1 présente des exemples de titres et leurs labels.

Tableau 1. Exemples de titres de publications et leurs labels

Titre	Label
La chloroquine, ce vieux médicament qui pourrait s'avérer utile contre le coronavirus	+1
Coronavirus : Donald Trump annonce prendre de l'hydroxychloroquine "depuis une semaine et demie"	0
On parle beaucoup de la chloroquine, qu'est-ce que c'est précisément ?	0
Dix hospitalisations après la prise d'hydroxychloroquine en automédication	-1

4.2 Visualisation des séries temporelles

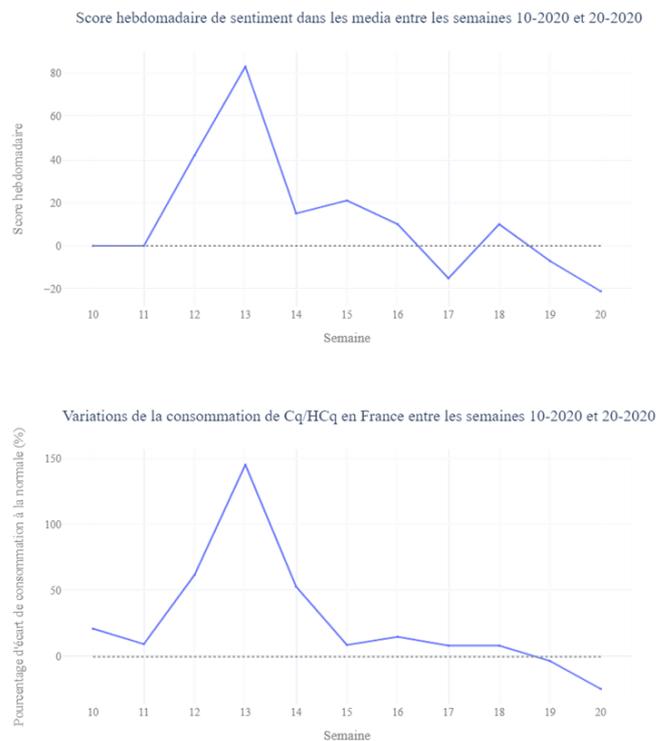


Figure 3. Séries temporelles de sentiment (en haut) et de variation de la consommation de Cq/HCq (en bas) entre les semaines 10-2020 et 20-2020

La visualisation des séries temporelles de *sentiment* et de variation de la consommation de Cq/HCq entre les semaines 10-2020 et 20-2020 (Figure 3) indique de fortes similitudes entre les deux variables. En effet, le score maximum de *sentiment* des médias, atteint en semaine 13, coïncide avec le pic de variation de consommation de Cq/HCq. Ainsi, alors que le nombre de publications positives concernant le traitement était particulièrement élevé, le nombre de prescriptions de Cq/HCq a été 150 % plus élevé que la normale. De même, lorsque le nombre de publications négatives a dépassé le nombre de publications positives à partir de la semaine 19 (score de sentiment en dessous de 0), l'écart de consommation de Cq/HCq est également tombé en dessous de 0 %, ce qui signifie que la consommation de ce médicament a alors été inférieure à la normale. Cette analyse suggère ainsi que la prescription de ce médicament a été fortement influencée par les médias.

4.3 Calcul de la corrélation

Le calcul du *coefficient de Pearson r* permet de mesurer la corrélation linéaire entre les variations de la consommation de Cq/HCq et l'*incidence* et les mêmes variations et le score de *sentiment* (Tableau 2). La valeur de r égale à 0.85, avec une p -valeur de $8.75e-4$ (sous le seuil de significativité défini à 0.01) confirme bien qu'il existe une corrélation significative, positive et forte entre la variation de la consommation de Cq/HCq et le nombre de publications concernant le médicament. De plus, la corrélation entre variation de la consommation de Cq/HCq et le score de *sentiment* ($r=0.93$) est particulièrement élevée, signifiant donc que l'analyse de sentiments des médias pourrait avoir un pouvoir prédictif fort sur les variations de consommation en temps de crise.

Tableau 2. Coefficients de corrélation (de Pearson r) entre variation de la consommation de Cq/HCq (var.cons) et incidence ; et entre var.cons. et sentiment

	r	p-valeur
var.cons/incidence	0.85	8.75e-4
var.cons/sentiment	0.93	2.49e-5

5 DISCUSSION

5.1 Apport du TALN

Le calcul des coefficients de Pearson met en évidence une corrélation plus forte avec les sentiments dans les médias qu'avec l'incidence (qui mesure la couverture médiatique). Cela démontre l'apport du TALN sur les médias d'information comparé à l'utilisation exclusive de données structurées ou semi-structurées (e.g. nombre de publications). Les résultats de ce cas d'application révèlent ainsi que l'analyse de sentiments des médias d'information couplée à un modèle prédictif de séries temporelles pourrait permettre d'obtenir un modèle de prévision de la demande performant et fiable en période de crise. Toutefois, il est à noter que les variables *incidence* et *sentiment* définies en 3.2. sont liées puisque le score hebdomadaire de sentiment prend en compte le nombre de publications. De plus, l'utilisation du score de *sentiment* dans un modèle d'analyse prédictive de série temporelle devrait être conditionnée par un nombre minimal de publications hebdomadaires, c'est-à-dire un seuil d'*incidence* qui devrait être déterminé.

5.2 Techniques d'analyse de sentiment

Le modèle d'analyse de sentiments développé s'organise en deux tâches consistant à (i) représenter un texte (c.-à-d. titre d'une publication) par un vecteur ; (ii) classifier chaque vecteur dans une catégorie (c.-à-d. positif, négatif, ou neutre). Plusieurs techniques ont été testées : *count vectorizer* et *tf-idf* pour la tâche (i) ; régression logistique, machine à vecteurs de support (*support vector machine*, SVM) et réseau de neurones Perceptron multicouches (*multi-layer Perceptron*, MLP) pour la tâche (ii). Le Tableau 3 présente les mesures de performance (score F1) pour chaque technique. Finalement, le modèle appliqué comprend (i) un *count vectorizer* ; (ii) un MLP à deux couches cachées, et atteint le score de performance de 0.79. Le Tableau 4 détaille les mesures de performance du modèle appliqué. Il montre que le modèle est moins performant pour identifier les publications neutres, ce qui est également dû à la forte hétérogénéité des données labellisées neutres. De futures recherches pourront ainsi explorer d'autres techniques pour améliorer la performance du modèle.

Tableau 3. Score F1 obtenu pour différentes techniques de représentation du texte et de classification supervisée

	Régression logistique	SVM	MLP
<i>Count vectorizer</i>	0.78	0.78	0.79
<i>Tf-idf</i>	0.79	0.78	0.77

Par ailleurs, dans ce cas d'application, seuls les titres des publications ont été analysés, car ceux-ci représentent la partie la plus importante pour l'analyse de sentiments, d'après Hui, Hoon, et Zainon (2017). Toutefois, inclure d'autres parties des publications, telles que les résumés ou les conclusions, pourrait également améliorer la performance du modèle.

Tableau 4. Mesures de performance du modèle d'analyse de sentiment

Label	Précision	Rappel	F1
-1	0.82	0.87	0.85
0	0.73	0.73	0.73
+1	0.80	0.76	0.78
Performance			0.79

5.3 Corrélation vs. causalité

La relation entre sentiments véhiculés par les médias et variation de la consommation de médicament a été mesurée quantitativement à l'aide du coefficient de corrélation de Pearson. Récemment, plusieurs auteurs ont discuté des limites inhérentes aux analyses de corrélation, qui ne prennent pas en compte les connaissances contextuelles, et ont argumenté pour la réintroduction de l'analyse causale dans les modèles statistiques Lemberger et Oblin (2020). Aussi, les recherches futures pourront analyser le lien causal entre ces deux variables à l'aide de réseaux Bayésiens par exemple. Toutefois, dans le cas de la crise du COVID-19, aucune donnée médicale officielle n'a conclu sur l'efficacité de ce traitement contre le COVID-19 en semaine 20 (World Health Organization 2020). Par conséquent, cela ne peut pas représenter une source commune corrélée à la fois à la variation de la consommation de Cq/HCq et aux sentiments. Ainsi, il semble donc que les sentiments véhiculés par les médias aient bien influencé la prescription de Cq/HCq pendant la crise du COVID-19 en France, confirmant l'analyse de Weill et al. (2020).

6 CONCLUSION

L'étude de cas présentée dans cet article s'est articulée en trois étapes : (1) d'abord, les publications des médias d'information et les données de consommation de Cq/HCq ont été collectées et un jeu de données a été construit ; (2) ensuite, un modèle d'analyse de sentiments basé sur de l'apprentissage automatique supervisé a été développé et deux séries temporelles ont été définies ; (3) enfin, une analyse statistique a permis de mesurer la corrélation entre les sentiments véhiculés par les médias et la variation de la consommation de Cq/HCq par rapport à la normale pendant la crise du COVID-19 en France. Les résultats ont mis en lumière une corrélation significative positive forte entre la variation de la consommation de Cq/HCq et les sentiments dans les médias. Par conséquent, les modèles d'analyse de sentiments, couplés aux modèles d'analyse prédictive de séries temporelles, devraient permettre de fournir des prévisions fiables de la demande en période de crise. Par ailleurs, les résultats semblent bien confirmer que les publications des médias d'information représentent une source de données intéressante pour les systèmes d'IA en période de crise : lors de telles situations, la couverture médiatique devrait être assez large (incidence suffisamment grande) pour assurer une quantité de données suffisante à l'application de modèles prédictifs. Cependant, ce cas d'application a plusieurs limites : d'abord, les données n'ont pas été labellisées par assez de personnes pour assurer une précision et une objectivité de l'analyse ; ensuite, le cas étudié s'est concentré sur la consommation d'un médicament pendant une période assez restreinte. Ainsi, plusieurs perspectives de recherche peuvent être dégagées : (i) cette étude sera étendue à d'autres produits et d'autres périodes ; (ii) des modèles prédictifs seront testés et couplés

aux modèles d'analyse de sentiments pour développer un modèle de prévision de la demande en période de crise ; (iii) les techniques état de l'art d'analyse de sentiments seront testées pour améliorer la performance et la finesse des résultats.

7 REFERENCES

- Asur, Sitaram and Bernardo A. Huberman. 2010. "Predicting the Future with Social Media." *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010* 1:492–99.
- Cakra, Trisedya. 2015. "Stock Price Prediction Using Linear Regression Based on Sentiment Analysis." *International Journal of Scientific & Engineering Research* 6(3):1655–59.
- Chan, Samuel W. K. and Mickey W. C. Chong. 2017. "Sentiment Analysis in Financial Texts." *Decision Support Systems* 94:53–64.
- Colón-Ruiz, Cristóbal and Isabel Segura-Bedmar. 2020. "Comparing Deep Learning Architectures for Sentiment Analysis on Drug Reviews." *Journal of Biomedical Informatics* 110(February):103539.
- Fan, Zhi Ping, Yu Jie Che, and Zhen Yu Chen. 2017. "Product Sales Forecasting Using Online Reviews and Historical Sales Data: A Method Combining the Bass Model and Sentiment Analysis." *Journal of Business Research* 74:90–100.
- Fang, Bing, Sheng Liang, Qian Zou, and Wei Huang. 2017. "Research on Sentiment Analysis of Financial Texts." *Proceedings of 2017 IEEE 7th International Conference on Electronics Information and Emergency Communication, ICEIEC 2017* 248–52.
- Gandomi, Amir and Murtaza Haider. 2014. "Beyond the Hype: Big Data Concepts , Methods , and Analytics." *International Journal of Information Management* 35(2):137–44.
- Gaspar, Rui, Cláudia Pedro, Panos Panagiotopoulos, and Beate Seibt. 2016. "Beyond Positive or Negative: Qualitative Sentiment Analysis of Social Media Reactions to Unexpected Stressful Events." *Computers in Human Behavior* 56:179–91.
- Gaudiaut, Tristan. 2018. "Les Médias d'information Les plus Utilisés En France." *Statista*.
- Grundmann, Tina, Carsten Felden, and Marco Pospiech. 2016. "Forecasting the Natural Gas Price Trend-Evaluation of a Sentiment Analysis." *Proceedings - 2016 4th International Conference on Future Internet of Things and Cloud Workshops, W-FiCloud 2016* (Section 1):160–64.
- Gupta, Kartick and Rajabrata Banerjee. 2019. "Does OPEC News Sentiment Influence Stock Returns of Energy Firms in the United States?" *Energy Economics* 77:34–45.
- Heaven, Will Douglas. 2020. "Our Weird Behavior during the Pandemic Is Messing with AI Models." *MIT Technology Review*.
- Hui, June Ling Ong, Gan Keng Hoon, and Wan Mohd Nazmee Wan Zainon. 2017. "Effects of Word Class and Text Position in Sentiment-Based News Classification." *Procedia Computer Science* 124:77–85.
- Ivanov, Dmitry. 2020. "Predicting the Impacts of Epidemic Outbreaks on Global Supply Chains: A Simulation-Based Analysis on the Coronavirus Outbreak (COVID-19/SARS-CoV-2) Case." *Transportation Research Part E: Logistics and Transportation Review* 136(March):101922.
- Ivanov, Dmitry and Alexandre Dolgui. 2020. "Viability of Intertwined Supply Networks : Extending the Supply Chain Resilience Angles towards Survivability . A Position Paper Motivated by COVID-19 Outbreak." *International Journal of Production Research* 7543.
- JLL. 2020. "Too Many Eggs: Supply Chain Shocks Arise from COVID-19." Retrieved July 10, 2020 (<https://www.jll.fr/fr/etudes-recherche/marches/too-many-eggs-covid-19-turns-focus-on-the-bullwhip-effect>).
- JP Morgan. 2020. "Media Consumption in the Age of COVID-19." Retrieved May 1, 2020 (<https://www.jpmorgan.com/global/research/media-consumption>).
- Kraaijeveld, Olivier and Johannes De Smedt. 2020. "The Predictive Power of Public Twitter Sentiment for Forecasting Cryptocurrency Prices." *Journal of International Financial Markets, Institutions and Money* 65:101188.
- Lemberger, Pirmin and Denis Oblin. 2020. "Reconciling Causality and Statistics." 1–22.
- Moreno-Ortiz, Antonio and Javier Fernández-Cruz. 2015. "Identifying Polarity in Financial Texts for Sentiment Analysis: A Corpus-Based Approach." *Procedia - Social and Behavioral Sciences* 198(Cilc):330–38.
- Sun, Qing, Jianwei Niu, Zhong Yao, and Hao Yan. 2019. "Exploring EWOM in Online Customer Reviews: Sentiment Analysis at a Fine-Grained Level." *Engineering Applications of Artificial Intelligence* 81(December 2018):68–78.
- Sun, Shiliang, Chen Luo, and Junyu Chen. 2017. "A Review of Natural Language Processing Techniques for Opinion Mining Systems." *Information Fusion* 36:10–25.
- Weill, Alain, Jérôme Drouin, David Desplas, Francois Cuenot, Rosemary Dray-Spira, and Mahmoud Zureik. 2020. *Usage Des Médicaments de Ville En France Durant l ' Épidémie de La Covid-19 – Point de Situation Après Les 8 Semaines de Confinement*.
- World Health Organization. 2020. "Coronavirus Disease (COVID-19): Hydroxychloroquine." *WHO Q&A*. Retrieved June 19, 2020 (<https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-hydroxychloroquine>).

Etude de la Perte Focale pour le Traitement du Langage Naturel en Maintenance dans l'Industrie 4.0

JUAN PABLO USUGA-CADAVID¹, SAMIR LAMOURI², BERNARD GRABOT³, ARNAUD FORTIN⁴

¹ LAMIH UMR CNRS 8201, ARTS ET METIERS - SCIENCES ET TECHNOLOGIES
151 Boulevard de l'Hôpital, 75013, Paris, France
juan_pablo.usuga_cadavid@ensam.eu

² LAMIH UMR CNRS 8201, ARTS ET METIERS - SCIENCES ET TECHNOLOGIES
151 Boulevard de l'Hôpital, 75013, Paris, France
samir.lamouri@ensam.eu

³ LGP, INP/ENIT
47 Avenue d'Azereix, 65000, Tarbes, France
bernard.grabot@enit.fr

⁴ IFAKT FRANCE SAS
8 Esplanade compans caffarelli, 31000, Toulouse, France
a.fortin@ifakt.com

Résumé – Exploiter les données de maintenance se trouvant dans les systèmes d'information peut s'avérer éprouvant. En effet, elles reposent principalement sur des commentaires et des descriptions rédigés par des opérateurs de maintenance. De ce fait, ces données textuelles contiennent des irrégularités comme des acronymes, des jargons, des fautes de frappe, entre autres, ce qui peut entraver leur traitement automatique. En outre, les jeux de données de maintenance présentent souvent des distributions déséquilibrées : quelques événements apparaissent plus souvent que d'autres, ce qui nuit à la performance des modèles classiques d'Apprentissage Automatique (AA). Ainsi, cet article explore l'usage d'une architecture récente d'Apprentissage Profond (AP) appelée Transformer, fournissant des résultats de pointe en Traitement du Langage Naturel (TLN). Pour aborder le déséquilibre de classes, une fonction de perte appelée la Perte Focale (PF) est explorée. Les résultats de l'étude suggèrent que quand toutes les classes ont la même importance, la PF n'améliore pas la performance de la classification. En revanche si l'objectif est de détecter la classe minoritaire, la PF permet d'atteindre la meilleure performance parmi les méthodes comparées. Cependant, cela dégrade la capacité du modèle à détecter la classe majoritaire.

Abstract – Harnessing data from historical maintenance databases may be challenging, as they tend to rely on text data provided by operators. Thus, they often include acronyms, jargon, typos, and other irregularities that complicate the automated analysis of such reports. Furthermore, maintenance datasets may present highly imbalanced distributions: some situations happen more often than others, which hinders the effective application of classic Machine Learning (ML) models. Hence, this paper explores the use of a recent Deep Learning (DL) architecture called Transformer, which has provided cutting-edge results in Natural Language Processing (NLP). To tackle the class imbalance, a loss function called Focal Loss (FL) is explored. Results suggests that when all the classes are equally important, the FL does not improve the classification performance. However, if the objective is to detect the minority class, the FL achieves the best performance, although by degrading the performance for the majority class.

Mots clés – Traitement du Langage Naturel, Maintenance Prédictive, Classification Déséquilibrée, Apprentissage Profond, Apprentissage par Transfert.

Keywords – Natural Language Processing, Predictive Maintenance, Imbalanced Classification, Deep Learning, Transfer Learning.

1 INTRODUCTION

L'un des leviers pour améliorer les systèmes de production est de mieux traiter les perturbations, comme les pannes de machine. En effet, les perturbations peuvent rendre infaisable le plan de production existant et entraîner des retards de livraison [Usuga Cadavid et al., 2019, 2020a]. Avec l'arrivée de l'Industrie 4.0 (I4.0), les compagnies sont plus disposées à créer des modèles permettant une meilleure caractérisation et prédiction de pannes de machines. Néanmoins, les données pour entraîner des modèles d'AA peuvent être difficiles à collecter dans un volume, fréquence et qualité suffisants pour satisfaire les besoins du métier. Même si l'Internet des Objets

peut se présenter comme une solution pour réaliser une collecte de données appropriée, cette technologie n'est pas encore adaptée à toutes les industries. Par exemple, quelques industries, comme celle des semi-conducteurs, sont capables de générer de grands volumes de données rapidement. Cependant, d'autres industries dont les processus dépendent plus des activités manuelles menées par des êtres humains peuvent avoir du mal pour atteindre les mêmes résultats [Kusiak, 2019].

Dans le contexte de la maintenance, une autre possibilité pour entraîner des modèles d'AA est d'utiliser des données déjà stockées dans des systèmes d'information : les historiques des

journaux de maintenance. Or, ces données présentent deux problèmes entravant leur utilisation : premièrement, ils reposent souvent sur des descriptions non structurées fournies par les opérateurs au moyen de champs de saisie en texte libre sur des logiciels de gestion de la maintenance. Secondement, ils ont la tendance à avoir des données intrinsèquement déséquilibrées. Cela veut dire que certains événements sont sous-représentés par rapport à d'autres. Par exemple, les problèmes sévères arrivent rarement par rapport à des problèmes peu graves de machines [Usuga Cadavid et al., 2020b]. Les jeux de données déséquilibrés affectent la performance des modèles d'AA, car ces derniers ont la tendance à mieux apprendre à détecter la classe majoritaire, en ignorant la classe minoritaire. En outre, appliquer des modèles d'AA doit assurer que l'être humain est inclus dans la boucle. Sinon, les industries ayant des opérations fortement dépendantes de l'activité humaine peuvent être réticentes à automatiser leurs activités de maintenance [Usuga Cadavid et al., 2020a]. De ce fait, l'exploitation des données se trouvant dans des journaux de maintenance pose trois défis principaux : traiter des données non structurées, mitiger l'effet du déséquilibre de classes, et garder l'être humain dans la boucle.

Pour traiter des données non structurées et garder l'être humain dans la boucle, des approches récentes de TLN utilisant l'AP permettent d'entraîner des modèles ayant une bonne performance avec un prétraitement de données minimale. Ainsi, les rapports de maintenance fournis par les opérateurs peuvent être utilisés directement, en changeant peu leur manière de travailler. Ces approches de TLN sont appelées Transformers et ils sont basés sur des mécanismes d'attention [Vaswani et al., 2017].

De manière globale, un Transformer est un réseau de neurones utilisant une architecture de type encodeur-décodeur et de multiples têtes d'attention. Ces modèles ont deux avantages principaux au moment de travailler avec des données de type texte : premièrement, ils peuvent considérer le contexte d'utilisation d'un mot dans une phrase en employant les autres mots autour. Secondement, les calculs pour entraîner un Transformer peuvent être facilement parallélisés sur plusieurs ordinateurs, contrairement aux réseaux récurrents de neurones. Le premier avantage permet de trouver des représentations numériques de mots qui sont plus riches que celles obtenues avec des méthodes de plongement lexical non contextualisées comme Word2Vec ou des méthodes "fréquentistes" comme TF-IdF [Martin et al., 2019; Usuga Cadavid et al., 2020b]. Le second avantage permet l'entraînement de grands modèles pour ensuite les ajuster à d'autres tâches avec peu d'effort. Cela est connu sous le nom d'Apprentissage par Transfert (AT) : préentraîner des modèles avec des données massives pour qu'ils apprennent des transformations de données pertinentes pour ensuite les ajuster à de nouvelles tâches (analyse de sentiments, moteurs de réponse, etc.). De ce fait, en utilisant l'AT un modèle peut atteindre une bonne performance en utilisant moins de données et de temps. Pour une étude détaillée sur l'AT, veuillez consulter l'article de Pan et Yang [2010].

Malgré les avantages fournis par l'AT en TLN, les défis imposés par le déséquilibre de classes peuvent encore nuire à l'apprentissage des modèles d'AP. De plus, la mitigation du déséquilibre de classes dans le TLN reste encore un domaine sous-étudié par rapport à la vision par ordinateur : l'enquête sur l'utilisation de l'AP pour le déséquilibre de classes réalisée

par Johnson et Khoshgoftaar [2019], affirme qu'environ 80 % de la recherche réalisée dans ce domaine se focalise sur l'étude des réseaux de neurones "convolutifs" (CNNs, par son acronyme en anglais) en vision par ordinateur. Il est donc important d'explorer des techniques de mitigation du déséquilibre de classes sur d'autres domaines et architectures, comme le TLN et les Transformers.

Johnson et Khoshgoftaar [2019] proposent trois types d'approches pour mitiger l'effet du déséquilibre de classes : des approches axées sur les données se basant sur des stratégies d'échantillonnage (par exemple, le sur- et sous-échantillonnage), des approches axées sur les algorithmes modifiant la manière dont les modèles d'AA apprennent (par exemple, l'apprentissage sensible au coût) et les approches hybrides, mélangeant les deux dernières. Le périmètre de cette étude se focalise exclusivement sur les approches axées sur les algorithmes. Plus spécifiquement, cette étude évalue l'utilisation de la PF, initialement proposée par Lin et al. [2017] pour des applications de vision par ordinateur, sur les Transformers. Pour cela faire, un jeu de données réelles provenant d'une entreprise est utilisé.

Cette recherche a trois objectifs principaux : premièrement, entraîner un modèle d'AP pour utiliser des descriptions de pannes de machine et déterminer si cela correspond à un problème qui arrêtera le système de production ou pas. Deuxièmement, explorer l'influence de plusieurs configurations de la PF sur la performance du modèle. Finalement, évaluer l'effectivité d'utiliser la PF pour des jeux de données déséquilibrés de maintenance en TLN.

Le premier objectif est abordé avec l'AT en ajustant un modèle appelé CamemBERT à classer des descriptions de panne de machine dans deux catégories possibles : *arrêt de production* et *non-arrêt de production*. Le deuxième objectif est traité en réalisant du réglage d'hyperparamètres sur la PF et en évaluant leur influence sur la performance de la classification. Finalement, le troisième objectif est abordé en comparant les résultats obtenus avec la PF avec d'autres fonctions de perte plus courantes (l'entropie croisée et sa forme pondérée). Cette comparaison sera complétée au moyen d'une visualisation utilisant l'Analyse de Composantes Principales (ACP) pour représenter les vecteurs de plongement lexical (ci-après dénommés « embeddings » par leur nom plus courant en anglais) appris par les modèles.

La suite de cette étude est articulée de la manière suivante : la section 2 explore d'autres travaux proches à cet article. La section 3 présente le jeu de données utilisé, le Transformer employé, les fonctions de perte à comparer, la politique d'entraînement et la stratégie de visualisation des embeddings. La section 4 présente et analyse les résultats. Enfin, la section 5 conclut cette étude et fournit des perspectives de recherche.

2 D'AUTRES TRAVAUX DANS LA LITTÉRATURE

Cette section présente quelques études mitigeant le déséquilibre de classes avec des approches axées sur les algorithmes en TLN. Pour chaque étude, le taux de déséquilibre est fourni (si l'étude le mentionne ou présente les éléments pour l'estimer). Le taux de déséquilibre peut être calculé utilisant l'équation 1.

$$\rho = \max_k\{C_k\} / \min_k\{C_k\} \quad (1)$$

Dans l'équation 1, ρ est le taux de déséquilibre et C_k est l'ensemble d'observations appartenant à la classe k . Ainsi, $\max_k\{C_k\}$ et $\min_k\{C_k\}$ représentent le nombre d'observations de la classe majoritaire et minoritaire, respectivement. Un taux de déséquilibre plus élevé représente un déséquilibre de classes plus sévère.

Dans le domaine du développement logiciel, Nnamoko et al. [2019] ont travaillé sur la prédiction de la criticité de bogues à partir de leur description. Dans leur étude, le jeu de données d'entraînement présentait un taux de déséquilibre de 45.5, suggérant que certains bogues arrivent plus souvent que d'autres. Pour mitiger le déséquilibre de classes, les auteurs ont proposé un arbre hiérarchique utilisant une stratégie d'apprentissage modifiée : l'arbre commence par apprendre à détecter la classe majoritaire contre les autres dans une classification binaire. Ensuite, il écarte la classe déjà apprise et il est entraîné à identifier la deuxième classe la plus nombreuse contre les classes restantes. Cette démarche se répète jusqu'à ce qu'il ne reste que deux classes, où un dernier arbre binaire est entraîné. Les résultats de cette étude ont montré que la démarche fournissait des résultats similaires à un entraînement standard.

L'étude de Wang et al. [2016] a proposé une nouvelle fonction de perte appelée la *fausse erreur quadratique moyenne* (mean false square error) pour mitiger l'effet du déséquilibre de classes. Les auteurs ont testé sa performance pour des réseaux de neurones profonds tant pour des applications de vision par ordinateur que de TLN. Ils ont artificiellement introduit du déséquilibre jusqu'à un taux de déséquilibre maximal de 20. Pour les applications en TLN, ils ont utilisé le jeu de données appelé *20 Newsgroup*, contenant des nouvelles classées en catégories par sujet. Les résultats suggèrent qu'utiliser la fonction de perte proposée a fourni de meilleurs résultats que le modèle de base, obtenu utilisant l'erreur quadratique moyenne.

Dans le domaine de la reconnaissance d'émotions, Li et al. [2019] ont cherché à reconnaître l'émotion des commentaires publiés par les utilisateurs d'un blog. Ils ont remarqué que les utilisateurs ont la tendance à partager plus de sentiments positifs que négatifs, résultant dans un taux de déséquilibre de 13.9. Pour mitiger l'effet du déséquilibre, les auteurs ont proposé une nouvelle fonction de perte inspirée de la PF et appelée *perte de correction de classe* (class correction loss). L'un des avantages de la fonction proposée, est qu'elle est adaptée à des cas de classification multiclasse, alors que la PF classique a été conçue pour des cas de classification binaire. La fonction proposée a été utilisée dans un réseau de neurones dense connecté à un CNN, un réseau récurrent de neurones et un Transformer pour générer des embeddings. Les résultats de l'étude suggèrent que l'approche proposée a mitigé le plus le déséquilibre, quand comparé à d'autres méthodes récentes de reconnaissance d'émotions.

Avec l'objectif final de générer automatiquement des romans, Iikura et al. [2021] ont exploré la segmentation automatique de paragraphes. Cela consiste à résoudre un problème de classification déterminant si deux phrases appartiennent ou pas au même paragraphe. Comme les romans ont la tendance à avoir plus de phrases que de paragraphes, les données d'entraînement présentaient un taux de déséquilibre de 5.63. Pour réaliser la classification, les auteurs ont utilisé de l'AT en ajustant un Transformer appelé BERT [Devlin et al., 2018]. Pour mitiger le déséquilibre de classes, la PF a été utilisée. Les

résultats obtenus avec la PF ont fourni de meilleurs résultats que les autres modèles entraînés avec l'entropie binaire croisée et sa version pondérée.

Dans l'industrie aéronautique, les travaux de Kao et al., [2018, 2019] ont abordé la reconnaissance automatique de pièces mentionnées dans des journaux de maintenance afin de les normaliser. En effet, il est courant de trouver que les opérateurs n'utilisent pas un langage standard pour nommer les pièces concernées dans leurs rapports. Par exemple, il a été trouvé que pour 10 termes génériques comme « circuit » ou « antenne », il existait un total de 418 variants dans le jeu de données [Kao et al., 2019]. Ainsi, les études ont eu pour but la reconnaissance de pièces dans le texte [Kao et al., 2018] et leur normalisation dans une forme canonique [Kao et al., 2019]. Encore dans l'aéronautique, Pham et al. [2020], ont exploré la reconnaissance automatique de texte à partir de rapports écrits à la main par les opérateurs. Ces rapports représentent un défi, ils sont souvent rédigés avec une calligraphie peu soignée et ils contiennent du bruit (tampons de validation, trame de la feuille, etc.). Pour reconnaître le texte, les auteurs ont employé une approche utilisant deux étapes successives : de la segmentation au moyen d'un R-FCN (*region-based fully convolutional networks*) et de la reconnaissance de texte avec un modèle appelé *CTCSeq2Seq*.

Enfin, l'étude d'Ansari [2020] a utilisé des systèmes experts pour supporter la prise de décision du planificateur de maintenance. Le but a été d'extraire automatiquement des éléments comme la charge de travail requise ou le coût estimé d'une intervention à partir de rapports écrits par des opérateurs. Malgré la validation de l'approche avec un démonstrateur, elle semble ne pas normaliser les noms des pièces, pouvant rendre difficile l'effectivité du système.

L'analyse d'autres études dans la littérature suggère que les applications trouvées en maintenance industrielle avec du TLN abordent rarement la problématique du déséquilibre de classes, même s'il s'agit d'un phénomène courant dans ce domaine. Dans des contextes différents à la manufacture, le déséquilibre de classes a été plus traité. Par exemple, la recherche réalisée par Iikura et al. [2021] est comparable à cet article, car ils utilisent aussi des Transformers avec la PF. Cependant, l'article ci-présent va plus loin en fournissant une vue détaillée du comportement du Transformer avec plusieurs configurations de la PF. De plus, cet article aborde la classification déséquilibrée des rapports de maintenance, ce qui peut s'avérer plus éprouvant que de travailler avec des romans. En effet, les jeux de données de TLN en maintenance imposent des défis particuliers, les opérateurs ont la tendance à être brefs au moment de rédiger leurs rapports. Cela peut dégrader la qualité des textes par rapport à des romans ou des commentaires sur des blogs, où une grammaire correcte est prioritaire.

3 METHODES ET MODELES UTILISES

3.1 Le jeu de données

Cette étude emploie un jeu de données industrielles provenant d'une société dont l'industrie et le nom ne sont pas mentionnés pour des raisons de confidentialité. Le jeu de données contient, après le nettoyage, 22709 observations. Chaque observation contient : le nom de l'équipement concerné fourni au moyen d'un champ de saisie en texte libre, une catégorie indiquant l'importance de l'équipement dans le processus de production, les symptômes rédigés par l'opérateur et un label indiquant si ce problème a arrêté le système de production ou pas.

Le seul prétraitement de données qui est réalisé pour chaque observation est de concaténer le nom de l'équipement, sa catégorie et les symptômes dans une seule chaîne de caractères. L'objectif est d'utiliser cette chaîne comme entrée unique pour la classer dans l'une des deux catégories suivantes : arrêt de production ou non-arrêt de production. La raison derrière cette concaténation est de montrer que les Transformers ont besoin de peu de prétraitement de données, même en travaillant avec des textes complexes. Quelques exemples d'observations sont fournis dans l'annexe I. Enfin, le taux de déséquilibre du jeu de données est de 10.5.

3.2 Transformer employé

Étant donné que les Transformers sont souvent préentraînés pour ensuite être ajustés à des tâches précises comme la classification de textes, cette étude utilise un modèle préentraîné appelé CamemBERT. Il a été créé par Martin et al. [2019] avec le but de traiter du texte en français. En fait, les Transformers fournissent de meilleurs résultats quand un modèle spécifique à une langue est utilisé au lieu de modèles génériques pour plusieurs langues. L'une des raisons expliquant cela est que les modèles spécifiques à une langue sont normalement entraînés sur plus de données, ce qui permet d'apprendre des représentations de données plus riches. Par exemple, CamemBERT a été entraîné sur 138Go de texte en français, contre 57Go utilisés pour la version générique multilingues de BERT [Martin et al., 2019]. Pour les expériences, la version de CamemBERT utilisée est celle disponible sur la librairie *Transformers* créé par Hugging Face [Wolf et al., 2019].

3.3 Fonctions de perte à comparer

Trois fonctions de perte seront comparées : l'entropie croisée (EC), sa forme pondérée (ECP) et la PF. Étant donné que l'EC peut être obtenue à partir de l'ECP, l'équation 2 présente cette dernière comme proposée par Ho et Wookey [2020]. Enfin, l'équation 3 présente la PF. Il est important de mentionner que cette étude utilise la forme pondérée de la PF, améliorant davantage sa performance sur des données déséquilibrées.

$$ECP = -\frac{1}{M} \sum_{k=1}^K \sum_{m=1}^M w_k * y_m^k * \log(h_k) \quad (2)$$

Sur l'équation 2, M est le nombre d'observations, K est le nombre de classes, w_k est le poids de la classe k , y_m^k est la sortie attendue de l'observation m appartenant à la classe k et h_k est la sortie du modèle après passage par la fonction softmax. Pour l'EC, le poids w_k de chaque classe est égal à 1.

$$PF = -\sum_{k=1}^K \sum_{m=1}^M w_k * (1 - h_k)^\gamma * y_m^k * \log(h_k) \quad (3)$$

Sur l'équation 3, M est le nombre d'observations, K est le nombre de classes, w_k est le poids de la classe k , y_m^k est la sortie attendue de l'observation m appartenant à la classe k , h_k est la sortie du modèle après passage par la fonction softmax et γ est le paramètre de concentration, avec $\gamma \geq 0$. Le terme $(1 - h_k)^\gamma$ est appelé le facteur de modulation : il donne un moindre poids aux observations faciles à classer et un poids plus élevé aux observations difficiles. Cela est déterminé en

fonction de la confiance du modèle. Pour les cas où le modèle est confiant de sa prédiction (exemple facile), $h_k \rightarrow 1$ résultant dans un facteur de modulation faible et donc qui réduit la perte. Dans le cas contraire, quand le modèle est peu confiant sur sa prédiction (exemple difficile), $h_k \rightarrow 0$ ce qui fournit un facteur de modulation élevé et donc une perte plus élevée. De ce fait, la PF aide le modèle à se concentrer sur les observations qui sont difficiles à classer pendant l'entraînement. Enfin, le paramètre de concentration augmente l'effet du facteur de modulation : des valeurs élevées de γ donneront plus d'importance aux exemples difficiles. Pour cette étude, l'effet de γ sera évalué avec $\gamma \in [0.5, 1, 2, 3, \dots, 9]$.

Sur l'équation 2 et 3, les poids w_k de chaque classe sont déterminés avec l'équation 4, car elle est la manière standard proposée sur la librairie de Python scikit-learn. Des travaux futurs de recherche évalueront d'autres heuristiques pour trouver ces poids.

$$w_k = M / (K * |C_k|) \quad (4)$$

Sur l'équation 4, M est le nombre d'observations, K est le nombre de classes et $|C_k|$ est le nombre d'instances de la classe k . Ainsi, w_k sera plus élevé pour les classes minoritaires et moins élevé pour les classes majoritaires.

3.4 Politique d'entraînement

Les données ont été divisées en 75 % et 25 % pour l'entraînement et validation, respectivement. Puis, l'ensemble d'entraînement a été divisé en 90 % pour l'entraînement effectif (données d'entraînement réduites) et 10 % pour la validation. Avec les données d'entraînement réduites, dix modèles utilisant la PF avec des valeurs différentes du paramètre de concentration γ seront entraînés et leur performance sera mesurée sur les données de validation afin de choisir le meilleur modèle. Ensuite, ce meilleur modèle est réentraîné en utilisant toutes les données d'entraînement. Après, il sera comparé avec un modèle utilisant l'EC et un autre employant l'ECP.

Pour choisir le meilleur modèle, le Coefficient de Corrélation de Matthews (CCM) est utilisé. Le CCM varie entre -1 et 1, où 1 implique une classification parfaite, 0 une prédiction aléatoire et -1 une prédiction contraire aux valeurs attendues. Le CCM est une meilleure mesure pour évaluer la performance sur des données déséquilibrées [Usuga Cadavid et al., 2020b]. Pour l'analyse, la spécificité et la sensibilité sont aussi fournies. Dans ce cas, la spécificité et la sensibilité mesurent le taux d'observations correctement classées pour le cas de non-arrêt et d'arrêt de la production, respectivement. La nomenclature suivante sera employée pour les modèles :

1. Cam_Base : CamemBERT ajusté utilisant l'EC. Ce modèle est celui de référence.
2. Cam_Pond : CamemBERT ajusté utilisant l'ECP.
3. PF_Cam_ γ : CamemBERT ajusté utilisant la PF avec des valeurs de γ différentes. Par exemple, si $\gamma = 6$, le nom du modèle est PF_Cam_6.

3.5 Réduction de dimensions pour la visualisation

Afin de contribuer au troisième objectif de cette étude qui est de fournir une meilleure comparaison entre les trois fonctions de perte, une visualisation en faibles dimensions est fournie pour les embeddings appris par les modèles. En fait, les

embeddings produits par les Transformers ont souvent plusieurs centaines de dimensions. Par exemple, CamemBERT produit des embeddings en 768 dimensions pour un document. Pour les réduire, l'ACP sera employée. Brièvement, l'ACP permet de trouver de nouvelles dimensions orthogonales entre elles qui retiennent la quantité la plus élevée de variance. Si les données sont projetées dans un espace avec moins de dimensions que celles du jeu de données initial, l'ACP réalise de la compression de dimensions [Jolliffe, 2011; Géron, 2019].

Pour cette étude, les embeddings pour l'ensemble de données de test sont générés avec chaque modèle utilisant une fonction de perte différente. Puis, les embeddings sont projetés dans un espace bidimensionnel à l'aide de l'ACP et leurs traces de densité sont analysées. En outre, les embeddings pour un modèle de CamemBERT sans ajustement sont aussi montrés afin de montrer l'effet de l'ajustement dans l'AT.

4 RESULTATS

Les modèles ont été implémentés avec PyTorch et entraînés sur Google Colab, avec deux cartes graphiques : une Tesla P100-PCI-E-16GB et une Tesla V100-SXM2-16GB, en fonction de leur disponibilité. Afin de considérer la nature stochastique des résultats, chaque modèle a été entraîné 20 fois et le meilleur modèle a été choisi par son CCM médian.

4.1 Sélection du meilleur paramètre de concentration

La figure 1 présente les boîtes à moustaches pour le CCM, la spécificité et la sensibilité de validation pour les 10 modèles entraînés avec la PF. Le meilleur modèle est entouré avec un carré rouge. Les valeurs pour le CCM, la spécificité et la sensibilité médians sont aussi fournies.

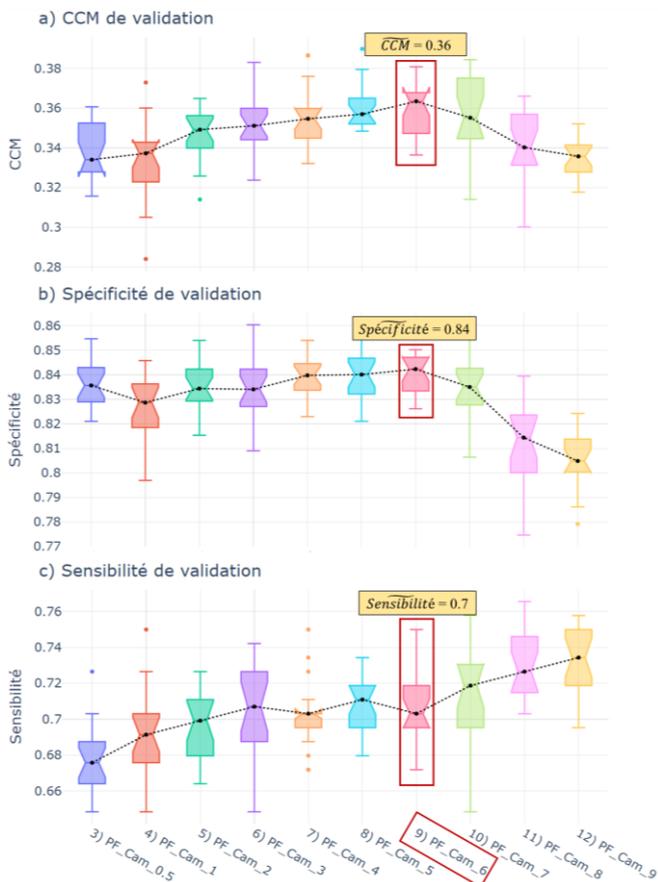


Figure 1. CMM (a), spécificité (b) et sensibilité (c) de validation pour des valeurs du paramètre de concentration explorées.

Les résultats suggèrent que le meilleur modèle, par CCM médian, est celui obtenu avec un paramètre de concentration $\gamma = 6$. La figure 1(b) et 1(c) montrent que $\gamma = 6$ fournit le meilleur équilibre entre la spécificité et la sensibilité, résultant dans le CCM le plus élevé. Après ce point, quand γ est augmenté, la modèle commence à donner trop d'importance à la classe minoritaire, ce qui améliore substantiellement la sensibilité. Néanmoins, cela dégrade significativement la spécificité, entraînant aussi une dégradation de la performance globale du modèle avec un CCM qui se réduit.

Il paraît qu'ajuster le paramètre de concentration γ est fondamental quand la PF est utilisée. Cela peut être un inconvénient : plus d'hyperparamètres à ajuster impliquent plus de charges de travail pour l'humain afin de trouver un bon modèle. De ce fait, des travaux futurs de recherche vont explorer s'il existe des rangs acceptables pour le paramètre de concentration dans des jeux de données de TLN en maintenance. En effet, il paraît que les résultats varient peu entre $\gamma \in [2, 6]$. Cela suggère qu'il peut y avoir des rangs de γ fournissant des résultats statistiquement similaires. Par la suite, le meilleur modèle (PF_Cam_6) sera comparé avec les autres modèles utilisant les deux autres fonctions de perte.

4.2 Comparaison de la PF avec d'autres fonctions de perte

La figure 2 présente les boîtes à moustaches pour le CCM, la spécificité et la sensibilité de test pour chacune des trois fonctions de perte comparées. Le meilleur modèle est entouré avec un carré rouge. Les valeurs pour le CCM, la spécificité et la sensibilité médians sont aussi fournies.



Figure 2. CMM (a), spécificité (b) et sensibilité (c) de test pour les fonctions de perte explorées.

Les résultats sur l'ensemble de tests suggèrent que le meilleur modèle, par CCM médian, est le modèle de référence : Cam_Base. Le CCM est une mesure équilibrée employée pour évaluer la performance dans des cas de déséquilibre de classes. Le CCM considère que toutes les classes ont la même importance. Dans ce cas, les approches axées sur les algorithmes s'avèrent moins efficaces que le modèle de base, sans modifications.

Dans le cas où la priorité est de détecter la classe minoritaire, la PF a des avantages sur les deux autres fonctions de perte. Cela peut être constaté dans la figure 2(c), où la sensibilité du modèle utilisant la PF (PF_Cam_6) est la plus élevée. Dans le contexte de cette étude, cela a du sens sous l'hypothèse que ne pas détecter une panne de machine qui arrêtera la production est le cas de figure le plus nuisible. De plus, la PF améliore la capacité de détection de la classe minoritaire plus que l'ECP, étant cette dernière une approche classique pour traiter le déséquilibre de classes. Néanmoins, il est important de remarquer que la PF a la tendance à dégrader la capacité du modèle à détecter la classe majoritaire. Cela peut être observé dans la figure 2(b). Sous certaines circonstances, cette situation est désavantageuse, car des problèmes n'arrêtant pas la production déclencheraient une fausse alarme d'arrêt.

Pour comparer comment chaque fonction de perte influence l'apprentissage du Transformer, la figure 3 fournit une visualisation des embeddings appris sous la forme de traces de densité pour les deux composantes principales obtenues au moyen de l'ACP.

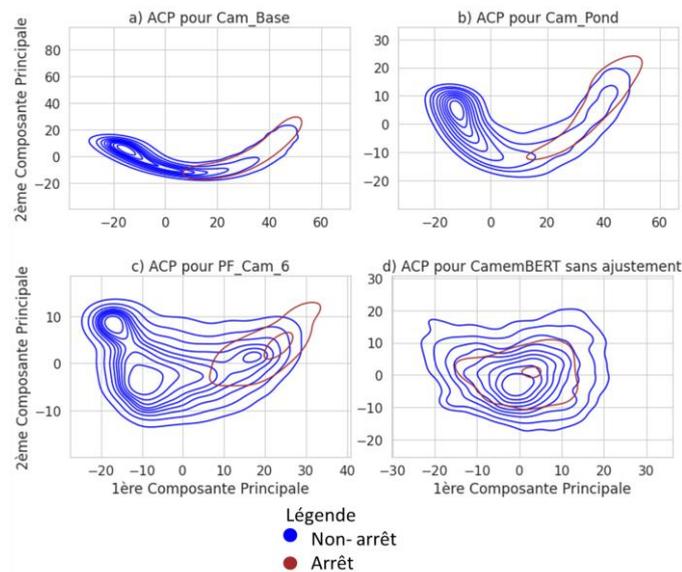


Figure 3. Traces de densité pour CamemBERT utilisant l'EC (a), l'ECP (b), la PF avec un paramètre de concentration de 6 (c) et CamemBERT sans ajustement (d)

La figure 3(d) montre que les embeddings initiaux du modèle sans ajustement ne sont pas représentatifs pour la tâche de classer les descriptions de maintenance : les deux classes se chevauchent sur des zones à haute densité. Cela met en avant l'importance de l'ajustement pour l'AT. Sur la figure 3(a), le modèle est ajusté avec l'EC. Dans ce cas, le modèle commence à essayer de créer une séparation entre les deux classes. Néanmoins, le déséquilibre de classes termine par générer un chevauchement de la classe minoritaire avec des régions denses de la classe majoritaire. Quand des approches axées sur les algorithmes sont appliquées, le Transformer sépare mieux les deux classes. Il est intéressant de constater que le modèle

utilisant la PF est le seul étant capable de générer une zone dense avec des observations de la classe minoritaire relativement loin des zones denses de la classe majoritaire. Cela peut expliquer sa sensibilité élevée. Cette distribution peut être observée dans le coin supérieur droit de la figure 3(c).

5 CONCLUSIONS

Cet article a exploré la classification de rapports se trouvant dans des journaux de maintenance à partir de la description textuelle fournie par les opérateurs. Plus spécifiquement, le but a été de déterminer si une description d'un problème sur une machine mènerait à un arrêt de production ou pas. Étant donné que les données de maintenance présentent des distributions déséquilibrées ayant majoritairement de cas qui n'arrêtent pas la production, une fonction de perte appelée PF et issue du domaine de la vision par ordinateur a été utilisée pour mitiger l'effet du déséquilibre de classes.

Pour évaluer l'influence de la PF sur des données déséquilibrées de TLN, elle a été utilisée sur un Transformer : un type d'architecture récente de réseau de neurones permettant l'AT en TLN. De plus, l'influence du paramètre de concentration, un hyperparamètre appartenant à la PF, a été évalué. Les résultats suggèrent qu'il doit être ajusté pour chaque cas d'application, car les résultats peuvent varier significativement avec des valeurs différentes du paramètre.

Dernièrement, la PF a été comparée avec deux autres fonctions de perte : l'EC et l'ECP. Les résultats ont montré que quand les deux classes sont aussi importantes, les approches axées sur des algorithmes évaluées dans cette étude n'ont pas pu améliorer la performance au-delà de celle du modèle de référence. Néanmoins, quand l'objectif est d'améliorer la capacité de détection de la classe minoritaire, qui correspondait à des problèmes menant à un arrêt de production, la PF a fourni les meilleurs résultats. Cependant, elle a aussi dégradé la capacité de détection de la classe majoritaire.

Des travaux futurs de recherche vont focaliser sur deux axes : premièrement, explorer l'utilisation de la PF sur d'autres jeux de données de TLN en maintenance ayant d'autres taux de déséquilibre. Le but sera de déterminer s'il existe des rangs standards du paramètre de concentration qui fournissent des bons résultats. Cela permettrait de gagner du temps au moment de créer des modèles. Secondement, explorer des approches heuristiques alternatives pour trouver les poids des classes utilisées dans des fonctions de perte comme l'ECP et la PF.

6 REMERCIEMENTS

Cette étude a été financée par le partenariat entre l'Association Nationale de la Recherche et de la Technologie (ANRT) et la société iFAKT France SAS, sous le dossier numéro 2018/1266.

7 ANNEXES

7.1 Annexe 1 : exemple des observations du jeu de données.

Pour des raisons de confidentialité, quelques éléments dans le texte ne sont pas affichés. Ils sont représentés par la marque [C] (caché) avec un commentaire sur l'information omise pour plus de clarté. Les fautes d'orthographe ont été laissées comme trouvées sur le jeu de données. Par exemple, il est possible d'observer que dans les exemples 2 et 3, les opérateurs utilisent les termes « HS » et « KC » pour dire « Hors Service » et « Cassé », respectivement.

Tableau 1. Exemple d'observations du jeu de données

ID	Texte concaténé	Label
1	[C – nom de la machine et catégorie]. Continu à se déplacer en mouvement transversal. Coupure du sectionneur et étiquette rouge sur boîtier de commande mise par [C – nom de l'opérateur] car jugé très dangereux.	Arrêt
2	[C – nom de la machine et catégorie]. IMPORTANT. LAME DE SCIE HS	Arrêt
3	[C – nom de la machine et catégorie]. BOUTON DE MISE SOUS TENSION KC [C – nom de l'opérateur]	Arrêt
4	[C – nom de la machine et catégorie]. PROBLEME DIMENSIONNELLE APRES COUPE. Lors de la coupe d'une plaque de 2 m , il y a une différence de 2mm d'un bout à l'autre. Différence entre l'affichage et la coupe réelle	Non-arrêt

8 REFERENCES

Ansari, F. (2020) Cost-based text understanding to improve maintenance knowledge intelligence in manufacturing enterprises, *Computers and Industrial Engineering*, 141. doi: 10.1016/j.cie.2020.106319.

Devlin, J., Chang, M.-W., Lee, K. et Toutanova, K. (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *arXiv e-prints*, p. arXiv:1810.04805.

Géron, A. (2019) *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition*. O'Reilly Media, Inc.

Ho, Y. et Wooley, S. (2020) The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling, *IEEE Access*, 8, pp. 4806–4813. doi: 10.1109/ACCESS.2019.2962617.

Iikura, R., Okada, M. et Mori, N. (2021) Improving bert with focal loss for paragraph segmentation of novels, in *Advances in Intelligent Systems and Computing*. L'Aquila, Italy, 7-10 October 2020, pp. 21–30. doi: 10.1007/978-3-030-53036-5_3.

Johnson, J. M. et Khoshgoftaar, T. M. (2019) Survey on deep learning with class imbalance, *Journal of Big Data*, 6(1), p. 27. doi: 10.1186/s40537-019-0192-5.

Jolliffe, I. (2011) Principal Component Analysis, in Lovric, M. (ed.) *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1094–1096. doi: 10.1007/978-3-642-04898-2_455.

Kao, A., Niraula, N. B. et Whyatt, D. (2018) PANDA - Discovering Part Name in Noisy Text Data, in *2018 IEEE International Conference on Prognostics and Health Management, ICPHM 2018*. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICPHM.2018.8448708.

Kao, A., Niraula, N. B. et Whyatt, D. (2019) Part name normalization, in *2019 IEEE International Conference on Prognostics and Health Management, ICPHM 2019*. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICPHM.2019.8819386.

Kusiak, A. (2019) Fundamentals of smart manufacturing: A multi-thread perspective, *Annual Reviews in Control*, 47, pp. 214–220. doi: 10.1016/j.arcontrol.2019.02.001.

Li, R., Si, Q., Fu, P., Lin, Z., Wang, W. et Shi, G. (2019) A multi-channel neural network for imbalanced emotion recognition, in *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*. Portland, OR,

USA, 4-6 November 2019, pp. 353–360. doi: 10.1109/ICTAI.2019.00057.

Lin, T.-Y., Goyal, P., Girshick, R., He, K. et Dollár, P. (2017) Focal Loss for Dense Object Detection, *arXiv e-prints*, p. arXiv:1708.02002.

Martin, L., Muller, B., Ortiz Suárez, P. J., Dupont, Y., Romary, L., Villemonte de la Clergerie, É., Seddah, D. et Sagot, B. (2019) CamemBERT: a Tasty French Language Model, *arXiv e-prints*, p. arXiv:1911.03894. Available at: <https://ui.adsabs.harvard.edu/abs/2019arXiv191103894M>.

Nnamoko, N., Cabrera-Diego, L. A., Campbell, D. et Korkontzelos, Y. (2019) Bug Severity Prediction Using a Hierarchical One-vs.-Remainder Approach, in Metais E. Meziane F., V. S. S. V. S. M. (ed.) *24th International Conference on Applications of Natural Language to Information Systems*. Salford, UK, 26–28 June 2019: Springer Verlag, pp. 247–260. doi: 10.1007/978-3-030-23281-8_20.

Pan, S. J. et Yang, Q. (2010) A survey on transfer learning, *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345–1359. doi: 10.1109/TKDE.2009.191.

Pham, H. et al. (2020) Robust Handwriting Recognition with Limited and Noisy Data, in *Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR*. Institute of Electrical and Electronics Engineers Inc., pp. 301–306. doi: 10.1109/ICFHR2020.2020.00062.

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R. et Fortin, A. (2019) Estimation of Production Inhibition Time Using Data Mining to Improve Production Planning and Control, in *2019 International Conference on Industrial Engineering and Systems Management (IESM)*. Shanghai, China, 25-27 September 2019, pp. 1–6. doi: 10.1109/IESM45758.2019.8948129.

Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R. et Fortin, A. (2020a) Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0, *Journal of Intelligent Manufacturing*, 31(6), pp. 1531–1558. doi: 10.1007/s10845-019-01531-7.

Usuga Cadavid, J. P., Grabot, B., Lamouri, S., Pellerin, R. et Fortin, A. (2020b) Valuing free-form text data from maintenance logs through transfer learning with CamemBERT, *Enterprise Information Systems*, (In Press), pp. 1–29. doi: 10.1080/17517575.2020.1790043.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. et Polosukhin, I. (2017) Attention is All you Need, in Guyon, I. et al. (eds) *Advances in Neural Information Processing Systems 30*. Long Beach, CA, USA, 4-9 December 2017: Curran Associates, Inc., pp. 5998–6008. Available at: <http://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>.

Wang, S., Liu, W., Wu, J., Cao, L., Meng, Q. et Kennedy, P. J. (2016) Training deep neural networks on imbalanced data sets, in *Proceedings of the International Joint Conference on Neural Networks*. Vancouver, BC, Canada, 24-29 July: Institute of Electrical and Electronics Engineers Inc., pp. 4368–4374. doi: 10.1109/IJCNN.2016.7727770.

Wolf, T. et al. (2019) HuggingFace's Transformers: State-of-the-art Natural Language Processing, *arXiv e-prints*, p. arXiv:1910.03771.

CIGI 2021

Une approche pour la formalisation des connaissances en PHM : application aux PME

NABIL OMRI^{1,2}, ZEINA AL MASRY¹, NICOLAS MAIROT², SYLVIAN GIAMPICCOLO² AND NOUREDDINE ZERHOUNI¹

¹ FEMTO-ST institute, Univ. Bourgogne Franche-Comté, CNRS, ENSMM,
24 rue Alain Savary, Besançon cedex, 25000, France

² SCODER
1 rue de la Forêt Z.A. l'Orée du Bois, Pirey 25480, France

Résumé – Le Prognostics and Health Management (PHM) est un concept émergent basé sur le management des données industrielles. La mise en œuvre du PHM dans les petites et moyennes entreprises (PME) est actuellement limitée en raison de la difficulté d'accès aux données. Cet écueil est surmonté en formalisant les connaissances des opérateurs et l'intégrer dans le système d'information des PME. Ainsi, la mise en place du PHM sera basée sur ce système d'information associant les données et les connaissances. À cette fin, nous proposons une approche PHM collaborative (X-PHM) pour assurer l'extraction des connaissances des opérateurs et leur intégration dans le processus PHM. Les décisions issues de cette approche sont restituées avec un souci d'intelligibilité. Ce papier propose une couche d'intégration des connaissances tout en se concentrant sur le management des données textuelles éditées par les opérateurs. Ce nouveau framework est appliqué pour prédire la durée de vie prévisionnelle (RUL) d'une presse mécanique dans une étude de cas réelle.

Abstract – Prognostics and Health Management (PHM) is an emerging concept based on industrial data management. PHM implementation in small and medium-sized enterprises (SMEs) is currently limited due to data accessibility difficulties. This pitfall is overcome by formalizing the operators' knowledge and integrating it into the SME's information system. Thus, the implementation of the PHM will be based on this information system associating data with knowledge. To this end, we propose a collaborative PHM approach (X-PHM) to ensure the extraction of operators' knowledge and its integration into the PHM process. The decisions resulting from this approach are restituted with a concern of explicability. This paper details the knowledge integration layer while focusing on managing the textual data edited by the operators. This new framework is applied to predict the remaining useful life (RUL) of a mechanical press in a real case study.

Mots clés - PHM, Traitement automatique du langage, PME, Analyse de données, Prédiction.

Keywords – PHM, Automatic language processing, PME, Data analysis, Prediction.

1 INTRODUCTION

Le domaine industriel a évolué rapidement au fil du temps à travers les différentes révolutions industrielles, de la machine à vapeur aux technologies du cloud computing. Tout au long de ces changements, les petites et moyennes entreprises (PME) ont toujours joué un rôle essentiel dans le commerce mondial. Les PME sont souvent définies en fonction de leur nombre d'employés et de leur chiffre d'affaires. Cependant, ces définitions varient d'un pays à l'autre [Moëuf et al., (2018)]. En France, les PME sont définies par la loi de modernisation de l'économie du 4 août 2008 comme l'ensemble des entreprises qui emploient plus de 10 et moins de 250 personnes, et qui réalisent un chiffre d'affaires annuel inférieur à 50 millions d'euros ou un bilan total n'excédant pas 43 millions d'euros [France-industrie, (2020)].

Les PME jouent un rôle très important dans l'économie française. En effet, elles représentent 90 % des entreprises françaises avec une employabilité de 6,3 millions de salariés et 43 % de valeur ajoutée [France-industrie, (2020)]. Malgré ces chiffres impressionnants, les PME françaises, comme toutes les PME du monde, rencontrent plusieurs difficultés sur la voie

de leur transformation digitale. La transformation digitale dans les PME est limitée essentiellement par deux facteurs [Omri et al., (2020)] : l'indisponibilité des ressources et l'organisation informelle dans les PME. Pour réussir une transformation digitale, des ressources spécifiques sont requises, plus particulièrement pour la collecte et l'analyse des données. Dans ce contexte et d'un point de vue technologique, seulement 5 % des équipements des entreprises sont à commande numérique [Pecht, (2009)]. Ce problème est principalement lié à l'absence de technologies de mesure adéquate pour répondre aux besoins de l'environnement manufacturier. Dans ce contexte, les PME peuvent être considérées comme un marché non desservi [Pecht, (2010)]. En outre, l'organisation des PME est moins formelle et la communication entre la direction et le personnel est étroite [Durst & Bruns, (2018)]. Ainsi, les informations sur le processus de production sont limitées dans le savoir-faire des opérateurs sans aucune documentation.

Certes les données digitalisées, qui sont la matière première pour les technologies basées sur l'intelligence artificielle (IA), sont peu disponibles dans les PME. Pourtant, les PME disposent d'une grande quantité de données sous forme d'un

remarquable savoir-faire qui peuvent être exploitées pour réussir le transfert digital. L'avantage dans les PME est que les opérateurs sont proches de l'outil, ce qui leur permet d'identifier leurs besoins et de les prioriser en termes d'urgence. La polyvalence du personnel leur permet d'analyser chaque problème par rapport à son impact sur l'ensemble du processus de production. Ainsi, la phase de modélisation du problème et d'identification des variables nécessaires à la résolution semble plus facile et plus efficace dans le cas des PME, ce qui réduit les coûts de mise en œuvre. Le véritable défi est de nature pédagogique pour convaincre le personnel de partager son savoir-faire avec un ensemble de technologies mystérieuses de type boîte noire.

Dans ce travail, nous proposons d'utiliser le Prognostics and Health Management (PHM) comme un cadre de travail pour assurer l'extraction des connaissances des opérateurs et leur intégration dans le système d'information des PME. Le PHM est une science qui étudie l'état de santé d'un équipement et prédit son évolution future (voir Figure 1). Ce concept permet de mieux contrôler les systèmes et de mettre en place des stratégies de maintenance adaptées [Pecht, (2009)]. Dans [Omri et al., (2019)], les auteurs définissent la PHM comme "un ensemble d'outils qui peuvent être utilisés en cascade ou séparément pour surveiller l'état de santé d'un système, prévoir son évolution future et/ou optimiser les décisions".

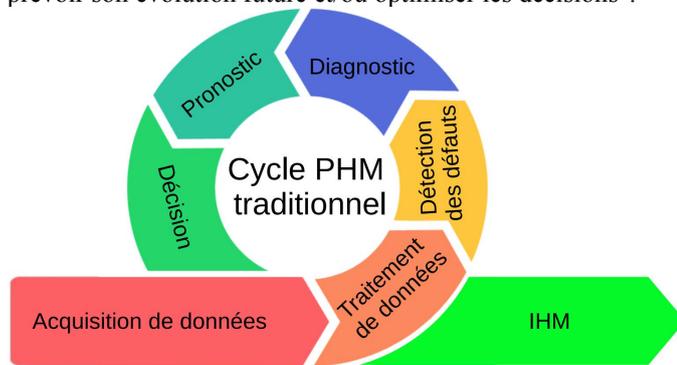


Figure 1. Le cycle PHM traditionnel.

Dans [Pecht, (2010)], l'auteur affirme que le PHM peut être mis en œuvre en utilisant des approches basées sur des modèles ou des données. La première approche consiste à construire des modèles analytiques qui sont directement liés aux processus physiques qui influencent l'état de santé des systèmes. Ainsi, une bonne compréhension du processus physique de dégradation et d'interaction des composants est nécessaire. La seconde approche consiste à utiliser des données de suivi historiques pour modéliser l'évolution du système

la qualité des données historiques [Omri et al., (2021)]. Une approche PHM vise à comprendre le comportement du système et à prédire son évolution future. Pour ce faire, les données décrivant le fonctionnement du système sont collectées et analysées afin d'extraire les connaissances nécessaires. En revanche, ces données présentent souvent des problèmes de qualité qui empêchent la satisfaction des objectifs. Pour éviter ce problème, le système d'information des PME est renforcé par les connaissances du personnel selon une démarche d'analyse de données informée. La Figure 2 détaille ce processus. En effet, le PHM orienté donné est basé sur les données collectées et qui décrivent le fonctionnement du système pour modéliser son comportement et prédire son évolution future. Dans ce contexte, l'expertise technique des opérateurs est mise à profit pour optimiser ce processus et améliorer ses résultats. Cependant, l'intégration de cette expertise est souvent implicite, non discutée, le mode d'intégration n'est pas formalisé dans le contexte du PHM.

L'objectif d'une telle approche est d'améliorer les performances du système et d'enrichir le système d'information de l'entreprise à travers des résultats d'analyse explicables. Cependant, deux problèmes se manifestent : (i) le traitement de la problématique des données industrielles et (ii) la formalisation et l'intégration de la connaissance humaine dans le processus d'analyse.

La problématique des données industrielles et leurs nettoyages ont été discutés dans plusieurs travaux de la littérature. Dans le prolongement de ces travaux, nous proposons ici de nous concentrer sur la connaissance (savoir-faire) du personnel. Le facteur humain est un élément important dans toute entreprise, quelle que soit sa taille. En revanche, dans le cas des PME le facteur humain est plus qu'important. En effet, les employés représentent le savoir-faire de l'entreprise et sont responsables de son développement et de son partage. En outre, la connaissance du personnel est une source importante de données qui peut être utilisée pour guider tout projet de numérisation. Cependant, deux problèmes principaux se posent dans ce contexte. Le premier est de nature pédagogique pour convaincre le personnel de partager son savoir-faire avec un ensemble de technologies de « type boîte noire » (c'est-à-dire qui n'explique pas son principe de fonctionnement et son raisonnement). Le second problème concerne la manière d'intégration du savoir-faire du personnel dans les outils d'analyse de données.

Nous sommes confrontés ici à deux problèmes complexes dans le domaine de l'analyse des données. Comme ils sont nommés

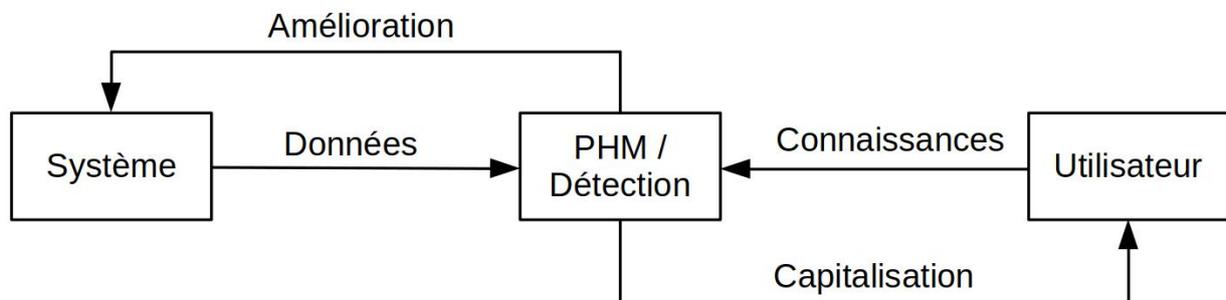


Figure 2. Problématique du PHM collaboratif.

jusqu'à ce qu'une défaillance se produise. Dans ce cas, la compréhension du processus physique du système ne pourrait pas être nécessaire, mais les résultats dépendent uniquement de

dans la littérature, ces problèmes sont les outils d'apprentissage automatique explicables [Guidotti et al., (2018)] et informés [von Rueden et al., (2019)]. Le terme

d'explication fait référence à la capacité de l'algorithme d'apprentissage automatique utilisé à expliquer les résultats et les modèles obtenus. Cela rend l'algorithme plus transparent et crée donc une relation de confiance entre lui et ses utilisateurs. Alors que « l'apprentissage automatique informé » se réfère à l'intégration des connaissances humaines dans le processus d'apprentissage afin de l'améliorer. Pour résumer, ce travail vise à développer une approche PHM collaborative (X-PHM) pour assurer l'extraction des connaissances des opérateurs et leur intégration dans le processus PHM. Les décisions issues de cette approche sont restituées avec un souci d'intelligibilité. Cependant, dans ce papier nous détaillons seulement la couche de formalisation et d'intégration des connaissances tout en se concentrant sur le management des données textuelles éditées par les opérateurs.

Le reste de ce document est organisé comme suit. La section 2 présente un état de l'art sur la formalisation des connaissances humaines et la gestion des données textuelles. La section 3 détaille le framework X-PHM proposé. Ce framework est appliqué à une étude de cas réelle dont les résultats sont présentés dans la section 4. Les conclusions et les perspectives de ce travail sont détaillées dans la section 5.

2 ÉTAT DE L'ART

Les connaissances humaines sont une source importante d'informations qui devraient être utilisées pour enrichir le système d'information des PME. Dans ce contexte, plusieurs approches de formalisation des connaissances humaines sont proposées [von Rueden et al., (2019)]. Cette section propose une revue de la littérature de ces approches tout en détaillant les différents types de connaissances humaines et les moyens de les formaliser pour une analyse de données informées.

2.1 Types de connaissances humaines

L'être humain interagit avec son environnement et en extrait des connaissances. De nombreuses études ont été menées dans la littérature pour caractériser ces connaissances sous différents points de vue (sociologie, psychologie, etc.). Ces études proposent plusieurs manières de catégoriser les connaissances humaines. Conformément à ces catégorisations, nous proposons ici de classer les connaissances humaines en trois types principaux : (i) Connaissances générales (ii) Lois scientifiques et (iii) Expertise.

- **Connaissances générales** : ce type de connaissance humaine fait référence à des informations de base connues et accessibles à toute l'humanité [von Rueden et al., (2019)]. Ces connaissances sont développées par l'intuition humaine et transmises d'une génération à l'autre.
- **Lois scientifiques** : cette connaissance représente l'ensemble des faits et des lois explicitement validés par des expériences. Habituellement, ces lois sont représentées sous la forme d'équations mathématiques qui relient les faits aux conséquences.
- **Expertise** : ce dernier type de connaissance humaine reflète le savoir-faire particulier développé par une personne dans un domaine particulier. Habituellement, cette sagesse se rapporte au domaine de travail dans lequel une personne passe beaucoup de temps. Les observations sont la principale source de ces

connaissances et peuvent être partagées entre les coéquipiers oralement ou sous forme de documents [Omri et al., (2019)]. Contrairement aux lois scientifiques, ces connaissances (si elles sont documentées) sont sauvegardées sous forme de textes.

2.2 Apprentissage informé et formalisation de la connaissance

L'apprentissage informé est défini comme la combinaison de données et de connaissances humaines pour un processus d'apprentissage plus efficace. Globalement, un processus d'apprentissage peut être divisé en trois étapes principales : (i) Prétraitement des données (ii) Data mining et (iii) Résultats. Ainsi, les connaissances humaines peuvent être intégrées dans le processus d'apprentissage tout au long de ces étapes.

- **Prétraitement informé des données** : cette tâche est divisée en plusieurs étapes telles que le nettoyage des données, la transformation des données et la sélection des variables. Dans ce contexte, les connaissances humaines peuvent être utilisées dans ces différentes tâches. En effet, la connaissance préalable du problème étudié peut guider la phase de sélection des variables. En outre, l'opérateur peut évaluer l'exactitude des données collectées et l'améliorer en conséquence. De plus, il peut être considéré comme une seconde source de données en définissant une fonction mathématique pour générer de nouvelles variables à partir de variables existantes ou en générant de nouvelles données à l'aide de distributions statistiques [Ladick'y et al., 2015]. Enfin, un être humain peut alimenter un algorithme d'apprentissage à travers une collection d'observations et d'événements rares qui sont difficiles à formaliser dans les données d'entrée.
- **Exploration de données informée** : l'exploration de données consiste à apprendre à partir des données et à en extraire des connaissances. Pour ce faire, des algorithmes d'intelligence artificielle sont utilisés pour prédire les sorties d'un ensemble de variables d'entrée. Comme exemple d'algorithmes utilisables, nous citons les réseaux de neurones artificiels (RNA), la machine vectorielle de support (SVM), les arbres de décision (DT), etc. Au niveau du data mining, la connaissance humaine peut identifier l'ensemble de paramètres de l'algorithme utilisé [Zemouri et al., (2019)] tel que l'architecture d'un RNA ou la profondeur de DT, etc. De plus, une connaissance préalable du problème étudié peut être utile pour modifier la fonction de perte de l'algorithme [Bhandari et al., 2020].
- **Résultats éclairés** : les résultats sont l'objectif du processus d'apprentissage. Ils représentent un ensemble de règles liant les entrées aux sorties. Ces règles peuvent être implicites ou explicites, simples ou complexes et elles acceptent un certain ordre d'incertitude. À ce niveau, l'homme peut intervenir pour évaluer la cohérence des règles extraites en les comparant aux lois scientifiques existantes ou simplement en utilisant sa propre expertise. De plus, des règles connues qui ne sont pas apprises à partir des données peuvent être injectées dans les résultats finaux pour les documenter.

Chaque type d'apprentissage informé nécessite une formalisation différente de la connaissance. Cependant, il est encore difficile de communiquer la connaissance humaine à

une machine. À cet effet, de nombreuses approches pour formaliser les connaissances humaines sous des formes compréhensibles par machine sont développées dans la littérature. Nous proposons ici de classer ces approches en trois grandes catégories : (i) Modèles mathématiques (ii) Règles logiques et (iii) Relations statistiques.

- **Modèles mathématiques** : les modèles mathématiques formalisent les connaissances sous forme de fonctions ou d'équations différentielles. En fait, les fonctions relient un ensemble de variables à une sortie ou elles peuvent être utilisées comme contraintes pour réduire l'espace des solutions.

2.3 Traitement automatique des langues

Le traitement du langage naturel (TLN) couvre plusieurs sujets et implique des domaines multidisciplinaires, tels que la linguistique, l'informatique et l'intelligence artificielle. L'histoire du TLN remonte au début des années 1950 avec le test de Turing. Depuis les années 1980, le TLN s'est appuyée sur les statistiques et les probabilités jusqu'à ces dernières années, lorsque l'utilisation de l'apprentissage profond et des réseaux de neurones artificiels (RNA) a atteint des résultats de pointe dans de nombreuses tâches de TLN. Ces sujets incluaient l'extraction d'informations utiles, la traduction automatique, la synthèse de textes, la réponse aux questions, la classification des documents et la reconnaissance vocale [Ott

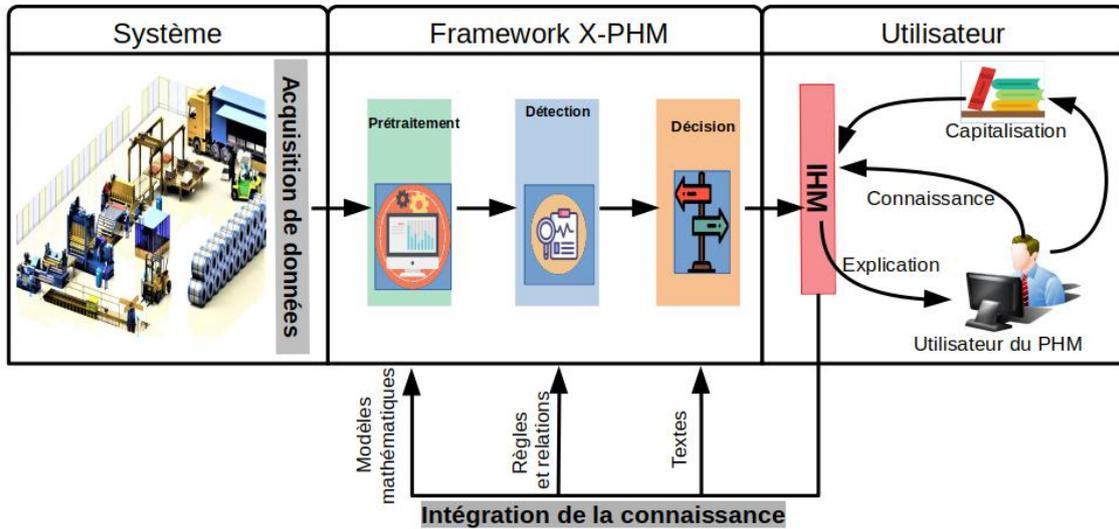


Figure 3. Explication de l'approche proposée.

- **Règles logiques** : au fil du temps, les connaissances humaines sont enregistrées sous forme de textes. Cependant, les phrases en langage naturel ne sont pas compréhensibles par les machines. Ainsi, la logique est utilisée pour transformer les faits et les conséquences représentés par des phrases en règles logiques formelles.
- **Relations statistiques** : les relations statistiques relient les variables de données les unes aux autres. En fait, les variables peuvent être corrélées ou elles peuvent être définies comme une distribution statistique. De plus, les relations statistiques peuvent faire référence à l'indépendance conditionnelle ou à la structure de corrélation de variables aléatoires ou même à une description complète des distributions de probabilité conjointes.

Idéalement, les modèles de formalisation de la connaissance sont les plus adaptés pour communiquer entre l'homme et les machines, car ils sont compréhensibles des deux côtés, mais toutes les connaissances humaines ne peuvent pas être modélisées selon ces formes. En effet, le domaine industriel se caractérise par une grande quantité de connaissances enregistrées sous forme de textes (rapports de suivi, procès verbaux de réunions, guides d'utilisation, etc.). L'enjeu est donc de pouvoir intégrer cette source de données dans le processus de gestion des données. Pour ce faire, plusieurs travaux ont été menés pour la valorisation des données textuelles dans le cadre de l'Industrie 4.0.

et al., (2018)]. Dans le contexte de l'Industrie 4.0, le TLN a été utilisé dans [Sharp et al., (2017)] pour extraire des informations supplémentaires sur le problème à partir des rapports de maintenance. Dans le même domaine [Usuga Cadavid et al. (2020)] appliquent un modèle TLN appelé camembERT pour prédire la criticité et la durée des pannes à partir de la description fournie dans les journaux de maintenance sous forme de données textuelles.

Dans ce travail, nous proposons d'intégrer la connaissance humaine dans un processus PHM collaborative (X-PHM) afin d'améliorer ses résultats et de surmonter le problème de manque de données dans les PME. La section suivante détaille le nouveau framework X-PHM.

3 APPROCHE PROPOSÉE

L'approche proposée consiste à développer une sorte de traducteur permettant de communiquer l'algorithme d'analyse de données avec son utilisateur. Cette communication consiste à intégrer les connaissances de l'utilisateur dans l'algorithme d'analyse des données d'un côté et à expliquer les décisions prises de l'autre côté.

Comme le montre la Figure 3, l'approche proposée consiste en un cadre interactif qui permet la communication entre l'utilisateur et l'algorithme d'apprentissage automatique. Cette interaction est représentée en termes (i) d'explication du processus de décision proposé et (ii) d'intégration de l'expertise de l'utilisateur dans la phase d'apprentissage. Ainsi, la connaissance humaine peut être validée à travers ce processus afin de la confirmer (ou la décliner). Les

connaissances capitalisées sont ensuite stockées dans une bibliothèque de règles qui seront utilisées à l'avenir pour soutenir la prise de décision et expliquer des phénomènes surprenants. Dans ce papier, nous nous limitons au processus d'intégration de la connaissance humaine, le reste de l'approche sera détaillé dans d'autres travaux.

Les connaissances humaines peuvent être intégrées dans le processus d'apprentissage tout au long de ses différentes étapes. Cependant, certaines tâches sont automatisées par des algorithmes sophistiqués pour faciliter la mission de l'utilisateur. En fait, des algorithmes évolutifs sont utilisés pour adapter l'architecture du RNA à la complexité du problème [Zemouri et al., (2019)]. Pour l'étape de sélection des variables, de nombreuses techniques sont développées et appliquées dans de nombreux domaines [Zemouri et al., (2018)]. En outre, d'autres méthodes sont développées pour apprendre efficacement à partir de données déséquilibrées ou pour évaluer automatiquement la cohérence des résultats. Pour cela, nous proposons d'intégrer les connaissances humaines dans la génération de nouvelles données et l'amélioration de leur qualité.

L'utilisateur est au centre du processus de préparation de données. En effet, il se charge de choisir les paramètres nécessaires. En outre, ses connaissances sont intégrées pour valider l'étape de génération et corriger les nouvelles données lorsqu'elles présentent des problèmes de précision ou de cohérence. De plus, l'utilisateur a une idée du problème, de ses causes et de ses conséquences et il lie généralement le problème à un ensemble de variables qui peut l'affecter (ou la variable qui n'a aucune influence). Ainsi, cette connaissance sera introduite dans l'algorithme sous forme de règles qui guideront la phase de préparation des données. Comme détaillé dans les sections précédentes, il existe plusieurs façons d'intégrer les connaissances humaines dans le processus d'apprentissage. Ici, nous proposons de valoriser les connaissances humaines dans la génération de nouvelles données et l'amélioration de leur qualité. Pour ce faire, nous proposons dans cette partie de valoriser la connaissance de l'utilisateur du PHM selon quatre manières :

- **Définition des nouvelles variables** : grâce à son expertise, l'utilisateur peut juger que la fusion d'un ensemble de variables peut générer une nouvelle variable plus pertinente pour la description du problème étudié. Ainsi, l'utilisateur est responsable de la définition de la fonction de calcul d'une telle variable.
- **Définition de l'importance des variables** : en continuité avec l'élément précédent, l'utilisateur a une idée du problème, de ses causes et de ses conséquences et il lie généralement le problème à un ensemble de variables qui peut l'affecter (ou la variable qui n'a aucune influence). Ainsi, cette connaissance sera introduite dans l'algorithme sous forme de règles qui guideront la phase de génération du voisinage en améliorant la qualité des nouvelles données et en augmentant leur similitude avec les réelles.
- **Description des variables** : il s'agit d'une description statistique permettant d'obtenir des informations sur chaque variable séparément. Nous distinguons deux types de variables : (i) les variables quantitatives (valeurs continues) et (ii) les variables qualitatives (le nombre de

valeurs possibles est limité). La description d'une variable quantitative est basée sur les statistiques suivantes : moyenne, médiane, variance, écart-type, quantiles, min et max. La description d'une variable qualitative est beaucoup plus sommaire. Une fois les modalités de la variable identifiées, il s'agit d'identifier le mode et d'étudier les proportions associées à chaque modalité. Bien entendu, ces informations sont calculables automatiquement. En revanche, dans le cas où nous n'avons pas d'historique des données enregistrées, la contribution de l'utilisateur devient indispensable pour avoir une idée complète de chaque variable.

- **Définition des relations statistiques** : les relations statistiques relient les variables de données les unes aux autres. En fait, les variables peuvent être corrélées ou elles peuvent être définies comme une distribution statistique. De plus, les relations statistiques peuvent faire référence à l'indépendance conditionnelle ou à la structure de corrélation de variables aléatoires ou même à une description complète des distributions de probabilité conjointes. Tout comme la description des variables, ces informations peuvent être calculées automatiquement. Cependant, nous nous intéressons aux cas où il n'y a pas d'historique des données enregistrées.

D'autre part, il reste un autre défi qui est la facilitation de cette tâche pour un utilisateur non spécialisé. Pour ce faire, nous proposons de formaliser ce travail dans une interface graphique ergonomique où l'utilisateur peut choisir entre plusieurs options pour communiquer ses connaissances. Cette interface permet à l'utilisateur de choisir entre des modèles mathématiques, des relations statistiques ou des règles logiques pour intégrer ses connaissances. En plus de ces formes d'intégration des connaissances, il existe des données textuelles éditées par les opérateurs sous forme de commentaires concernant des événements survenus pendant la production. Ainsi, un module de traitement des langages naturels est utilisé pour extraire des informations sur le processus de production qui renforcent les données collectées automatiquement par les capteurs installés sur le système.

4 CAS D'ÉTUDE ET RÉSULTATS

Le framework X-PHM proposé a été appliqué dans l'entreprise SCODER. SCODER est une PME française spécialisée dans le découpage des pièces automobiles à haute précision. Cette application vise à prédire la durée de vie résiduelle (RUL) d'une presse mécanique afin de planifier les actions de maintenance et ainsi augmenter sa productivité. L'application s'est déroulée en trois étapes principales: préparation des données, intégration des connaissances et prédiction du RUL.

4.1 préparation des données

Rappelons que la machine est équipée d'un système d'acquisition de données intégré qui contrôle les micro-arrêts de la machine. Ces données sont fusionnées avec l'historique d'intervention de l'équipe de maintenance pour prédire la RUL de la presse. La Figure 4 présente l'approche proposée qui consiste à équiper le personnel de maintenance par des tablettes afin d'enregistrer en temps réel toutes interventions sur la presse. Les données collectées sont ensuite traitées via un module de TLN afin d'extraire le type d'intervention, le composant concerné et le travail effectué. En outre, le même processus est appliqué afin de valoriser les commentaires de l'équipe de production. Les informations extraites concernent

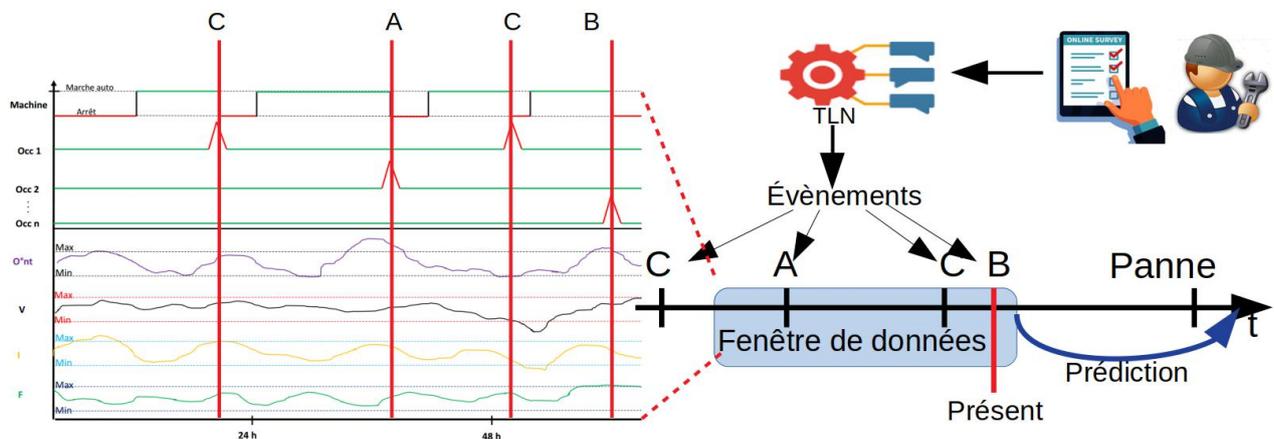


Figure 4. Illustration de l'approche de préparation de données.

essentiellement les événements extérieurs qui ne sont pas enregistrés par le système d'acquisition de données. Ces informations sont ensuite croisées avec les micro-arrêts enregistrés automatiquement par la machine.

4.2 Intégration des connaissances

Dans cette étape, et en utilisant le processus défini dans la section 3, le savoir-faire des opérateurs est utilisé à différents niveaux pour améliorer la qualité des données collectées:

- **Sélection des variables:** à partir des données collectées, un sous-ensemble de variables est sélectionné comme important. Pour des raisons de confidentialité, ces variables seront notées $\text{Var}_i, \{i = 1, \dots, I\}$.
- **Définition de nouvelles variables:** sur la base de l'expertise du personnel de la société SCODER, de nouvelles variables sont définies à partir des données existantes. Ces variables sont considérées comme plus pertinentes pour l'application.
- **Limites des variables:** à ce niveau, les connaissances humaines sont utilisées pour fixer les limites des variables afin d'éviter les valeurs aberrantes dans l'imputation des données manquantes.

4.3 Prédiction du RUL

La base de données résultante est analysée avec des algorithmes d'IA afin d'identifier le lien entre les micro-arrêts et les grandes pannes et ainsi prédire la prochaine défaillance. Ce travail a été réalisé avec l'équipe de maintenance qui a participé à la numérisation de l'historique d'interventions avant l'installation des tablettes. En outre, cette collaboration a permis d'identifier facilement les modes de défaillance et ainsi une capitalisation efficace de la connaissance.

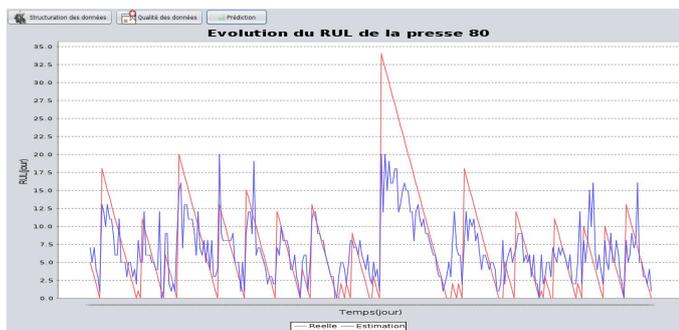


Figure 5. Évolution du RUL prédit versus le RUL réel. En rouge, la courbe d'évolution du RUL réel. En bleu, l'évolution du RUL prédit.

Plusieurs techniques d'IA ont été testées et comparées pour cette application. Dans ce contexte, l'algorithme de forêts aléatoires a donné les meilleurs résultats avec un RMSE de 0.3. La Figure 5 montre l'évolution du RUL prédit versus le RUL réel.

Cette application est implémentée dans un logiciel appelé "SCODER Data System (DS2)". DS2 est développé pour faciliter la tâche de prédiction pour les utilisateurs non spécialistes.

5 CONCLUSION

Ce travail présente un framework X-PHM qui permet d'assurer l'extraction des connaissances des opérateurs et leur intégration dans le système d'information des PME. Ainsi, la difficulté d'accès aux données dans les PME est surmontée à travers ce système d'information associant les données et les connaissances. Ce nouveau framework est appliqué pour prédire le RUL d'une presse mécanique dans une étude de cas réelle. Les résultats prouvent que la mise en place du PHM est possible même dans les PME avec des infrastructures de collecte de données peu développées.

Le travail de formalisation des connaissances présenté dans ce travail n'est qu'une partie d'un projet ambitieux qui vise à faciliter la transformation digitale au sein des PME. Dans ce contexte, les données textuelles sont une source d'information importante à valoriser. Ainsi, les perspectives concernent la valorisation de ces données même si elles sont « griffonnées » sur papier.

6 RÉFÉRENCES

Bhandari, B., Alsadoon, A., Prasad, P., Abdullah, S., and Haddad, S. (2020). Deep learning neural network for texture feature extraction in oral cancer : enhanced loss function. *Multimedia Tools and Applications*, 79(37) :27867–27890.

Durst, S. and Bruns, G. (2018). Knowledge management in small and medium-sized enterprises. In *The Palgrave Handbook of Knowledge Management*, pages 495–514. Springer.

France-industrie (2020). L'industrie en France c'est : Tous les chiffres clés. <https://www.franceindustrie.org/>, page (accessed : 04.12.2020).

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., and Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5) :1–42.

- Ladick`y, L., Jeong, S., Solenthaler, B., Pollefeys, M., and Gross, M. (2015). Data-driven fluid simulations using regression forests. *ACM Transactions on Graphics (TOG)*, 34(6) :1–9.
- Moeuf, A. (2018). Identification des risques, opportunités et facteurs critiques de succès de l'industrie 4.0 pour la performance industrielle des PME (Doctoral dissertation, Université Paris-Saclay (ComUE)).
- Omri, N., Al Masry, Z., Giampiccolo, S., Mairot, N., and Zerhouni, N. (2019). Data management requirements for phm implementation in smes. In *2019 Prognostics and System Health Management Conference (PHM-Paris)*, pages 232–238. IEEE.
- Omri, N., Al Masry, Z., Mairot, N., Giampiccolo, S., and Zerhouni, N. (2020). Industrial data management strategy towards an sme-oriented phm. *Journal of Manufacturing Systems*, 56 :23–36.
- Omri, N., Al Masry, Z., Mairot, N., Giampiccolo, S., & Zerhouni, N. (2021). Towards an adapted PHM approach: Data quality requirements methodology for fault detection applications. *Computers in Industry*, 127, 103414.
- Otter, D. W., Medina, J. R., & Kalita, J. K. (2018). A survey of the usages of deep learning in natural language processing. *CoRR abs/1807.10854* (2018).
- Pecht, M. (2009). Prognostics and health management of electronics. *Encyclopedia of Structural Health Monitoring*.
- Pecht, M. G. (2010). A prognostics and health management roadmap for information and electronics-rich systems. *IEICE ESS Fundamentals Review*, 3(4) :4_25–4_32.
- Sharp, M., Sexton, T., & Brundage, M. P. (2017, September). Toward semi-autonomous information. In *IFIP International Conference on Advances in Production Management Systems* (pp. 425-432). Springer, Cham.
- Usuga Cadavid, J. P., Grabot, B., Lamouri, S., Pellerin, R., & Fortin, A. (2020). Valuing free-form text data from maintenance logs through transfer learning with CamemBERT. *Enterprise Information Systems*, 1-29.
- von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Pfrommer, J., Pick, A., Ramamurthy, R., et al. (2019). Informed machine learning—a taxonomy and survey of integrating knowledge into learning systems. *arXiv preprint arXiv :1903.12394*.
- Zemouri, R., Omri, N., Devalland, C., Arnould, L., Morello, B., Zerhouni, N., and Fnaiech, F. (2018). Breast cancer diagnosis based on joint variable selection and constructive deep neural network. In *2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME)*, pages 159–164. IEEE.
- Zemouri, R., Omri, N., Fnaiech, F., Zerhouni, N., and Fnaiech, N (2019). A new growing pruning deep learning neural network algorithm (gp-dlnn). *Neural Computing and Applications*, pages 1–17.

**Session 4.4 - Session spéciale 13 :
L'excellence opérationnelle via la
digitalisation des processus industriels**

Chairman : Loubna Benabbou - Marie-Anne Le Dain

Algorithme génétique pour l'optimisation intégrée du contrôle de la maintenance et de la qualité dans les systèmes manufacturiers

IMEN HARBAOUI DRIDI¹, NADIA BAHRIA², ANIS CHELBI³, HANEN BOUCHRIHA⁴

¹ Université de Tunis El Manar, LACS
École Nationale d'Ingénieurs de Tunis (ENIT)
Institut National des Sciences Appliquées et de Technologie (INSAT)
imen.harbaoui@issatkr.rnu.tn

² Université de Tunis El Manar, OASIS
École Nationale d'Ingénieurs de Tunis (ENIT)
BP 37, le Belvédère 1002 Tunis - Tunisie
Ecole Polytechnique Internationale Privée de Tunis, PI-LAB
Rue du Lac d'Annecy, Tunis 1053
nadia.bahria@enit.utm.tn

³ Université de Tunis, CEREP
École Nationale Supérieure d'Ingénieurs de Tunis (ENSIT)
5 Av. Taha Hussein, 1008 Tunis – Tunisie
anis.chelbi@planet.tn

⁴ Université de Tunis El Manar, LACS
École Nationale d'Ingénieurs de Tunis (ENIT)
hanen.bouchriha@enit.rnu.tn

Résumé – Cet article présente une nouvelle approche basée sur les algorithmes génétiques pour l'optimisation intégrée de la maintenance et du contrôle qualité dans les systèmes manufacturiers. Le système de production étudié est considéré comme imparfait. Le contrôle de la qualité est effectué à l'aide d'une carte de contrôle x-bar. L'objectif est de déterminer les valeurs optimales de la taille de l'échantillon, l'intervalle d'échantillonnage, le nombre d'échantillons et les limites de surveillance et de contrôle de la carte de contrôle pour minimiser le coût total moyen par unité de temps. L'algorithme génétique proposé permet la génération des meilleures valeurs des cinq variables de décision minimisant la fonction objectif.

Abstract – This paper presents a new approach based on genetic algorithms for the integrated optimization of maintenance and quality control in manufacturing systems. The production system studied is considered imperfect. Quality control is done using an x-bar control chart. The objective is to determine the best values of sample size, sampling interval, number of samples, and monitoring and control limits of the control chart in order to minimize the average total cost per time unit. The proposed genetic algorithm allows the generation of the best values of the five decision variables minimizing the objective function.

Mots clés – Système manufacturier, Algorithme génétique, Optimisation, Maintenance, contrôle qualité.

Keywords – Manufacturing Systems, Genetic Algorithm, Optimization, Maintenance, Quality Control.

1 INTRODUCTION

Dans les systèmes manufacturiers, la gestion des opérations repose essentiellement sur trois fonctions principales : la gestion de la production, la planification des activités de maintenance et le contrôle de la qualité. La conception et l'optimisation conjointe de ces trois fonctions améliore l'efficacité des unités de production et réduit les coûts. Les unités de production sont généralement soumises à plusieurs processus de dégradation qui perturbent les activités de production et affectent la qualité des produits. Ainsi, la mise en œuvre de stratégies de maintenance préventive (MP) devient déterminante pour réduire la probabilité de panne et les pertes de production correspondantes.

De plus, un équipement mal entretenu peut produire des produits défectueux, entraînant des coûts de réparation.

Par conséquent, nous proposons une nouvelle approche, basée sur des algorithmes génétiques, intégrant les fonctions maintenance et qualité pour les systèmes de production imparfaits et produisant des unités défectueuses. Cette approche vise à minimiser le coût total moyen par unité de temps. Le contrôle de la qualité est effectué à l'aide d'une carte de contrôle x-bar. En fonction des résultats de la carte de contrôle, il est décidé de réaliser ou non une action de maintenance préventive ou corrective.

Le coût total est exprimé en fonction de cinq variables décisives, qui sont : la taille de l'échantillon, l'intervalle

d'échantillonnage, le nombre d'échantillons et les limites de surveillance et de contrôle de la carte de contrôle. Le but de l'approche développée est de déterminer les valeurs optimales des variables de décision qui minimisent la fonction objectif. Le nombre de combinaisons possibles entre ces valeurs rend la résolution du problème assez complexe, d'où l'utilisation de métaheuristiques comme les algorithmes génétiques (AG). Ces algorithmes évolutifs permettent d'obtenir une solution approchée d'un problème d'optimisation dans un temps raisonnable.

Le reste de cet article est organisé comme suit : La section 2 présente un état de l'art sur le problème traité et l'utilisation des algorithmes génétiques pour le résoudre. Dans la section 3, le modèle mathématique est développé exprimant le coût total moyen par unité de temps. La section 4 est consacrée à l'approche proposée basée sur des algorithmes génétiques pour résoudre la problématique. Les résultats sont présentés et commentés afin de démontrer l'intérêt de l'utilisation de l'algorithme d'optimisation proposé. Enfin, la section 5 conclut cet article.

2 ETAT DE L'ART

La performance d'un système de production dépend de sa capacité à répondre simultanément à plusieurs exigences telles que la satisfaction de la demande, la qualité des produits, les faibles coûts d'exploitation, qui dépendent fortement des performances en atelier. Pour répondre à toutes ces exigences, de nombreuses recherches se sont concentrées sur la résolution et l'optimisation conjointe de la production, de la maintenance et du contrôle qualité dans les systèmes manufacturiers. Les fonctions de production et de maintenance sont directement liées. Lorsqu'une défaillance survient au hasard, elle réduit la productivité du système et peut rendre le plan de production actuel obsolète. La révision du plan de production en cas d'urgence peut entraîner des coûts excessifs. Pour réduire la probabilité de panne et les pertes de production correspondantes, des stratégies de maintenance préventive sont mises en œuvre. [Sortrakul et Cassidy, 2007], [Yulan et al., 2008], [Wang, 2013] et [Liu et al., 2015].

[Dellagi et al., 2017] ont considéré un système de fabrication sujet à des pannes aléatoires et soumis à une politique de MP périodique avec une réparation minimale en cas de panne. Ils ont développé un modèle analytique pour déterminer un plan de production et de maintenance intégré quasi optimal en tenant compte de l'influence des fluctuations de la cadence de production. [Bahria et al., 2018a] ont développé une stratégie intégrée de maintenance et de contrôle qualité pour les systèmes de fabrication qui minimise le coût total moyen par unité de temps. Ces auteurs ont utilisé des cartes de contrôle pour le contrôle de la qualité. Ce travail a été étendu par l'optimisation de la taille du stock tampon en intégrant la fonction production [Bahria et al., 2018b].

[Salmasnia et al., 2017] ont intégré les cartes de contrôle et la maintenance en considérant un échantillonnage non uniforme. Une action de MP est effectuée à la fin du cycle de production lorsque le processus est sous contrôle. Lorsqu'une cause attribuable se produit (signal d'avertissement ou un autre type d'alarme), une action de maintenance corrective (MC) est exécutée. L'objectif est de déterminer les paramètres optimaux de la carte de contrôle. [Nourelfath et al., 2016] ont abordé un problème de dimensionnement des lots de multi-produits pour intégrer conjointement la production, la planification des actions de MP et la qualité. Des actions de MP imparfaites sont entreprises pendant les périodes d'inspection. Cependant, lorsque le système commence à produire des unités non

conformes ou à la fin du cycle de production, une action de MC est exécutée. L'objectif principal de ce travail est de déterminer le plan de production optimal et la stratégie de maintenance préventive (MP).

[Lesage et Dehombreux, 2015] ont proposé un modèle de simulation permettant d'optimiser conjointement les stratégies de maintenance et de contrôle qualité. Le modèle développé par ces auteurs considère un système de production continue qui réalise des articles de bonne ou de mauvaise qualité et soumis à des actions de maintenance préventive ou corrective basées sur les résultats du contrôle de la qualité.

Les métaheuristiques, comme les algorithmes génétiques, ont également été utilisés pour résoudre des problèmes d'optimisation des systèmes manufacturiers. En effet, les algorithmes génétiques sont des algorithmes d'optimisation stochastique basés sur les mécanismes de sélection naturelle et de génétique. La simplicité de leurs mécanismes, la facilité de leur mise en œuvre et leur efficacité même pour des problèmes complexes ont conduit à un nombre croissant de travaux de recherche dans le domaine de la gestion de la production.

[Ye et Ma, 2015] ont optimisé conjointement la planification de la production et de la maintenance préventive dans un atelier job-shop flexible. Ce modèle d'optimisation conjointe multi-objectifs prend en compte deux facteurs, qui sont le temps d'exécution maximum et les coûts de maintenance par unité de temps. Le modèle est résolu à l'aide d'un algorithme génétique.

[Abraham et al., 2015] ont proposé une approche basée sur des algorithmes génétiques pour un problème intégré de production et de distribution, dans un système multi-produits. Le but du problème est de maximiser le profit. Le coût total comprend le coût de production, le coût de détention des stocks, les frais de transport, les frais de réapprovisionnement et le coût de rupture de stock.

[Ettaye et al., 2018] ont intégré des stratégies de maintenance et de production dans un système de fabrication peu fiable. Une action de MP est effectuée à chaque cycle de maintenance et une réparation minimale en cas de panne imprévue. Les auteurs ont utilisé un algorithme génétique afin d'obtenir les valeurs optimales de la quantité à produire et du temps de mise en production ainsi que le stock et le niveau de rupture à la demande, afin de minimiser le coût total.

Plusieurs travaux de la littérature ont utilisé les algorithmes génétiques pour l'optimisation et le contrôle des systèmes manufacturiers. [Chou et al., 2006], [Chen et al., 2007], [Tornig et al., 2009], [Niaki et al., 2010], [Franco et al., 2012], [Lee et al., 2012], [Franco et al. 2014]. La mise en œuvre de cette méthode métaheuristique a montré une meilleure résolution des problèmes traités, ce qui encourage les chercheurs à adopter les algorithmes génétiques. En effet, ils ont la capacité d'éviter d'être piégés dans l'optimum local, ils utilisent des règles de sélection probabilistes et non déterministes et ils travaillent sur le chromosome, qui est une version codée des paramètres des solutions potentielles.

En conséquence, nous sommes intéressés à développer une approche basée sur des algorithmes génétiques qui minimise le coût total moyen par unité de temps pour un système de fabrication intégrant les coûts de maintenance et de qualité. Cette approche a permis de réduire les coûts opérationnels par rapport à l'ensemble des travaux antérieurs.

3 STRATEGIE ET MODELE MATHEMATIQUE

3.1 Stratégie d'optimisation intégrée de la maintenance et du contrôle qualité

Nous considérons un système de fabrication composé d'une seule machine et produisant des articles par lots de taille Q . L'état de la machine peut être évalué en mesurant une caractéristique clé de qualité x du produit fabriqué. Au cours de la production, à chaque unité de temps Is , un échantillon de taille n est prélevé et vérifié pour suivre l'évolution de l'indicateur x et surveiller ainsi l'état de dégradation du processus. Les résultats d'échantillonnage sont consignés dans une carte de contrôle 'x-bar' (Figure 1) avec m échantillons pour chaque lot de production.

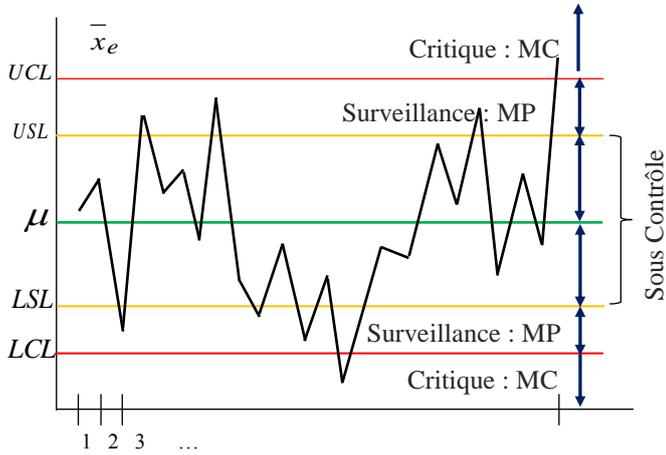


Figure 1. Carte de contrôle 'x-bar' et évolution de l'indicateur de qualité pendant un cycle de production

Avec :

$$LCL = \mu - k_c \times \frac{\sigma}{\sqrt{n}} : \text{Limite de contrôle inférieure} \quad (1)$$

$$UCL = \mu + k_c \times \frac{\sigma}{\sqrt{n}} : \text{Limite de contrôle supérieure} \quad (2)$$

$$LSL = \mu - k_p \times \frac{\sigma}{\sqrt{n}} : \text{Limite de surveillance inférieure} \quad (3)$$

$$USL = \mu + k_p \times \frac{\sigma}{\sqrt{n}} : \text{Limite de surveillance supérieure} \quad (4)$$

La moyenne μ et l'écart type σ du processus sont connus ou bien estimés. La valeur moyenne \bar{x}_e des mesures de l'indicateur x est inscrite sur la carte de contrôle pour chaque échantillon e . Elle s'exprime par :

$$\bar{x}_e = \frac{1}{n} \sum_{i=1}^n x_i \quad (5)$$

k_p et k_c sont respectivement les coefficients de limites de surveillance et du contrôle de la carte de contrôle 'x-bar'.

Un premier scénario est envisagé, si la moyenne des mesures de l'échantillon de l'indicateur de qualité se situe entre les limites de surveillance et de contrôle, l'unité de production est considérée alors à l'état de surveillance. Dans ce cas, la production est interrompue et une action de MP de durée moyenne μ_p est maintenue.

Un deuxième scénario est pris en compte, si la moyenne de l'échantillon des mesures de l'indicateur qualité dépasse les limites de contrôle, l'unité de production est considérée alors dans un état critique. Dans ce cas, la production est

interrompue et une action de MC de durée moyenne μ_c doit être effectuée.

Si la moyenne des mesures de l'échantillon de l'indicateur de qualité se situe entre les limites de surveillance, le processus est considéré alors sous contrôle. Dans ce cas, la production continue normalement. Ceci, constitue le troisième scénario.

Pour notre étude, nous considérons les hypothèses suivantes :

H1 : le système produit un seul type d'article;

H2 : L'indicateur de qualité contrôlé est distribué conformément à la loi Normale;

H3 : Les articles non conformes ne sont pas réinsérés dans le processus de production;

H4 : La non-conformité des produits est due à la dégradation de la machine;

H5 : Les différents coûts unitaires sont connus et constants;

H6 : Les ressources nécessaires pour effectuer les opérations de maintenance sont toujours disponibles;

H7 : Les durées des actions de maintenance préventive et corrective sont considérées plus courtes que le temps de production d'un lot;

H8 : Les opérations de maintenance sont considérées comme parfaites et l'équipement est remis à l'état neuf.

H9 : L'unité de production ne tombe jamais en panne avant la prochaine opération de maintenance.

3.2 Modélisation Mathématique

Notre problème est caractérisé par les paramètres suivants:

Q : Taille du lot de production;

n : taille de l'échantillon;

m : nombre d'échantillons contrôlés lors de la production d'un lot;

PCD : durée moyenne du cycle de production;

δ_p : Amplitude moyenne de la dérive de l'indicateur de qualité si les limites de surveillance sont dépassées;

δ_c : Amplitude moyenne de la dérive de l'indicateur de qualité si les limites de contrôle sont dépassées;

Is : intervalle d'échantillonnage;

$F(\cdot)$: Fonction de distribution de la loi Normale centrée réduite;

C_i : coût de l'inspection d'un produit;

C_r : coût d'un produit défectueux;

C_{MP} : coût d'une action MP;

C_{MC} : coût d'une action de MC;

C_F : Coût d'une fausse alerte.

La durée moyenne du cycle de production PCD s'exprime comme suit :

$$PCD = m \times \left\{ \begin{aligned} & \left((1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n})) \times \mu_p \right) \\ & + \left((1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n})) \times \mu_c + Is \right) \end{aligned} \right\} \quad (6)$$

Le coût total moyen de maintenance pendant un cycle de production est égal à la somme des coûts des actions de maintenance préventive et corrective. En considérant n_p et n_c sont respectivement le nombre d'actions de MP et MC effectuées pendant la production d'un lot. Ce coût s'exprime comme suit:

$$C_M = C_{MP} \times n_p + C_{MC} \times n_c \quad (7)$$

Ce qui donne :

$$\Gamma M = \left\{ \begin{aligned} & C_{MP} \times \left((1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n})) \right) \\ & + C_{MC} \times \left((1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n})) \right) \end{aligned} \right\} \times m \quad (8)$$

Le coût total moyen de la qualité est égal à la somme du coût des fausses alarmes, du coût de l'échantillonnage et du coût du rejet d'items non conformes pendant la période de production d'un lot. Il est exprimé comme le montre l'équation 9.

$$\Gamma Q = \left\{ \begin{array}{l} 2 \times C_F \times \left\{ \sum_{i=0}^m \left\{ \begin{array}{l} i \times \left(\begin{array}{l} F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) \\ + F(k_c - \delta_c \times \sqrt{n}) - F(-k_c - \delta_c \times \sqrt{n}) - 1 \end{array} \right) \\ \times \left(\begin{array}{l} 2 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \\ - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \end{array} \right) \end{array} \right\} \right\} \\ \times F(-k_p) + C_i \times n \times m \\ + C_r \times \left\{ \begin{array}{l} \left(1 - F\left(\frac{k_p}{\sqrt{n}} - \delta_p\right) - F\left(-\frac{k_p}{\sqrt{n}} - \delta_p\right) \right) \\ \times \left(1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \right) \\ + \left(1 - F\left(\frac{k_c}{\sqrt{n}} - \delta_c\right) - F\left(-\frac{k_c}{\sqrt{n}} - \delta_c\right) \right) \\ \times \left(1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \right) \\ + \left(1 - F(k_p) - F(-k_p) \right) \\ \times \left(\begin{array}{l} F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) \\ + F(k_c - \delta_c \times \sqrt{n}) - F(-k_c - \delta_c \times \sqrt{n}) - 1 \end{array} \right) \end{array} \right\} \times Q \end{array} \right. \quad (9)$$

En combinant les équations (8) et (9) et en divisant par la durée moyenne d'un cycle de production donnée par l'équation (6), nous obtenons le coût total moyen par unité de temps, comme le montre l'équation suivante :

$$\Gamma_{tot} = \left\{ \begin{array}{l} \left(\frac{1}{m} \times \left\{ \begin{array}{l} \left(1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \right) \times \mu_p \\ + \left(1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \right) \times \mu_c + Is \end{array} \right\} \right) \\ \left(\begin{array}{l} C_{MP} \times \left(1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \right) \\ + C_{MC} \times \left(1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \right) \end{array} \right) \times m \\ + 2 \times C_F \times \left\{ \sum_{i=0}^m \left\{ \begin{array}{l} i \times \left(\begin{array}{l} F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) \\ + F(k_c - \delta_c \times \sqrt{n}) - F(-k_c - \delta_c \times \sqrt{n}) - 1 \end{array} \right) \\ \times \left(\begin{array}{l} 2 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \\ - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \end{array} \right) \end{array} \right\} \right\} \\ \times F(-k_p) + C_i \times n \times m \\ \times \left\{ \begin{array}{l} \left(1 - F\left(\frac{k_p}{\sqrt{n}} - \delta_p\right) - F\left(-\frac{k_p}{\sqrt{n}} - \delta_p\right) \right) \\ \times \left(1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \right) \\ + \left(1 - F\left(\frac{k_c}{\sqrt{n}} - \delta_c\right) - F\left(-\frac{k_c}{\sqrt{n}} - \delta_c\right) \right) \\ \times \left(1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \right) \\ + \left(1 - F(k_p) - F(-k_p) \right) \\ \times \left(\begin{array}{l} F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) \\ + F(k_c - \delta_c \times \sqrt{n}) - F(-k_c - \delta_c \times \sqrt{n}) - 1 \end{array} \right) \end{array} \right\} \times Q \end{array} \right. \quad (10)$$

Nous considérons ici trois probabilités $p1$, $p2$ et $p3$, qui représentent respectivement les trois scénarios possibles définis dans la section 3.1.

La probabilité que le système se trouve dans l'état de surveillance (scénario 1 - action MP) est donnée par l'équation 11.

$$p_1 = 1 - F(k_p - \delta_p \times \sqrt{n}) + F(-k_p - \delta_p \times \sqrt{n}) \quad (11)$$

La probabilité que le système se trouve dans un état critique (scénario 2 - action MC) est donnée par :

$$p_2 = 1 - F(k_c - \delta_c \times \sqrt{n}) + F(-k_c - \delta_c \times \sqrt{n}) \quad (12)$$

La probabilité que le système soit à l'état contrôlé (scénario 3 - aucune intervention de maintenance) est donnée par :

$$p_3 = \left\{ \begin{array}{l} F(k_p - \delta_p \times \sqrt{n}) - F(-k_p - \delta_p \times \sqrt{n}) + F(k_c - \delta_c \times \sqrt{n}) \\ - F(-k_c - \delta_c \times \sqrt{n}) - 1 \end{array} \right\} \quad (13)$$

Les phases de développement des différents coûts sont détaillées dans nos travaux [Bahria et al., 2018a]. Ainsi, le coût total est exprimé en fonction des cinq variables de décision : n , m , Is , k_p et k_c .

Ces variables de décision varient selon des intervalles définis. La recherche de leurs valeurs optimales, pour minimiser le coût total, rend notre problème assez complexe. Nous proposons donc d'utiliser une méthode métaheuristique pour sa résolution.

La section suivante présente notre approche basée sur les algorithmes génétiques pour calculer les valeurs optimales des cinq variables de décision.

4 APPROCHE PROPOSEE BASEE SUR LES ALGORITHMES GENETIQUES

4.1 Codage du chromosome

Notre objectif est de déterminer les valeurs des cinq variables de décision : kp , kc , Is , n et m qui minimisent la fonction objectif. En effet, le codage représente les solutions sous forme de chromosome. La figure 2 montre le codage utilisé pour notre solution.

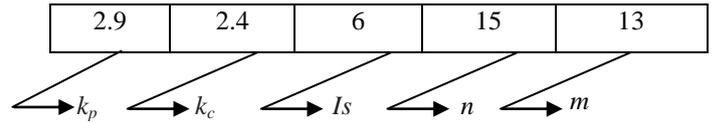


Figure 2. Codage du chromosome

4.2 Opération de croisement

Suite à la génération aléatoire de la population initiale, on procède à la phase de croisement, qui assure la recombinaison des gènes parentaux pour former de nouveaux descendants. Pour cela, on utilise le croisement à un point, on considère p la position de croisement et on traite cette phase comme le montre la figure 3.

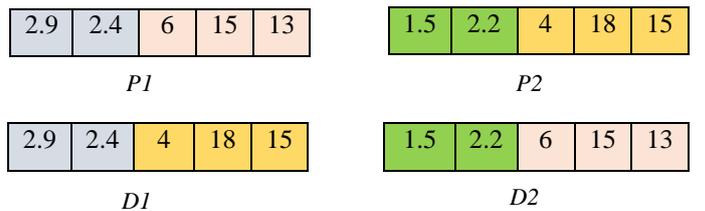


Figure 3. Opération de croisement

4.3 Opération de mutation

La mutation consiste à choisir deux allèles, au hasard, dans un chromosome, et à échanger leurs valeurs respectives. Toutefois, pour notre cas chaque allèle désigne une variable spécifique, nous ne pouvons donc pas échanger deux allèles, chacun représentant une variable différente. Nous procédons alors comme le montre la figure 4, pour l'opération de mutation.

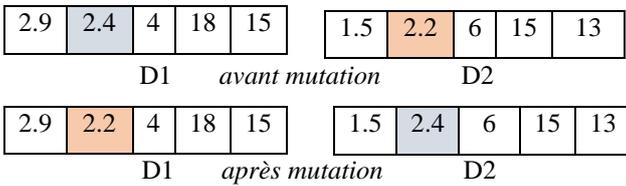


Figure 4. Opération de mutation

4.4 Procédure de l'approche proposée pour la minimisation du coût total

La génération de la population initiale permet de produire une population d'individus qui servira de base aux générations futures. Le choix de la population initiale est important car il peut accélérer la convergence vers l'optimum global.

A partir de cette population initiale de taille N , une nouvelle population est créée suite aux applications des opérations de croisement et de mutation. Nous conservons une copie de la population initiale, car elle peut contenir de bonnes solutions. Cette phase est illustrée par la figure 5.

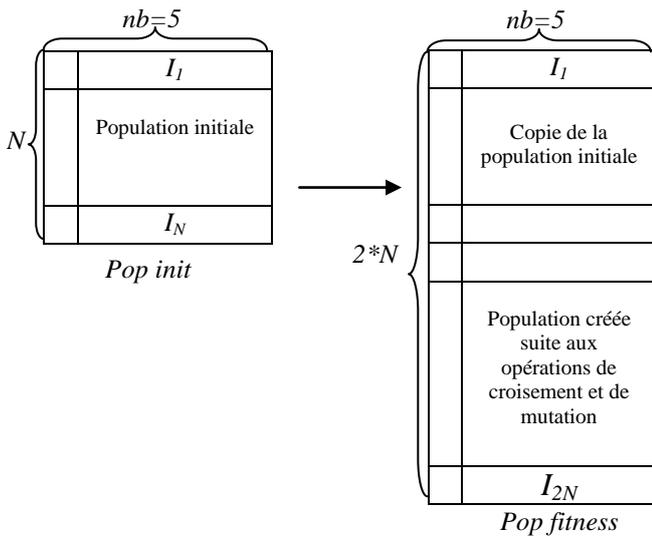


Figure 5. Création des populations

Les différentes étapes de la procédure d'optimisation du coût total sont détaillées dans l'algorithme suivant (Figure 6).

nb : nombre des variables de décision; N : taille de la population; μ_p : durée moyenne de MP ; μ_c : durée moyenne de MC ; C_{PM} : coût de MP ; C_{CM} : coût de MC ; C_i : coût d'inspection; C_r : coût de rejet; C_f : coût de fausse alarme; δ_p : dépassement des limites de surveillance; δ_c : dépassement des limites de contrôle; Q : taille du lot de production;

Etape 1: Création de la population initiale Pop_{init} de taille N
Tant que le nombre de générations n'est pas atteint **faire**

Début

Etape 2 : Création de la population intermédiaire $Pop_{fitness}$ de taille $2N$ avec les individus de copie, de croisement et de mutation ;

Etape 3: Déterminer la valeur du fitness pour chaque individu de la $Pop_{fitness}$ (coût total) Equation 10;

Copier la meilleure solution ;

Si la solution courante est meilleure **alors**

La meilleure solution = la solution courante ;

Fin

Incrémenter le nombre de générations ;

Fin

Fin

Figure 6. Algorithme génétique proposé

Pour chaque individu de la deuxième population de taille $2N$, nous calculons la valeur du fitness pour déterminer celui qui minimise notre fonction objectif.

5 RESULTATS ET SIMULATION

Pour l'exemple de simulation, les valeurs des variables d'entrée sont les suivantes :

- Le nombre de variables de décision nb est 5 ;
- La durée moyenne de maintenance préventive μ_p est de 0,5 heure ;
- La durée moyenne de maintenance corrective μ_c est de 2 heures ;
- Le coût de la maintenance préventive au C_{MP} est de 200 \$;
- Le coût de la maintenance corrective C_{MC} est de 1000 \$;
- Le coût de l'inspection du produit C_i est de 15 \$ / unité inspectée ;
- Le coût du produit rejeté C_r est de 70 \$ / unité rejetée ;
- Le coût en fausse alarme C_f est de 120 \$ / fausse alarme ;
- Le dépassement des limites de surveillance δ_p est de 0,4 ;
- Le dépassement des limites de contrôle δ_c est égal à 1 ;
- La taille du lot de production Q est de 1000 unités.

Notre approche a été implémentée en utilisant DevC ++. Le tableau 1 présente les meilleurs résultats obtenus avec notre algorithme génétique. Le tableau 2 illustre le meilleur résultat donné par une autre approche basée sur une procédure d'optimisation numérique itérative qui dépend des incréments, développée par Bahria et al (2018a).

Tableau 1 : Meilleure solution avec les AGs

n	m	$Is(h)$	k_p	k_c
36	45	16.25	2.134	2.989
IQ (\$)	IM (\$)	PCD (h)	Γ_{tot} (\$/h)	
80394.4	50386.5	834.72	156.67	

Tableau 2 : Meilleure solution avec la procédure itérative

n	m	$Is(h)$	k_p	k_c
18	15	4	1.5	2.2
IQ (\$)	IM (\$)	PCD (h)	Γ_{tot} (\$/h)	
12258	1060.4	62.33	213.67	

En comparant les résultats obtenus par les deux méthodes, nous constatons que notre approche a réduit le coût total de 26,68%. La réduction du coût total moyen par unité de temps par rapport à l'approche itérative est justifiée par le balayage d'un espace de solutions réalisables plus étendu, en comparaison avec celui fixé dans le travail de Bahria et al (2018a).

Notons que notre approche basée sur les algorithmes génétiques a été comparée à la procédure itérative, en prenant en compte les mêmes valeurs des données d'entrée.

En effet, la taille de l'échantillon et le nombre d'échantillons ont augmenté. Cependant, la fréquence d'échantillonnage a diminué. Cela a augmenté le coût moyen d'échantillonnage : $C_{ech} = C_i * n * m$. Ce coût était de 4050 \$ et a été augmenté à 24300 \$.

Nous constatons que le coût de maintenance a augmenté, nous effectuons donc plus d'actions de maintenance afin de mieux maîtriser le processus et réduire le nombre d'unités non conformes. Cela a réduit le coût total moyen par unité de temps des produits non conformes à 67.2 \$/h contre 131.68 \$/h. Par conséquent, si nous effectuons plus d'actions de maintenance, la durée moyenne d'un cycle de production augmente également.

L'avantage des algorithmes génétiques se révèle principalement dans la variété et le nombre de combinaisons possibles des paramètres à traiter. Le tableau 3 montre d'autres résultats obtenus par notre approche, réduisant ainsi le coût total.

Tableau 3 : Simulations avec les AGs

n	m	$Is(h)$	k_p	k_c
11	19	5.47	2.8	3.6
$\Gamma Q (\$)$	$\Gamma M (\$)$	$PCD (h)$	$\Gamma_{tot} (\$/h)$	
15925.5	7433.6	118.9	196.4	

n	m	$Is(h)$	k_p	k_c
11	43	3.26	2.3	3.9
$\Gamma Q (\$)$	$\Gamma M (\$)$	$PCD (h)$	$\Gamma_{tot} (\$/h)$	
17166.1	12445.8	165.8	178.5	

n	m	$Is(h)$	k_p	k_c
13	15	18.02	2.05	2.43
$\Gamma Q (\$)$	$\Gamma M (\$)$	$PCD (h)$	$\Gamma_{tot} (\$/h)$	
43647.4	13987.5	298.7	192.9	

Pour la première partie du tableau 3, moins d'inspections ont été notées, réduisant le coût d'inspection qui est égal à 3135 \$.

Les coefficients des limites de la carte ont augmenté et les limites de surveillance et de contrôle s'éloignent pour effectuer moins fréquemment des actions de maintenance préventive et corrective. Par conséquent, le coût de maintenance moyen diminue.

La gravité de la carte de contrôle étant plus faible, la proportion d'articles non conformes augmente, ce qui entraîne une augmentation des coûts de rejet à 107.6 \$/h. La durée moyenne d'un cycle de production est réduite suite à la diminution du nombre d'actions de maintenance préventive et corrective.

Pour la deuxième partie du tableau 3, le nombre d'articles inspectés au cours du cycle de production passe à 473 ($n \times m$) avec une fréquence d'échantillonnage plus élevée. Ainsi, le coût d'échantillonnage augmente à 10067.8 \$.

Le coefficient des limites de surveillance a très légèrement augmenté et on constate une variation significative du

coefficient des limites de contrôle. Par la suite, la carte de contrôle est moins serrée et moins d'actions de maintenance sont entreprises, ce qui explique la diminution du coût moyen de maintenance.

Pour la troisième partie du tableau 3, le nombre d'échantillons contrôlés lors de la production d'un lot est passé à 195. Cela se traduit par une diminution du coût d'inspection et par la suite du coût de qualité. L'équation 6 exprime la durée moyenne du cycle de production. Ce coût a diminué suite à la réduction du nombre d'échantillons prélevés.

6 CONCLUSIONS

Cet article présente une approche basée sur les algorithmes génétiques pour la résolution et l'optimisation conjointe de la maintenance et du contrôle de la qualité dans les systèmes manufacturiers.

Un état de l'art sur les différentes stratégies de résolution des problèmes d'optimisation des systèmes manufacturiers a été proposé. De nombreux chercheurs ont utilisé les algorithmes génétiques donnant de bons résultats, d'où notre intérêt pour cette méthode métaheuristique. Par la suite, nous avons présenté notre stratégie basée sur les cartes de contrôle, ainsi que le modèle mathématique développé.

L'utilisation des algorithmes génétiques ainsi que les différentes phases de l'approche proposée ont ensuite été détaillées, afin de déterminer les valeurs optimales des variables de décision minimisant le coût total moyen par unité de temps. Notons ici que nous avons développé une opération de mutation heuristique appropriée à notre problème, puisque chaque gène, au sein du chromosome, représente un paramètre bien défini.

Dans la dernière partie, nous avons présenté les résultats de simulation en utilisant les mêmes valeurs de paramètres d'entrée que celles de la littérature pour pouvoir comparer. Les résultats obtenus indiquent que l'approche que nous proposons a produit de meilleurs résultats. Cela confirme l'efficacité des différentes phases développées pour notre algorithme génétique.

Pour nos travaux futurs, nous proposons d'étudier d'autres métaheuristicques, comme la recherche taboue ou le recuit simulé et de comparer les résultats des différentes méthodes évolutives. Une autre direction importante pour nos recherches futures est d'étudier les systèmes de fabrication multi-produits ou les lignes de production composées de plus d'une machine.

7 REFERENCES

- Sortrakul, N., & Cassady, C. R. (2007). Genetic algorithms for total weighted expected tardiness integrated preventive maintenance planning and production scheduling for a single machine. *Journal of Quality in Maintenance Engineering*, vol. 13(1), 49-61.
- Yulan, J., Zuhua, J., & Wenrui, H. (2008). Multi-objective integrated optimization research on preventive maintenance planning and production scheduling for a single machine. *International Journal of Advanced Manufacturing Technology*, vol. 39(9-10), pp. 954-964.
- Wang, S. (2013). Integrated model of production planning and imperfect preventive maintenance policy for single machine system. *International Journal of Operational Research*, vol. 18(2), pp. 140-156.
- Liu, X., Wang, W., & Peng, R. (2015). An integrated production, inventory and preventive maintenance model for a multi-product production system. *Reliability*

- Engineering & System Safety*, vol. 137, pp. 76-86.
- Dellagi, S., Chelbi, A., & Trabelsi, W. (2017). Joint integrated production-maintenance policy with production plan smoothing through production rate control. *Journal of Manufacturing Systems*, vol. 42, pp. 262-270.
- Bahria, N., Chelbi, A., Harbaoui Dridi, I. and Bouchriha, H. (2018a). Maintenance and quality control integrated strategy for manufacturing systems, *European Journal of Industrial Engineering*, vol. 12(3), pp. 307-331.
- Bahria, N., Chelbi, A., Bouchriha, H., & Dridi, I. H. (2018b). Integrated production, statistical process control, and maintenance policy for unreliable manufacturing systems. *International Journal of Production Research*, 1-23.
- Salmasnia, A., Abdzadeh, B., Namdar, M. (2017), A joint design of production run length, maintenance policy and control chart with multiple assignable causes, *Journal of Manufacturing Systems*. vol. 42, pp. 44–56.
- Nourelfath, M., Nahas, N., & Ben-Daya, M. (2016), Integrated preventive maintenance and production decisions for imperfect processes. *Reliability Engineering and System Safety*, vol. 148, pp. 21–31.
- Lesage, A., & Dehombreux, P. (2015). Optimisation conjointe des stratégies de maintenance et de contrôle qualité : modèle intégrant la maintenance imparfaite. *11ème congrès international de génie industriel–GIGI 2015*, 26-28.
- Ye, J., & Ma, H. (2015). Multiobjective joint optimization of production scheduling and maintenance planning in the flexible job-shop problem. *Mathematical Problems in Engineering*, <http://dx.doi.org/10.1155/2015/725460>.
- Abraham, A. J., Kumar, K. R., Sridharan, R., & Singh, D. (2015). A genetic algorithm approach for integrated production and distribution problem. *Procedia-Social and Behavioral Sciences*, 189, 184-192.
- Ettaye, G., El Barkany, A., Jabri, A., & El Khalfi, A. (2018). Optimizing the integrated production and maintenance planning using genetic algorithm. *International Journal of Engineering Business Management*, vol 10, p 1-9.
- Chou, C. Y., Chen, C. H., & Chen, C. H. (2006). Economic design of variable sampling intervals T2 control charts using genetic algorithms. *Expert Systems with Applications*, 30(2), 233-242.
- Chen, Y. K., Hsieh, K. L., & Chang, C. C. (2007). Economic design of the VSSI X control charts for correlated data. *International journal of production economics*, 107(2), 528-539.
- Torng, C. C., Lee, P. H., Liao, H. S., & Liao, N. Y. (2009). An economic design of double sampling X charts for correlated data using genetic algorithms. *Expert Systems with Applications*, 36(10), 12621-12626.
- Niaki, S. T. A., Ershadi, M. J., & Malaki, M. (2010). Economic and economic-statistical designs of MEWMA control charts—a hybrid Taguchi loss, Markov chain, and genetic algorithm approach. *The International Journal of Advanced Manufacturing Technology*, 48(1-4), 283-296.
- Franco, B. C., Costa, A. F. B., & Machado, M. A. G. (2012). Economic-statistical design of the X chart used to control a wandering process mean using genetic algorithm. *Expert Systems with Applications*, 39(17), 12961-12967.
- Lee, P. H., Torng, C. C., & Liao, L. F. (2012). An economic design of combined double sampling and variable sampling interval X control chart. *International Journal of Production Economics*, 138(1), 102-106.
- Franco, B. C., Celano, G., Castagliola, P., & Costa, A. F. B. (2014). Economic design of Shewhart control charts for monitoring autocorrelated data with skip sampling strategies. *International Journal of Production Economics*, 151, 121-130.

CIGI 2021

Maturité du système physique de production : modèle de diagnostic 4.0

DAVID DAMAND¹, SAMIA GAMOURA², MARC BARTH³

¹ EM Strasbourg Business School, Université de Strasbourg, Humanis (EA 7308)
61 avenue de la Forêt Noire, Strasbourg, France
damand@unistra.fr

² EM Strasbourg Business School, Université de Strasbourg, Humanis (EA 7308)
61 avenue de la Forêt Noire, Strasbourg, France
samia.gamoura@em-strasbourg.eu

³ INSA Strasbourg, Humanis (EA 7308)
61 avenue de la Forêt Noire, Strasbourg, France
marc.barth@insa-strasbourg.fr

Résumé – Le concept d'industrie 4.0 ou industrie du futur est associé à une évolution de l'organisation des ressources de production. Cette (r)évolution se caractérise par une interconnexion des moyens de production et des systèmes à la fois au sein des entreprises et vers l'extérieur (fournisseur, client, filiales, etc.). Compte tenu de l'impact sur les enjeux financiers, une question fondamentale concernant la qualité des conditions initiales à une transformation 4.0. Dans ce papier, la question de la qualité du système physique de production est discutée. La littérature propose un nombre significatif de modèles d'analyse du système physique de production. Cependant, la littérature ne propose pas une méthode intégrée et systématique basée sur ces modèles d'analyse. L'objectif de ce papier est de discuter les potentialités d'une telle méthode. Les premiers résultats sont présentés.

Abstract – The concept of Industry 4.0 or industry of the future is associated with an evolution in the organization of production resources. This (r) evolution is characterized by an interconnection of means of production and systems both within companies and to the outside (supplier, customer, subsidiaries, etc.). Given the impact on financial issues, a fundamental question concerns the quality of the initial conditions for a 4.0 transformation. In this paper, the issue of the quality of the physical production system is discussed. The literature offers a significant number of models for analyzing the physical production system. However, the literature does not offer an integrated and systematic method based on these analysis models. The objective of this paper is to discuss the potentials of such a method. The first results are presented.

Mots clés – système physique de production, conception, diagnostic, propriétés d'analyse.

Keywords – production system, conceptual design, diagnostic, analysis properties.

1 INTRODUCTION

Un des défis de la 4^{ème} révolution industrielle sous de le nom d'industrie 4.0 est de répondre à cette exigence de produits uniques et personnalisés tout en conservant des coûts équivalents à une production de masse et cela, malgré les faibles volumes de production engendrés. La réponse 4.0 se traduit par la connexion des besoins du client au système physique de production via l'évolution des systèmes de décisions et d'information (vers des systèmes cyber-physiques). Il est généralement admis que la performance de ces systèmes évolués de décisions et d'information est fortement dépendante des conditions initiales associées à la qualité de conception du système physique de production (SPP) (Koren et al., 1998 ; Rouse et al., 2003). Le cas échéant, et compte tenu des enjeux financier (Kaspar et Schneider 2015), il faut pouvoir (re) évaluer et/ou () concevoir le SPP.

Un SPP efficient, dit « Lean », permettra de concevoir un système de décision et d'information efficient (Singh et al., 2006 ; Amoako-Gyampah et al, 2008). A contrario, la non-performance d'un SPP, dit « non-Lean » est difficilement compensable par la performance du système de décision et d'information (Olhager, 2013, Nicoletti, B. 2013).

La littérature scientifique dénombre un nombre important de travaux sur la conception de système physique de production (SPP). Ces travaux proposent une décomposition séquentielle systématique du processus de conception du SPP en trois phases (Raoot et al.,1993 ; Phillips,1997 ; Irani et al., 2000 ; Tompkins et al 2003 ; Singh et al., 2006 ; Muther, 1973) : Elaboration des agencements conceptuels du SPP, Elaboration des agencements physiques du SPP, Implémentation du SPP.

Les travaux de ce papier se positionnent au niveau de la phase d'élaboration des agencements conceptuels du SPP, appelée aussi phase de conception préliminaire du SPP. Cette phase présente des enjeux stratégiques et économiques majeurs dans le cycle de vie du SPP (Phillips, 1997 ; Alfness et al., 2006). Généralement différents principes de solutions établies selon différents points de vue (processus, surface disponible, coût, ...) sont évalués afin d'aider les concepteurs à sélectionner les solutions jugées les plus pertinentes, c'est à dire, qui correspondent aux besoins formulés par les concepteurs (Pahl, 1996).

La difficulté majeure est essentiellement liée au nombre important de points de vue à prendre en compte : législation, sécurité, humain, manutention, gestion, génie civil, maintenance, ..., et procédé c'est à dire le flux physique des produits. Le flux physique des produits représente l'élément fondateur du SPP et est étroitement lié à la performance du SPP (Tompkins et al., 2003), (Leone et al., 2002). Le point de vue considéré par la suite est le point de vue du flux physique des produits modélisé dans les gammes de production (GP), c'est à dire le flux des produits entre poste de charge (PDC). Le type de GP considéré dans ce papier est généralement de grande taille et très varié c'est à dire : multi-part, multi-niveaux, multi-opérations, multi-poste de charge et multi-processus de production.

Aussi, les relations entre PDC créées par le flux de produits sont potentiellement complexes et la solution d'implémentation du SPP n'est pas caractérisée de solution évidente. Notons qu'avant de prendre en charge une solution, le concepteur doit d'abord l'accepter. La confiance du concepteur dans une solution est souvent liée à la compréhension qu'il en a (Buchanan and Daellenbach, 1987). Plus le concepteur comprend et approuve une solution, plus il a confiance dans la solution. Cela renvoie à la question de l'intelligibilité d'une solution.

Pour comprendre les propriétés (relations spécifiques) des données de flux entre PDC et tendre vers des principes de solutions, le processus de compréhension est usuellement décomposé en 3 activités (Tableau 1).

L'activité 2 (Tableau 1) représente le cœur du processus de compréhension, et également le cadre d'analyse de ce papier. L'analyse concerne la qualité de la solution d'agencement conceptuelle et tente de répondre à des questions du type : pourquoi telle solution se révèle-t-elle pertinente ? Quelles propriétés caractérisent telle solution ? C'est à ce niveau que se fait le lien avec l'extraction progressive des structures visuelles (SV) contenu dans les Données de Flux (DF).

L'analyse repose essentiellement sur une problématique de description (Roy, 1985) de la relation 'Propriété d'Analyse (PA) – Structure Visuelle (SV)'. Aussi, le déroulement de cette problématique de description induit deux questions génériques de recherche (Q1, Q2) : Quel est l'ensemble des relations ' (1) PA – (1) SV' applicable aux DF (Q1) ? Quel est l'ensemble des relations 'nPA – nSV' potentiellement pertinente à transformer en VUES ?

Tableau 1. Méthode usuelle de compréhension des propriétés des données de flux

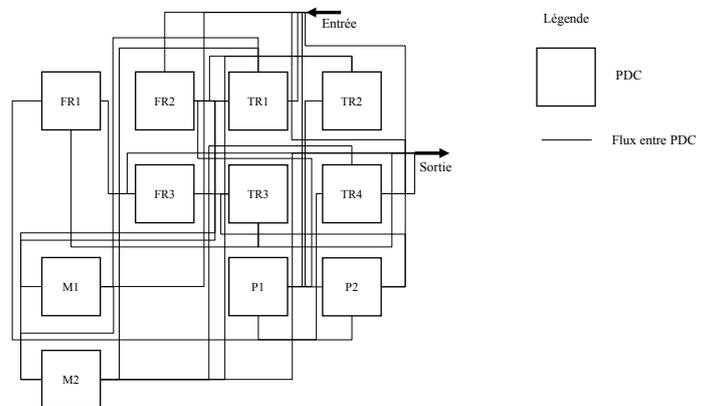
Données/ Activités ↓	Description
Données (0) : Gamme de Production	Les Gammes de Production représentent les données brutes.
Activité 1	Transformation des Gammes de Production en données de flux.
Données (1) : Données de Flux	Les Données de Flux résultant sont des tables de données : Matrice Produit/PDC, Matrice PDC/PDC.
Activité 2	Transformation des Données de Flux en structure visuelle.
Données (2) : Structures Visuelles	Les structures visuelles correspondent aux résultats d'application des propriétés d'analyse aux Données de Flux. Une Structure Visuelle (SV) est issue du résultat d'application d'une Propriété d'Analyse (PA) (relation 1PA-1SV) ou à la combinaison de plusieurs résultats d'analyse (relation nPA-nSV).
Activité 3	Transformation des Structures Visuelles en VUES. L'objectif est de rendre visibles les Structures Visuelles.
Données (3) : Vue	Une VUE représente un agencement conceptuel des PDC. Une VUE est un dessin de graphe orientés et pondérés (Les sommets et arcs du graphe représentent respectivement les postes de charges et le flux entre les postes de charges). Une VUE constitue l'interface permettant au concepteur d'observer et d'interpréter les Données de flux.

↓ Selon la taxonomie de Card (1999).

L'exemple ci-dessous illustre les données (0)(1)(2)(3) de la méthode usuelle décrite dans le Tableau 1. La pertinence de la propriété d'analyse d'ordre total est considérée.

Etat initial (Figure 1) : plan d'implantation des PDC et représentation des flux entre PDC.

Figure 1. Plan d'implantation



Données (0) : Gamme de Production

Les Gammes de Production concernent 10 produits et 11 PDC (Tableau 2).

Tableau 2. Gammes de Production

Produits	Gammes de Production (GP)
A01	P2 → FR1 → FR3 → TR3
A02	TR4 → P1 → TR2 → M2
A03	P1 → TR1 → M2 → TR4
A04	P2 → FR1 → TR3 → FR3
A05	M1 → FR2 → TR2 → M2
A06	P1 → TR4
A07	P1 → P2 → TR3
A08	TR1 → M1 → FR2 → M2
A09	P2 → FR1 → FR3
A10	FR2 → TR1 → M1 → M2

Données (1) : Donnée de Flux

Les données de Flux sont représentées par la matrice d'adjacence pondérée PDC/PDC (Tableau 3). Une cellule du tableau représente la quantité de flux (nombre d'ordres de fabrication).

Tableau 3. Matrice d'adjacence pondérée

PDC/ PDC	FR1	FR2	FR3	M1	M2	P1	P2	TR1	TR2	TR3	TR4
FR1	0		60							5	
FR2		0			40			60	60		
FR3			0							50	
M1		100		0	60						
M2					0						10
P1						0	10	10	10		40
P2	65						0			10	
TR1				100	10			0			
TR2					70				0		
TR3			5							0	
TR4						10					0

Données (2) : Structure Visuelle

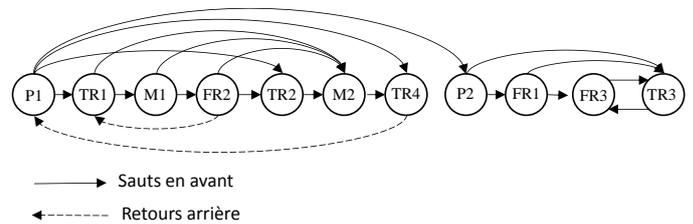
La Structure Visuelle est obtenue à la suite de l'application de la propriété d'ordre total (Tableau 4). L'algorithme appliquée minimise la quantité de flux sous la diagonale.

Tableau 4. Structure visuelle

Rang	1	2	3	4	5	6	7	8	9	10	11
PDC/ PDC	P1	TR1	M1	FR2	TR2	M2	TR4	P2	FR1	FR3	TR3
P1	0	10			10		40	10			
TR1		0	100			10					
M1			0	100		60					
FR2			60		0	60	40				
TR2					0	70					
M2						0	10				
TR4	10							0			
P2									0	65	10
FR1									0	60	5
FR3										0	50
TR3										5	0

Données (3) : la Figure 2 correspond à la visualisation de la structure visuelle sous forme de tracé de graphe.

Figure 2. VUE : tracé de graphe de flux de la relation d'ordre total entre les PDC



2 REVUE DE LA LITTÉRATURE

La littérature scientifique (Irani et al., 2000 ; Tompkins et al 2003) décrit une typologie de flux déclinée en 6 types de flux (Tableau 5) : Type 1. Cellule, Type 2. Ligne, Type 3. Niveau, Type 4. Proximité, Type 5. Zone de lancement, Type 6. Zone de fin de production. Chaque type de flux est caractérisé selon : le principe organisationnel du Lean Management, le principe technologique, le type de relation PA-SV recensées dans la littérature scientifique.

Les 10 types de relation PA-SV (10 PA, resp. 10 SV) recensées dans la littérature scientifique sont respectivement : classe, ordre total, ordre partiel, niveau antériorité, relation, Squelette, Source totale, Source partielle, Puits total, Puits partiel. Pour chaque type de PA, un nombre important d'algorithmes d'analyse sont proposés (Daita et al., 2000 ; Irani et al., 2000). Les type de PA les plus étudiés concernent principalement les types « classe et ordre ». La classe regroupe en famille les PDC ; ce qui permet lors de la phase d'élaboration des agencements physique du SPP de tirer profit de la similarité des séquences des opérations des différents produits pour diviser en unité de production plus petite (agencement physique : cellule) et ainsi simplifier la gestion (De Guio et al., 1999). L'ordre permet de déterminer un « sens » dominant de flux de produit entre les PDC ; ce qui favorise la hiérarchisation des opérations suivant les procédés (agencement physique en ligne), et simplifie ainsi les déplacements (Zhou et al., 1996).

Tableau 5. Dénombrement des relations PA-SV de la littérature scientifique

Typologie de flux	Principe organisationnel du Lean Management	Principe technologique	Relation PA-SV	Références
Type 1. Cellule	Simplifier la gestion	Diviser l'ensemble des PDC en unité de production plus petite	1. Classe	(Mahseh et al., 2002), (Yin et al., 2006), (Selim et al., 1998), (De Guio et al., 1999),
Type 2. Ligne	Simplifier le déplacement	Ordonner les postes suivant les procédés	2. Ordre total 3. Ordre partiel	(Sarker et al., 2000), (Huang, 2003) (Hollier, 1963) (Irani et al., 2000a) (Barth et al., 1998), (Zhou et al., 1996), (Zhou, 1998), (De Guio et al., 1996)
Type 3. Niveau		Hiérarchiser les postes suivant les procédés	4. Niveau	(Lee et al., 2001), (Lee et al., 2004)
Type 4. Proximité	Diminuer le déplacement	Rapprocher les postes	5. Relation	(Deba and Bhattacharyya, 2002)
Type 1, 2, 3, 4	Simplifier le déplacement	Standardiser les moyens de déplacement	6. Squelette	Zhou et al., 1996
Type 5. Zone de lancement	Simplifier la gestion des lancements de production	Regrouper les postes de début de GP	7. Source totale 8. Source partielle	(Zhou, 1998), (Lee et al., 2004), (Barth et al., 1989), (Zhou et al., 2003), (Lee et al., 2001), (De Guio et al 1996).
Type 7. Zone de fin de production	Simplifier la gestion des fins de production	Regrouper les postes fin de GP	9. Puits total 10. Puits partiel	

La littérature ne propose pas de papier de type « revue de la littérature » qui capitalisent l'ensemble des PA à appliquer. Les travaux de la littérature étudient principalement l'application d'une PA à la fois ; ils décrivent la SV résultante et l'indicateur de performance associé. Et, plus rarement, des travaux étudient l'application d'une combinaison de PA. Par exemple, Lee (2004) propose l'étude de l'application du PA de niveau combinée avec les PA de « puits et sources » ; ce qui permet respectivement de regrouper les PDC de début de GP (agencement physique : zone de lancement) et de fin de GP (agencement physique : zone de fin de production) et ainsi simplifier le lancement et la fin de production.

Les problèmes traités dans la littérature sont formalisés sous la forme de problèmes d'optimisation combinatoire. Pour chaque problème plusieurs méthodes d'optimisation sont proposées. Avec l'accroissement des vitesses de calcul durant ces 30 dernières années, ces méthodes ont trouvé leur pleine application. A titre d'illustration pour le PA classe, la littérature dénombre plusieurs types d'algorithmes comme : Array-based methods, clustering methods, mathematical programming based methods, Graph theoretical methods, Neural network based methods, Search methods, Genetic Algorithm optimisation method, incremental cell-formation method.

Les travaux de la littérature mettent plutôt l'accent sur l'analyse d'algorithmiques et leur performance que sur la synthèse. LEE et al. (2004) ont développés 6 méthodes de recherche heuristique basées sur 'simulated annealing' et 'tabu search'. Les algorithmes sont comparés par rapport au temps de calcul de recherche de solutions optimales. Yin et Yasuda (2006) développent une étude comparative des différents coefficients de similarité recensés dans les algorithmiques de construction de classe. Ils dénombrent 17 coefficients de similarité. Il propose une taxonomie pour mettre en correspondance les différents coefficients de similarité.

En réponses aux 2 questions de recherche Q1 et Q2 :

Q1. Le Tableau 5 construit dans ce papier à partir d'une revue de la littérature est une réponse à Q1. Le Tableau 5 recense respectivement 10 types de relations 1PA-1SV. Cependant il est probable, compte tenu de la diversité des procédés, que toutes les PA ne soit pas pertinente dans le contexte d'une étude de cas. En effet, en considérant par exemple la PA_{ordre_total} , la SV_{ordre_total} résultante peut être caractérisé par un nombre de retours arrière supérieur au nombre de saut en avant dans les DF. Dans cet exemple, la PA est jugée non pertinente. Il est donc probable que la pertinence d'une PA soit contingente. C'est ce que la première hypothèse de recherche tente de valider.

Hypothèse 1 : la pertinence d'une PA est contingente.

Q2. Le Tableau 5 ne répond que très partiellement à Q2. Les combinaisons entre SV sont très peu explicitées. Les travaux de la littérature ne mentionnent qu'occasionnellement l'étude de la combinaison de SV. Seul un papier (Lee, 2004) a été recensé. Les décisions de conception du SPP (Quel type de flux mettre en œuvre ?) sont étroitement couplées aux principes organisationnels et ne peuvent pas être analysées de manière isolée. La juxtaposition des VUES mono- SV ne suffit pas pour décrire des VUES multi- SV dû au fait de leur combinaison. A titre de comparaison, dans le domaine chimique, les propriétés de l'eau ne résultent pas des propriétés de l'hydrogène et de l'oxygène, éléments chimiques qui la composent.

Selon la thèse ci-dessus, pour décrire les propriétés contenues dans les GP, l'analyse mono- SV est nécessaire, mais ne se suffit pas à elle seule. Cependant, aucun papier ne s'inscrit dans la globalité qui s'est construite à partir de la prise en compte : ni de l'ensemble des principes organisationnels, ni de l'ensemble des PA, ni de l'ensemble des SV et combinaison de SV.

L'objectif est de faire émerger des caractéristiques nouvelles et non apparentes contenu dans les DF à partir de combinaisons multi- SV. La combinaison idéale est la combinaison qui tient compte de l'ensemble des principes organisationnels et de l'ensemble des relations PA-SV (Tableau 5).

S'inscrire dans la globalité des 10 relations PA-SV recensés revient à étudier un modèle générique de 1013 combinaisons de SV (modèle générique :

$$1013 = C_{10}^2 + C_{10}^3 + C_{10}^4 + C_{10}^5 + C_{10}^6 + C_{10}^7 + C_{10}^8 + C_{10}^9 + C_{10}^{10}.$$

Il est difficilement concevable (compte tenu du temps d'analyse et des limites cognitives) de mener 1013 analyses des DF à partir de 1013 combinaisons de SV.

Notons que la phase de conception préliminaire du SPP est contrainte par la règle de prise en compte de l'ensemble des postes de charges des GP (Burbidge, 1989 ; Zhou et al., 2003 ; Tompkins et al., 2003). Ceci est essentiel puisque la phase de conception détaillée (layout design) exige le positionnement de l'ensemble des postes de charge sur la surface géométrique dédié au SPP étudié. Compte tenu de cette règle d'instanciation du modèle générique, des combinaisons de SV sont non applicables. Par exemple, la combinaison (SV_{puits} \wedge SV_{source}).

Il est donc probable que le nombre de 1013 combinaisons puisse être réduit en un modèle de référence de taille raisonnable eu égard à la prise en charge par le concepteur. C'est ce que la deuxième hypothèse de recherche tente de valider.

Hypothèse 2 : Il est possible de réduire le degré de généralité pour obtenir un modèle de référence de taille raisonnable.

Finalement, dans l'affirmative des hypothèses 1 et 2, il est probable d'obtenir un modèle particulier qui dénombre les SV et combinaison de SV pertinentes et particulières à transformer en VUES. C'est ce que la troisième hypothèse de recherche tente de valider.

Hypothèse 3 : Le modèle de référence peut être utilisé comme base pour obtenir un modèle particulier.

Dans le cadre de ce papier, seul le travail de validation de l'hypothèse 1 est présenté.

3 VALIDATION DE L'HYPOTHESE 1 : LA PERTINENCE D'UNE PA EST CONTINGENTE.

La confirmation ou l'infirmité de l'hypothèse 1 consiste à évaluer la pertinence des PA dans différents contextes d'étude.

Cette évaluation est réalisée en 3 étapes :

- Etape 1 : recueil et préparation des données (Activité 1, Tableau 1)
- Etape 2 : Application des PA (Activité 2, Tableau 1)
- Etape 3 : Evaluation de la pertinence des PA

3.1 Etape 1 : recueil et préparation des données

Les contextes d'étude représentent 19 GP issues de 19 cas industriels (Cas 1...Cas 19). Les secteurs d'activité des entreprises étudiées concernent principalement : l'aéronautique, l'automobile et l'équipement électrique. Le

Tableau 6 présente une description des 19 DF extraites respectivement à partir des 19 GP industrielles.

Tableau 6. Description des 19 DF industriels étudiés

	Cas d'étude industriels (Ci)																		
	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11	C 12	C 13	C 14	C 15	C 16	C 17	C 18	C 19
Nombre de PDC	12	40	16	28	26	46	19	73	87	14	17	104	29	21	90	43	31	23	25
Nombre de produits	7	30	18	2	32	19	206	47	47	51	30	183	3019	2262	110	110	85	85	85
Nombre d'arcs	12	54	22	31	43	86	99	462	495	48	31	386	296	205	479	312	141	83	94
Densité [%]	9	4	9	4	6	4	27	8	7	25	10	4	35	46	6	17	14	15	15

Pour chaque DF industrielle relative à une étude de cas, le Tableau 6 spécifie le nombre des PDC, le nombre des produits, le nombre d'arcs représentant le flux des produits reliant les PDC. La dernière ligne de ce tableau donne une mesure de la densité des arcs dans les DF industrielles étudiées. Dans l'ensemble des 19 DF industrielles étudiées, la densité varie de 4 % à 46%. Ces valeurs de densité sont caractéristiques des graphes industriels : ce sont des graphes à faible densité.

3.2 Etape 2 : application des PA

L'étape 2 consiste à appliquer les 10 PA recensées aux 19 DF industrielles. Cette étape est déclinée en 2 temps (spécifications et résultats de calcul).

Pour chaque PA :

- Spécification de l'algorithme appliqué, de l'indicateur de performance (IP) mesuré, du domaine de définition de l'IP, de la mesure optimale de l'IP (MO : valeur cible). Le Tableau 7 présente ces spécifications de calcul.

Tableau 7. Spécifications de calcul

PA	Références algorithme	Nom de l'indicateur	Domaine de définition	MO
classe	(De Gujo, 1999)	Nombre de classes	[0, Nbre PDC [Nbre de PDC
		% inter classe B	[0,50[0
		% inter classe P	[0,50[0
Ordre total	(Zhou et al., 1996)	% FASB	[0,50[0
		% FASP	[0,50[0
Ordre partiel	(Zhou et al., 1996)	% Ordre partiel	[0,100[100
Niveau	(Kahn, 1962)	% Niveau	[0,100]	100
Relation		Flux pondéré entre PDC : Test d'adéquation loi de Pareto (principe des 80-20)	{rejet : 0 ; non-rejet : 1}	1
Squelette	(Zhou et al., 1996)	% Transitif B	[0, 100[100
		% Transitif P	[0, 100[100
Source totale		% Source T] 0, 100]	0
Source partielle	(Zhou et al., 1996)	% Source P	[0, 100]	0
Puits total		% Puits T] 0, 100]	0
Puits partiel		% Puits P	[0, 100]	0

FAS : Feedback arc set

- Application des différents algorithmes de mesure consignés dans le Tableau 7 aux 19 DF industrielles. Le Tableau 8 présente le résultat de l'application.

Tableau 8. Résultats de calcul

PA	IP	Cas d'étude industriels (Ci)																		
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
Classe	n classes	2	7	2	2	4	4	3	7	3	3	2	3	10	9	2	2	4	3	3
	% Inter classe B	3	8	22	11	14	10	10	40	12	12	16	20	40	36	11	9	22	32	31
	% Inter classe P	8	11	18	9	19	13	2	34	6	12	19	20	49	37	7	7	27	30	29
Ordre total	% FAS B	0	3	0	9	2	14	2	31	31	11	1	13	5	5	18	24	22	20	20
	% FAS P	0	7	0	9	1	19	29	35	37	33	3	19	27	29	30	35	35	31	34
Ordre partiel	% Ordre partiel	16	70	82	64	77	71	42	76	81	35	76	81	58	52	71	62	74	39	88
Niveau	% Niveau	83	52	29	50	42	46	84	37	33	80	47	26	68	81	46	63	55	48	48
Relation	Test Pareto	1	0	0	1	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0
Squelette	% Transitif B	0	15	0	3	16	14	51	41	37	38	23	44	51	73	44	45	34	34	34
	% Transitif P	0	8	0	3	12	14	60	41	41	44	25	48	73	85	38	46	24	50	43
Source totale	% Source T	8	8	6	7	27	4	0	0	0	7	6	8	0	0	1	0	3	13	16
Source partielle	% Source P	8	8	0	0	27	4	11	4	2	14	18	15	31	29	3	19	29	26	28
Puits total	% Puits T	8	3	6	7	12	2	11	0	1	7	12	7	3	5	1	0	0	0	0
Puits partiel	% Puits P	8	5	0	0	31	9	52	16	14	7	18	44	38	43	0	7	19	22	28

3.3 Etape 3 : Evaluation de la pertinence des PA

L'étape 3 consiste à discuter la question de la pertinence des PA. Cette étape est déclinée en 2 points de vue : Répartition des valeurs d'IP par rapport aux MO (Tableau 9) ; Répartition des valeurs d'IP par rapport à une zone de pertinence

- Répartition des valeurs d'IP par rapport aux MO (Tableau 9). Il s'agit de comparer les différentes valeurs d'IP des 19 cas d'études aux MO respectives. Les comparaisons sont réalisées sur la base des histogrammes de répartition des valeurs d'IP (Tableau 8). La Figure 3 présente les histogrammes de répartition des valeurs d'IP associés à la pertinence des PA Classe et Ordre total.

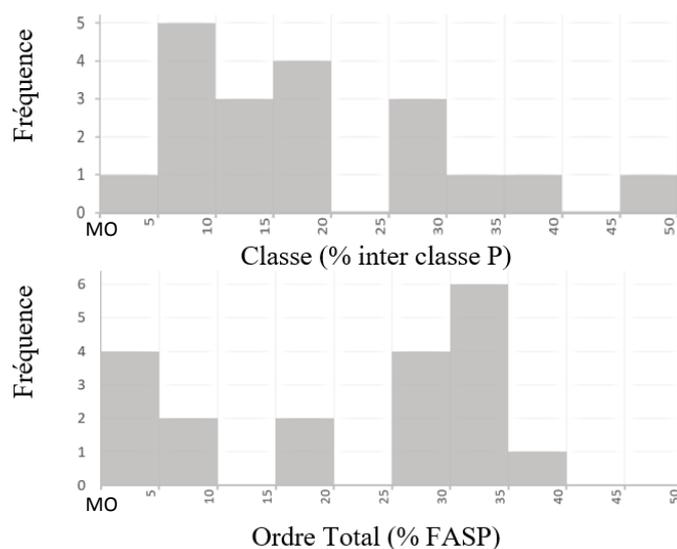


Figure 3. Histogramme de répartition (Classe, Ordre)

Sur les histogrammes ci-dessus, la répartition des mesures pour les 19 cas industriels *n'est pas homogène*. Des valeurs remarquables (proche de la MO) sont observées pour les 2 IP. Dans le cas du PA Classe : le taux de flux inter-cellules est inférieur à 30 % dans 68 % des cas industriels (soit 13 cas d'étude sur un total de 19 cas). Dans le cas du PA Ordre, le taux de retours arrière est inférieur à 30% dans 63 % des cas industriels (soit 12 Cas d'étude sur un total de 19 cas).

- Répartition des valeurs d'IP par rapport à une zone de pertinence. Pour prolonger la discussion, une zone de pertinence est définie selon 2 règles :
 - 1. Pour chaque IP, mesurer l'écart entre la MO et la valeur d'IP la plus proche (Tableau 9),
 - 2. Identifier le plus grand écart calculé.

Tableau 9. Identification de la zone de pertinence

PA	IP	Domaine de définition	MO	Meilleure mesure obtenue	# Cas	Ecart entre la MO et la meilleure valeur obtenue
classe	% inter class B	[0,50[0	2	C1	2
	(% inter class P)	[0,50[0	3	C7	3
Ordre total	% FASB	[0,50[0	0	C1,C2	0
	(% FASP)	[0,50[0	0	C1,C2	0
Ordre partiel	% Ordre partiel	[0,100[100	88	C19	12
Niveau	% Niveau	[0,100]	100	84	C7	16
Relation	Test Pareto	{ 0 ; 1 }	1	1	C1,C4,C7,C9,C15,C16	0
Squelette	% Transitif B	[0, 100[100	59	C14	27
	% Transitif P	[0, 100[100	73	C14	15
Source totale	% Source T] 0, 100]	0	0	C7,C8,C9,C13	0
Source partielle	% Source P	[0, 100]	0	0	C3, C4	0
Puits total	% Puits T] 0, 100]	0	0	C8, C16, C17,	0
Puits partiel	% Puits P	[0, 100]	0	0	C3, C4	0

Selon la définition ci-dessus, la zone de pertinence identifiée considère au moins une occurrence par IP. Aussi aucun PA n'est exclu de la discussion. La zone de pertinence est égale à l'écart maximum (Tableau 9) : 27 %, arrondi à 30% pour simplification. Notons que cette valeur couvre 60% du domaine de définition des IP Classe et Ordre.

A noter : En référence à la MO, selon le PA et l'IP considéré, le numéro de Cas le plus proche de la MO est variable. Ce constat incite à observer cette variabilité dans la zone de pertinence calculée.

Pour accélérer l'observation (Tableau 10) :

- Si la valeur de l'IP est incluse dans la zone de pertinence alors la valeur « 1 » (signifiant pertinent) est assignée à la cellule ;
- Sinon, la valeur « 0 » (signifiant non-pertinent) est assignée à la cellule.

Tableau 10. Impact de la zone de pertinence

PA	IP	Cas d'étude industriels (Ci)																	Total		
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17		C18	C19
Classe	% Inter classe B	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	15
	% Inter classe P	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	0	0	14
Ordre total	% FAS B	0	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	15	
	% FAS P	0	1	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	6	
Ordre partiel	% Ordre partiel	0	0	1	0	1	0	0	1	1	0	1	1	0	0	0	0	1	0	8	
Niveau	% Niveau	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	4	
Relation	Test Pareto	1	0	0	1	0	0	1	0	1	0	0	0	0	1	1	0	0	0	6	
Squelette	% Transitif B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	
	% Transitif P	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	
Source totale	% Source T	1	1	1	1	0	1	0	0	0	1	1	1	0	0	1	1	1	1	12	
Source partielle	% Source P	1	1	0	0	0	1	1	1	1	1	1	0	0	1	1	0	1	0	12	
Puits total	% Puits T	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	0	14	
	% Puits P	1	1	0	0	0	1	0	1	1	1	1	0	0	0	1	1	1	0	10	
Total		8	8	5	7	6	8	7	3	7	8	9	8	2	5	7	6	6	4	3	

Observations du Tableau 10 par ligne (par IP) :

- Les PA ne sont pas tous pertinents dans tous les 19 Cas.
- Le total par IP est variable (1...15).
- Aucune séquence (ligne) n'est superposable.

Observations du Tableau 10 par colonne (par Cas) :

- Les PA ne sont pas tous pertinentes.
- Le total par Cas est variable (2...9).
- Aucune séquence (colonne) n'est superposable.

En guise de conclusion dans cette section, les connaissances capitalisées ne permettent pas de rejeter l'Hypothèse 1.

4 CONCLUSION

Le travail de ce papier s'inscrit dans le cadre de la mesure de la maturité du système physique de production (SPP). L'objectif final est l'élaboration d'une méthode intégrée d'aide au diagnostic. Le travail de recherche est séquentiel et consiste à valider 3 hypothèses de recherche (H1...H3) : H1. La pertinence d'une PA est contingente ; H2. Il est possible d'obtenir un modèle de référence pour le diagnostic du SPP ; H3. : Le modèle de référence peut être utilisé comme base pour obtenir un modèle particulier pour un SPP particulier. Le travail de ce papier a permis de valider H1. La validation des hypothèses H2 et H3 est en cours de réalisation.

5 REFERENCES

Alfnes, E. & Martinsen, "Modeling and design of flow manufacturing systems for SMEs" Intelligent Computation in manufacturing Engineering, 2006

Amoako-Gyampah, K., and M. Acquah (2008). Manufacturing Strategy, Competitive Strategy, and Firm Performance: An Empirical Study in a Developing Economic Environment. International Journal of Production Economics, 2, 575-92.

Barth M., De Guio R., Industrial Implementation of Production Flow Analysis. In IRANI A.S. (ed.) Handbook of Cellular Manufacturing Systems, John Wiley & Sons. 1999,

Barth M., De Guio R., Zhou J., " A graph theoretic approach

for the production line formation problem " International Conference on Engineering Design and Automation, Maui, Hawaii, August 9-12, 1998.

Buchanan J.T., Daellenbach H.G., "A comparative evaluation of interactive solution methods for multiple objective decision models", European Journal of Operational Research, 24, 353-359, 1987.

Card, S.K., J.D. Mackinlay, and B. Shneiderman, Readings in Information Visualization: Using Vision to Think. San Francisco, California: Morgan-Kaufmann, 1999

Daita, S.T., Irani, S.A. and Kotamraju, S. Algorithms for production flow analysis. Int.J.Prod.Res, Vol 37(11), 2000

De Guio and M. Barth, 1999. Cell Formation Using Production Flow Analysis. Handbook on Cellular Manufacturing. J. Wiley, 1999.

De Guio R., M. Barth and J. Zhou, "Checking Algorithms for the Weighted Feedback Arc Set problem with Order Constraints", IEEE proceedings CESA'96, Lille, France July 9-12, 1996.

Deb, S.K. and Bhattacharyya B., 2002. Fuzzy decision support system for manufacturing facilities layout planning. Decision Support Systems 40 (2005) 305– 314.

Hollier R.H, « The layout of Multi-Product Lines », Int. J. Prod. Res, vol 2, 47-57, 1963.

Irani, S. A., Zhang, H., Zhou, J., Huang, H., Udai, T. K. & Subramanian S. Production Flow Analysis and Simplification Toolkit. Int. J. Prod. Res, 38(8), 1855-1874. (2000).

Kaspar, S., and M. Schneider. (2015). Lean Und Industrie 4.0 in Der Intralogistik: Effizienzsteigerung Durch Kombination Der BeidennAnsätze. Productivity Management 20(5), 17–20.

Koren, Y., Hu J., and Weber T. (1998) Impact of Manufacturing System Configuration on Performance , In: CIRP Annals, 1, 689-698.

Lee, D.H., Kiritsis, D., Xirouchakis, P., " Iterative approach to operation selection and sequencing in process planning" Int J Prod Res. N. 42, V.22,p. 4745-4766, 2004.

Leone, G. and Rahn, R.D. Fundamentals of flow manufacturing. Flow publishing Inc. Colorado, USA. 2002

Mahseh O. Srinivasan G. « Incremental cell formation considering alternative machines » int. j. prod. res., vol. 40, no. 14, 3291-3310, 2002.

Muther, R. Systematic Layout Planning, second ed., Cahners Books, Boston, 1973.

Nicoletti, B. 2013. Lean and Automate Manufacturing and Logistics. In Advances in Production Management Systems, APMS, 2013, 278–285. State College, PA: Springer.

Olhager, J. 2013. Evolution of Operations Planning and Control: From Production to Supply Chains. International Journal of Production Research, 51 (23–24), 6836–6843.

Pahl G., Beitz W. Engineering Design: A Systematic Approach, Springer-Verlag, 2nde edition, 1996.

Phillips, E.J. (1997)."Manufacturing Plant Layout: Fundamentals and Fine Points of Optimum Facility Design " Dearborn, MI: Society of Manufacturing Engineers.

Raoot, A.D. and Rakshit, A. (1993). "A 'linguistic pattern' approach for multiple criteria facility layout problems", In: International Journal of Production Research, 31:203-222.

Rouse, P and Putterill M. (2003). An integral framework for performance measurement. Management Decision 41(8), 791-805.

Roy B., Multicriteria methodology for decision aiding. Boston: Springer, 1996. DOI 10.1007/978-1-4757-2500-1.

- Sarker, B.R. and Y. Xu. " Designing multi-product lines: job routing in cellular manufacturing systems ", IIE Transactions. V32 (3), p. 219-235. 2000
- Selim, H.M., Askin, R.G., Vakharia, A.J.,"Cell formation in group technology: review, evaluation and directions for future research", Int. J. Comput. Ind. Eng, Vol.34, 1998.
- Singh, S.P and Sharma R.R.K. (2006). "A review of different approaches to the facility layout problem", In: The International Journal of Advanced Manufacturing Technology, Vol 3 425-433.
- Tan, K. C. (2001). A framework of supply chain management literature. European Journal of Purchasing & Supply Management, 7(1), 39-48.
- Tompkins, J. A., White, J. A., Bozer, Y. A., Tanchoco, J. M. A., Facilities planning. John Wiley & Sons, 2003.
- Yin, Y., Yasuda, K. (2006). "Similarity coefficient methods applied to the cell formation problem: A taxonomy and review", In: International Journal of Production Economics, Vol. 101, n° 2, p. 329–352.
- Zhou J, Irani S. A new flow diagramming scheme for mapping and analysis of multi- product flows in facility. Journal of Integrated Design and Process Science, March, Vol. 7, No. 1, pp. 25-58. 2003.
- Zhou J., Barth M. and DeGuio R .Quasi-source heuristic for the production line formation of manufacturing system. IEEE International Conference on SMC, China, 14-16, 1996.

CIGI 2021

La transformation digitale de la chaîne d'approvisionnement

MOHAMED-ILIASSE MAHRAZ¹, LOUBNA BENABBOU², ABDELAZIZ BERRADO¹

¹DEPARTEMENT GENIE INDUSTRIEL, ECOLE MOHAMMADIA D'INGENIEURS,
Mohammed V University of Rabat, Morocco
M.iliasse.mahraz@gmail.com, Berrado@emi.ac.ma

² DEPARTEMENT SCIENCES DE LA GESTION,
Université du Québec à Rimouski (UQAR) Campus de Lévis, Québec, Canada
Loubna_benabbou@uqar.ca

Résumé – L'émergence de nouveaux outils basés sur l'Intelligence Artificielle, et certaines transformations technologiques basées sur les Big Data, ainsi que le développement de l'utilisation de l'Internet des Objets, poussent à un changement profond affectant tous les processus métiers de l'entreprise. Aujourd'hui, nous en parlons également de la transformation digitale dans le domaine de la chaîne d'approvisionnement. Dans un monde où la circulation des biens et des informations est de plus en plus dense et rapide, les entreprises sont confrontées à un défi. En effet, elles doivent intégrer ces nouvelles technologies afin de faire évoluer leur organisation et leur logistique pour développer l'agilité de la chaîne d'approvisionnement. La révolution digitale a créé de nouveaux besoins, mais elle a également permis d'optimiser les méthodes et les processus. Gagner en agilité dans la chaîne d'approvisionnement est essentiel pour relever les nouveaux défis du marché, où les flux s'accélèrent, et cela ne peut se faire que par une transformation digitale. L'objectif principal de cette étude est de fournir une revue de littérature systématique de la transformation digitale dans la chaîne d'approvisionnement. Une revue des articles relatifs à ces sujets a été réalisée à travers la base de données "Scopus". Elle donne également un état des lieux complet de la transformation digitale grâce à une revue de 84 articles qui peuvent servir de base à une feuille de route pour la transformation digitale dans la chaîne d'approvisionnement.

Mots clés – Chaîne d'approvisionnement, Transformation digitale, Revue systématique

Abstract –The emergence of new tools based on Artificial Intelligence, and some technological transformations based on Big Data, as well as the development of the use of Internet of Objects, push for a profound change affecting all the business processes of the enterprise. Today, we also talk about it in the field of the supply chain. In a world where the flow of goods and information is getting denser and faster, companies are facing a challenge. Indeed, they must integrate these new technologies in order to evolve their organization and their logistics to develop the agility of the Supply Chain. The digital revolution has created new needs, but it has also allowed the optimization of methods and processes. The implementation of tools to both improve, optimize supply by lowering costs and delivery times, but also other tools to track products in different stages of the supply chain . Gaining agility in the supply chain is essential to meet the new challenges of trade, where flows are accelerating, and this can only be done through a digital transformation. The main objective of this study is to provide a systematic literature review of digital transformation in the supply chain. A review of articles related to the topics was done within "Scopus". It also give a comprehensive state of the art of digital transformation through a review of 84 that can serve as a basis for a roadmap for digital transformation in the supply chain.

Keywords – Supply Chain, Digital transformation, Systematic Review

1 INTRODUCTION

Aujourd'hui, les entreprises se transforment grâce aux nouvelles technologies de l'information et de la communication, mais peu ont compris comment tirer parti de ce phénomène. La transformation digitale (TD) est essentielle pour toutes les entreprises, quels que soient leur taille et leur secteur d'activité. Au-delà de la dématérialisation des processus de travail, cette approche permet aux entreprises d'optimiser leurs opérations, et de gagner en performance, efficacité et compétitivité par l'adoption d'un nouveau mode de

gestion, de nouveaux outils, de nouvelles méthodes de travail mais aussi de nouvelles réflexions et organisations.

La TD est devenu un sujet de préoccupation important et un enjeu stratégique pour toutes les organisations. Il offre de nouvelles opportunités aux entreprises, au-delà de leurs activités traditionnelles, en accélérant leur croissance et en créant des avantages compétitifs durables et la sécurité des opérations.

La TD augmente la rentabilité de l'entreprise en simplifiant les processus et les interactions au sein de l'entreprise. Les technologies liées au TD telles que la Big Data, l'intelligence

artificielle, le Cloud Computing, les réseaux sociaux et l'Internet des objets, offrent de nouveaux usages basés sur l'innovation et axés sur les besoins du consommateur. L'utilisation de ces outils est devenue un enjeu vital pour les entreprises de tous types d'activités, de toutes tailles et de toutes nationalités. Toutefois, la mise en œuvre des nouvelles technologies ne suffit pas à faire de l'entreprise digitale un phénomène plus large, qui suppose non seulement l'utilisation de nouveaux outils digitaux mais aussi la modification des éléments critiques de l'entreprise, notamment la stratégie, le modèle commercial, les processus opérationnels, les structures organisationnelles et la culture organisationnelle [Arribas et Alfaro, 2018]. C'est une façon de s'approprier la culture, non seulement au sein des entreprises mais aussi dans la vie de tous les jours. La stratégie, ainsi que les changements apportés à une organisation, y compris sa structure, ses processus et sa culture, sont nécessaires pour générer la capacité de créer de nouvelles voies de création de valeur.

Ce phénomène, qui est né dans les années 2000 avec la démocratisation d'Internet, a un impact majeur sur la chaîne d'approvisionnement. La chaîne d'approvisionnement désigne la chaîne logistique d'approvisionnement et de livraison. En principe, les règles de la chaîne d'approvisionnement étaient assez simples. Trois groupes d'acteurs la composent : le fournisseur, le distributeur et le client final, avec un nouvel acteur : Le prestataire logistique, avec pour objectif d'évaluer avec précision les besoins, la disponibilité et les capacités de chaque maillon de la chaîne logistique et de fabrication, afin de mieux les synchroniser et de servir les clients dans les meilleures conditions possibles.

Dans un environnement de plus en plus concurrent, la transformation digitale de la chaîne d'approvisionnement est aujourd'hui plus nécessaire que jamais. Jusqu'à présent, la tâche de la chaîne d'approvisionnement était simple et la rentabilité est l'un des facteurs les plus importants de la stratégie d'approvisionnement. Aujourd'hui, les flux d'informations sont tout aussi critiques que les flux de marchandises, et les entreprises doivent donc s'adapter à la rapidité et à l'exactitude des données, qui sont essentielles pour atteindre les objectifs commerciaux. En outre, il s'avère que la collaboration est plus importante que jamais, voire essentielle. Ainsi, la transformation digitale et la digitalisation des flux d'information est un élément clé pour les entreprises et devient la clé de voûte pour être compétitif sur le marché.

Ces dernières années, un grand nombre d'études ont tenté de donner une définition et une solution pour gérer la transformation digitale dans la chaîne d'approvisionnement. En dépit de ce nombre important d'études qui ont été menées, on a tenté de traduire systématiquement ces conclusions en un examen complet. Cependant, il manque encore des cadres et des lignes directrices pour les entreprises sur la manière de naviguer dans des changements aussi radicaux [Bygstad et Aanby, 2017] [Chaniyas et Hess, 2016].

Dans cet article, nous avons l'intention de fournir une vue globale sur la transformation digitale de la chaîne d'approvisionnement avec de nouvelles perspectives, basée sur une revue systématique de la littérature qui nous permet d'explorer tous les aspects de la littérature existante et des preuves empiriques. Pour cela, nous avons analysé la littérature universitaire et professionnelle récente afin de trouver des recherches dans ce domaine. Nous donnerons un aperçu de la méthodologie de recherche et des résultats dans la prochaine section.

2 REVUE SYSTEMATIQUE SUR LA TRANSFORMATION DIGITALE DE LA CHAÎNE D'APPROVISIONNEMENT :

2.1 Phase de planification :

- **Repérage des recherches :**

Pour atteindre les objectifs de ce document, une approche d'examen systématique est adoptée. Comme première étape de notre étude, nous avons précisé un problème et reformulé notre objectif sous la forme des questions claires, structurées et sans ambiguïtés. La transformation digitale de la chaîne d'approvisionnement, ainsi que son impact sur les entreprises industrielles, présentent un intérêt particulier. Pour cela, nous avons étudié les questions suivantes.

QR1 : Comment pouvons-nous définir la transformation digitale et la chaîne d'approvisionnement ?

QR2 : Quel est l'impact de la transformation numérique sur la chaîne d'approvisionnement ?

2.2 Phase de sélection :

- **La stratégie de recherche :**

Dans cet article, nous avons classé et passé en revue les études pertinentes afin de mieux comprendre la transformation digitale de la chaîne d'approvisionnement. La revue de littérature s'est concentrée sur les articles scientifiques publiés en anglais entre 2015 et 2020, et est examinée dans la principale base de données en ligne de littérature évaluée par des pairs : Scopus.

Les mots clés n'étaient pas prédéterminés avant la recherche, mais ils ont été progressivement sélectionnés au cours du processus de sélection par le biais d'articles et de documents antérieurs dans le même domaine ou avec une portée similaire, et les termes finaux utilisés peuvent être vus dans le tableau 1. La méthode de recherche utilisée dans notre étude est appelée une recherche booléenne (AND/OR). Nous avons essayé de combiner différents termes en utilisant les opérateurs logiques OR et AND. Un ensemble de mots clés pertinents a été sélectionné par le biais d'articles et de documents précédents dans le même domaine ou avec une portée similaire. Les différentes compositions de termes ont été recherchées dans les mots clés, les résumés et le titre. La chaîne de recherche utilisée est : [(“Digital transformation” OR “Smart” OR “Digital” OR “Agility” OR “Flexibility” OR “Durability”) AND (“Supply chains” or “logistics”)].

Figure 1. Méthodologie d'examen



Grâce à la base de données Scopus, nous avons effectué nos recherches en utilisant différentes méthodes avec une configuration spécifique de la chaîne de recherche. Scopus est considéré comme la plus grande base de données de citations, de résumés de bibliographies évaluées par des pairs et de sites web de qualité, et couvre un large éventail de disciplines et de

sujets. Elle comprend des outils intelligents permettant de suivre, d'analyser et de visualiser les recherches.

Tableau 1. Processus d'examen systématique

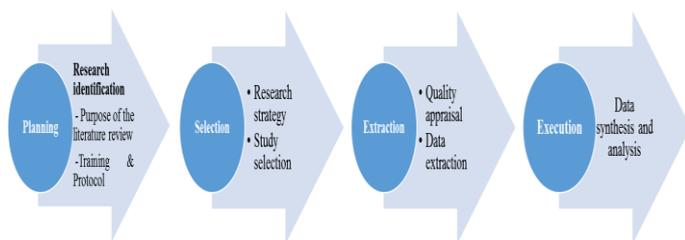
Institut pour les informations scientifiques – SCOPUS	
Critères	Filtres
Restrictions	Sujet (Titre, Résumé, Mots clés)
Type de documents	Articles de journal et Actes de conférences
Language	Anglais
Année	2015-2020

Cette étude vise à examiner la littérature sur la transformation digitale de la chaîne d'approvisionnement. Pour cela, nous avons utilisé une approche systématique suivant le protocole d'Okoli et Schabram (2010) pour examiner un grand nombre d'articles publiés entre 2015 et 2020, dont on a pu extraire un grand nombre de documents relatifs à la transformation digitale, et qui ont été sélectionnés pour être analysés.

L'examen systématique nous a permis de rassembler toutes les informations pertinentes et valides concernant la transformation digitale grâce à son approche rigoureuse et bien définie qui suit des règles strictes et un protocole clairement défini, celui-ci vise à limiter l'erreur systématique (biais), principalement en essayant d'identifier, d'évaluer et de synthétiser toutes les études pertinentes en utilisant une méthodologie spécifique.

Notre revue systématique a comporté quatre étapes suivant le protocole d'Okoli et Schabram :

Figure 1. Protocole d'examen systématique d'Okoli et Schabram (2010)



• Évaluation de la qualité :

La sélection initiale n'a pas été limitée et a inclus dans les résultats des articles d'autres domaines de recherche puisque nous n'avons pas appliqué de restrictions. Étant donné que l'objectif de ce document est de fournir une vue d'ensemble des domaines de recherche qui se concentrent sur les questions relatives aux contextes de transformation digitale et d'explorer son impact sur les entreprises industrielles, nous avons affiné les résultats de recherche en fonction de certaines mesures de qualité.

Tout d'abord, une vérification manuelle du contenu des articles complets, des résumés et du titre a été effectuée. Ensuite, nous avons défini les critères de sélection des études à inclure ou à exclure.

-Critères d'inclusion :

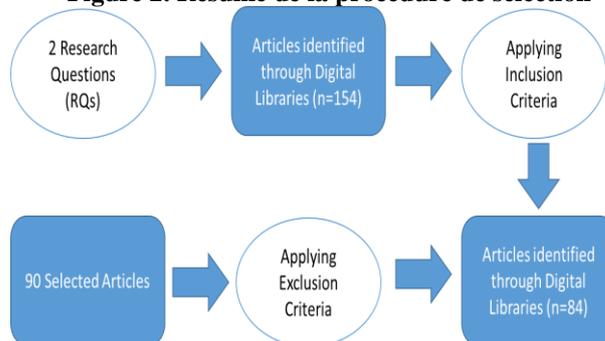
- L'examen doit être lié à la transformation digitale et doit être en rapport avec la question de recherche.
- La publication doit avoir une méthodologie transparente et calibrée.
- La recherche doit provenir de ressources et de revues fiables
- L'année de publication des articles doit être comprise entre 2015 et 2020.

-Les articles doivent être rédigés en anglais.

-Critères d'exclusion :

- L'étude est un éditorial, un Keynote, une opinion, un tutoriel, un atelier, un rapport de synthèse, un poster ou un article tel que des articles non publiés, des thèses de master et des livres. Ces documents ont été exclus car ces articles ne sont généralement pas évalués par des pairs.
- Le texte intégral de l'étude n'est pas disponible.

Figure 2. Résumé de la procédure de sélection



Les critères d'exclusion ont été appliqués en utilisant l'opérateur logique OU entre eux.

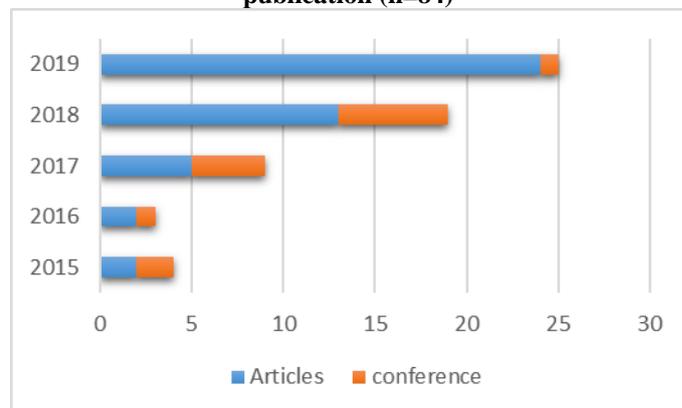
L'application de la chaîne de recherche dans la base de données Scopus nous a permis de générer 154 articles. Nous avons commencé la phase de filtrage des articles en appliquant les critères d'inclusion. Nous avons limité notre champ d'étude aux articles de journaux et aux actes de conférences publiés entre 2015 et 2020 et rédigés en anglais. Ensuite, en fonction de leurs titres, mots-clés, résumés et peut-être le texte complet, nous avons pu sélectionner les articles qui nous permettront de répondre aux QRs. Le processus d'exclusion a donné lieu à 84 articles issus de la base de données Scopus.

2.3 Phase d'extraction :

• L'occurrence des articles par année de publication :

Cette recherche a montré que le concept de transformation digitale de la chaîne d'approvisionnement est à ses débuts de développement et de recherche dans les milieux universitaires, alors qu'il est bien connu et discuté par les praticiens. De 2015 à 2020, un grand nombre d'articles sur la TD de la chaîne d'approvisionnement sont publiés chaque année. La répartition des articles au cours de la période de recherche (2015-2020) est présentée sur la figure 2.

Figure 3. La fréquence des articles par année et type de publication (n=84)



68 % du nombre total de documents sont des articles de journaux et 32 % des actes de conférences. L'analyse de la figure 2. montre qu'un nombre élevé de publications a été enregistré entre 2018, 2019 et 2020, sachant que le travail se poursuit cette année, et l'on s'attend à l'augmentation du nombre d'articles publiés d'ici la fin de l'année.

- **Répartition des articles par revue :**

Les auteurs ont utilisé 60 types de revues différentes pour diffuser les résultats de leurs recherches. La revue qui a eu le plus grand nombre de publications dans ce domaine est l'International Journal of Supply Chain Management, qui publie des recherches empiriques de haute qualité sur la gestion de la chaîne d'approvisionnement qui ont un impact significatif sur les théories et les pratiques de la chaîne d'approvisionnement.

Tableau 2. Répartition des articles par revue

Top 5 Publications Journal	Count	% of 84
International Journal of Supply Chain Management	11	13.10%
Production Planning and Control	7	8.33%
IEEE Transactions on Engineering Management	4	4.73%
Business Process Management Journal	2	2.83%
Computers in Industry	2	2.83%
Autres	58	69.04%

- **Les 5 premiers auteurs sur le domaine par numéro de citation :**

Nous avons également extrait les cinq premiers auteurs de la matière en fonction du nombre de citations.

Tableau 3. Les cinq premiers auteurs sur le domaine

Auteurs	Publication	Citations	Citations/Années
Kache et al. (2017)	International Journal of Operations and Production Management	65	32.50
Ivanov et al. (2019)	International Journal of Production Research	25	25
Büyüközkan et al. (2018)	Computers in Industry	16	16
Ali et al. (2016)	Computer Journal	15	5
Bechtsis et al. (2018)	Journal of Cleaner Production	12	12

3 LA TRANSFORMATION DIGITALE :

3.1 La Transformation digitale comme un nouveau concept :

La transformation digitale est un concept relativement nouveau et a atteint une grande popularité parmi les chercheurs et les praticiens au cours des dernières années. En perpétuelle évolution, nous assistons à une véritable révolution

introduite par les organisations. Pour bien comprendre le concept de transformation digitale, nous devons d'abord définir et différencier les termes de numérisation, digitalisation et transformation digitale [Fichman et al., 2014]. L'utilisation de ces termes jusqu'à présent crée une certaine ambiguïté au sein de la communauté scientifique puisque la plupart des chercheurs utilisent les termes de manière interchangeable, alors qu'en réalité, il existe une différence significative entre les termes. Il est nécessaire de faire la distinction entre ces trois termes. En effet, la numérisation signifie la conversion d'un format analogique en un format numérique ; c'est la représentation d'informations avec un langage binaire (0 et 1). C'est la représentation de l'information dans tout système informatique. Par exemple, on peut scanner une photo avec un scanner pour la stocker sur un CD.

Alors que la digitalisation signifie l'utilisation de données et de technologies numériques pour automatiser le traitement des données et optimiser les processus, c'est l'utilisation d'un système informatique pour automatiser ou semi-automatiser les processus [Bharadwaj et al., 2013]. De ce point de vue, l'informatisation est un terme proche de la digitalisation et est parfois considérée comme un synonyme.

Cependant, la transformation digitale est un nouveau concept fréquemment utilisé par les universitaires et les praticiens, mais qui change de forme très souvent en fonction du contexte dans lequel il est utilisé. En parcourant la littérature, nous avons remarqué que le terme "transformation digitale" est très largement utilisé, parfois à tort, car il n'existe pas de définitions formelles dans la littérature universitaire. De nombreux auteurs ont tenté de le définir et de discuter de son concept, le tableau.1 illustre certaines définitions uniques.

Tableau 4. Quelques Définitions actuelles de la transformation digitale

Définitions	Sources
L'utilisation de nouvelles technologies numériques (médiâs sociaux, mobiles, analytiques ou appareils intégrés) pour permettre des améliorations économiques importantes (comme l'amélioration de l'expérience client, la rationalisation des opérations ou la création de nouveaux modèles économiques).	Liere-Netheler et al. (2018)
L'utilisation de la technologie pour améliorer radicalement les performances ou la portée des entreprises.	Westerman et al. (2015) Westerman et al. (2014)
La transformation digitale englobe la digitalisation des canaux de vente et de communication et la digitalisation des offres des entreprises (produits et services), qui remplace ou augmente les offres physiques. En outre, la transformation digitale implique des mouvements commerciaux tactiques et stratégiques qui sont déclenchés par des informations basées sur des données et le lancement de modèles commerciaux digitaux qui permettent de nouvelles façons de saisir la valeur.	Horlach et al. (2017)

Le terme est assez ancien, il est apparu pour la première fois en 2000 [Patel et McCarthy, 2000] et était lié à la digitalisation, mais il désigne maintenant un phénomène lié

aux nouveaux usages des consommateurs et aux objets uniques qui ont un impact direct sur les modèles commerciaux et les organisations actuels. La première définition stricte du terme apparaît dans un livre de Fors et Stolterman [Fors et Stolterman, 2004], où ils décrivent la transformation digitale comme les changements que la technologie numérique entraîne ou influence dans tous les aspects de la vie humaine. D'autres chercheurs tels que [Westerman et al., 2011] définissent la transformation digitale comme l'utilisation de la technologie pour améliorer radicalement la performance ou la portée des entreprises, et la création de nouvelles opportunités commerciales par l'utilisation de données et de technologies numériques.

Ce ne sont plus seulement les processus commerciaux ou les macro-processus qui sont affectés par le changement, mais tous les processus commerciaux, les processus de soutien et la conception organisationnelle elle-même. Ces processus sont entièrement dédiés à la création de valeur. En effet, les entreprises sont appelées à repenser leurs opérations et la manière dont elles interagissent avec leurs parties prenantes, à s'adapter aux nouveaux comportements du marché ou à tirer parti des outils technologiques en termes d'innovation ou de productivité. La TD est un processus de transformation perturbateur ou incrémentiel. Il commence avec l'adoption et l'utilisation des technologies numériques, puis évolue vers un changement holistique implicite d'une organisation, ou délibérément vers la création de valeur [Henriette et al. 2016].

3.1 Composantes et technologies de la transformation digitale de la chaîne d'approvisionnement :

Aujourd'hui omniprésent dans notre vie quotidienne, le digital est plus développé dans les entreprises. La chaîne d'approvisionnement, fonction stratégique de l'entreprise, n'échappe pas à ce constat. Les technologies digitales offrent différentes solutions qui répondent aux enjeux de la chaîne d'approvisionnement des entreprises et aux attentes du marché. De nombreuses tendances significatives de la numérisation pourraient être appliquées dans une chaîne d'approvisionnement pour améliorer sensiblement l'avenir. Neuf différents facilitateurs des technologies digitales seront explorés dans cette étude. Ces neuf domaines sont illustrés et expliqués sur le tableau 4.

Tableau 7. Les technologies utilisées dans la transformation digitale de la chaîne d'approvisionnement

Technologie	Description
Big Data	<i>BD refers to a large or complex amount of continuously changing data that can no longer be analyzed with conventional analysis methods and data processing tools. On the other hand, this also includes the sum of the data processing technologies and analysis methods that have been developed in recent years to collect and evaluate a large or complex amount of data. [Wang et al., 2016]</i>
Internet of Things (IoT)	<i>IoT refers to the networking of physical objects via sensors and actuators with the Internet or other networked systems. It allows objects to communicate with each other independently and perform tasks for the user. In addition, users can track and monitor the status of objects and even remotely control them. Besides,</i>

	<i>networked sensors also monitor the environment, humans [Fedyk, 2016][Manyika et al., 2015].</i>
Social Media (SM)	<i>Social media can be used to create relationships and can be used in a supply chain to build and expand relationships between business stakeholders. The information and knowledge accumulated through the use of social media by supply chain partners can be a valuable source of insight into the various issues in the supply chain. It can also be used to build relationships and identify key performance indicators. Some companies use social media to solicit feedback from their consumers to make improvements. [Perrin,2015]</i>
Robotics (R)	<i>R technology in Logistics is a branch of engineering that involves the conception ,design ,manufacture, and operation of R. [Siciliano & Khatib,2016]</i>
3D printing	<i>3D printing, also known as additive manufacturing – AM (the terms 3D printing and additive manufacturing have become interchangeable), is an additive technology used for making three dimensional solid objects up in layers from a digital file without the need for a mould or cutting tool. [Kubáč et al., 2017]</i>
Nanotechnology (NT)	<i>Nano Technology is the manipulation of functional systems on an atomic, molecular, and supramolecular scale. [Schmidt et al., 2016]</i>
Augmented Reality (AR)	<i>Augmented Reality refers to the layering of computer simulation models over the physical layout of current surroundings. In a sense, this is the hallmark of virtual reality, but AR refers to using this information to improve the efficiency of today's processes as they relate to the supply chain. [Merlino et al., 2017]</i>
Sensor Technology (ST)	<i>The adoption of sensor technology is essential and can make viable those unit processes that are inherently so complex or unstable that they demand continuous monitoring and control to ensure acceptable yields.</i>
Cloud Computing (CC)	<i>Cloud computing is an IT service model where computing services (both hardware and software) are provided on demand to customers through self-service, regardless of device and location. [Schmidt et al.,2016]</i>

3.2 Les traits de la transformation digitale :

La transformation digitale est l'un des sujets qui occuperont longtemps l'agenda des études et de la gestion de nombreuses entreprises. Il s'agit d'une tendance permanente qui est constamment renouvelée par les nouvelles générations de technologies numériques. Trois caractéristiques peuvent décrire la transformation numérique.

Premièrement, la transformation digitale est irréversible. Les nouvelles technologies numériques ou les nouveaux concepts de déploiement de technologies peuvent être moins efficaces au début de leur mise sur le marché que les technologies existantes ou leurs concepts de déploiement. Cependant, la situation s'améliore à mesure que la transformation arrive à maturité, en retirant les solutions préétablies de leur position dominante sur le marché. Ainsi, l'utilisateur d'une innovation numérique ne veut pas renoncer au confort que son utilisation lui procure. Renoncer à l'utilisation des nouvelles technologies est impensable pour beaucoup de gens, même temporairement. Deuxièmement, la transformation digitale est inévitable. Le contexte sociétal et économique actuel impose l'utilisation innovante des technologies numériques pour répondre aux défis de l'urbanisation et de la mondialisation des activités économiques. Aujourd'hui, toutes les entreprises sont appelées à effectuer cette transformation pour mieux intégrer le nouvel écosystème numérique et réaliser ce saut vers la modernité. Les progrès des technologies numériques permettent aujourd'hui aux entreprises de développer des solutions à des défis complexes. La transformation digitale a atteint un seuil de maturité, à partir duquel de nombreux problèmes jusqu'ici insolubles peuvent être résolus.

Enfin, les processus de transformation digitale sont encore incertains, mais inévitables et progressent extrêmement rapidement. Le rythme rapide du développement des technologies numériques et leurs effets dans de nombreux secteurs rendent difficile de prévoir quelles entreprises réussiront leur transformation. On peut également constater une accélération de l'innovation et du changement. Se réinventer constamment est devenu la revendication de nombreuses entreprises. La compétitivité a depuis longtemps cessé de reposer uniquement sur le produit physique, mais plutôt sur son intégration dans un écosystème de services intelligents.

La transformation digitale est inévitable, irréversible et inéluctable. Ces trois caractéristiques illustrent le fait que la transformation digitale de l'économie et de la société est un processus qui ne peut être arrêté et qui est en pleine expansion. Alors qu'il est en train de modifier profondément l'économie et le monde du travail d'une manière générale, le numérique accompagne les mutations profondes et rapides des métiers. Les possibilités offertes par les nouvelles technologies numériques poussent les entreprises à évaluer constamment leur potentiel de développement des modèles commerciaux existants.

3.3 L'impact de la transformation digitale sur la chaîne d'approvisionnement :

La transformation digitale de la chaîne d'approvisionnement permet aux entreprises de répondre aux nouvelles exigences des clients, aux défis de l'approvisionnement et aux attentes en matière d'amélioration de l'efficacité. Ça permettra aux entreprises de bénéficier d'un avantage qui se définit par l'amélioration des performances tant en interne qu'en externe.

La TD est un défi majeur pour les entreprises car il permet d'intégrer le client au cœur de la chaîne d'approvisionnement. C'est un domaine clé de différenciation qui peut également

améliorer de manière significative la compétitivité des entreprises.

Cette transformation permet aux entreprises d'être plus réactives et de mieux répondre aux besoins des clients finaux et d'améliorer le service à l'industrie et aux consommateurs. L'accès rapide à l'information est un facteur majeur.

L'accès rapide à des données fiables et pertinentes devrait contribuer à rendre la prise de décision plus efficace grâce à des données de grande taille et à l'intelligence artificielle, permettant ainsi de gagner en compétitivité et de s'adapter à la concurrence. Cette source d'information permet aux fabricants d'être plus efficaces dans la variation de leur activité et de limiter l'augmentation des coûts. La mise en place d'outils numériques transversaux permet de rassembler les entités et favorise l'adaptabilité et la réactivité des entreprises.

La TD est un axe d'amélioration de la performance globale des entreprises. Il génère des opportunités d'amélioration du fonctionnement des organisations.

Tableau 5. Avantages de la transformation digitale de la chaîne d'approvisionnement

Avantages	Descriptions
Flexibilité et réactivité accrue	-Amélioration de la fiabilité, de la réactivité et de la flexibilité grâce à l'échange rapide d'informations en temps réel [Dweekat et al. 2017] -Visibilité en temps réel [Kumar et al. 2016] -Amélioration du contrôle et de la traçabilité des produits [Uddin et Sharif, 2016] -Visibilité en temps réel des fluctuations de la demande et des capacités [Schrauf et Bertram, 2016] -Augmenter la réactivité [Hanifan et al. 2014]
Un engagement de bout en bout avec le client	-Améliorer le service aux clients [Shee et al. 2017] -Améliorer l'expérience du client [Tahiduzzaman et al. 2017]
Réduction significative des coûts	-Une intégration plus efficace des processus internes [Ferretti et Schiavone, 2016] [Mann, 2015] -l'efficacité opérationnelle [Li et Li, 2017] -Mieux surveiller et gérer les stocks [Reaidy et al. 2015]
Partager l'information et adopter des analyses avancées	-Améliorer la collaboration entre les parties prenantes du secteur, y compris la chaîne d'approvisionnement. [Menon et al. 2018] -Meilleure prise de décision [Menon et al. 2018] -Mieux comprendre et prévoir la demande et résoudre des problèmes jusqu'alors insolubles, voire inconnus, tout au long de la chaîne d'approvisionnement. [Menon et al., 2018]

4 CONCLUSION

Des changements révolutionnaires auront lieu suite à la transformation digitale de la chaîne d'approvisionnement qui est devenue plus sophistiquée, avec l'implication croissante et plus intégrée des nouvelles technologies. En conséquence, avec l'introduction d'Industry 4.0, le modèle traditionnel de la chaîne d'approvisionnement est rapidement transformé en un modèle de digitalisation et d'intégration de l'ensemble du processus, depuis la conception, la fabrication, le marketing, la livraison jusqu'au service client. Ce nouveau modèle de chaîne d'approvisionnement a amélioré l'efficacité, la transparence, la communication, la collaboration, la flexibilité et la réactivité entre les acteurs de la chaîne d'approvisionnement afin de

fournir des prévisions précises de la demande, de réduire les délais et les coûts des commandes, d'améliorer la qualité des produits et des services et de renforcer la satisfaction des clients. Il est important et essentiel que les entreprises soient très conscientes que la chaîne d'approvisionnement est devenue le cœur et le système circulatoire de tous les processus de la chaîne de valeur de l'entreprise qui doivent être entièrement numérisés. La transformation digitale se produit lorsque la chaîne d'approvisionnement est intégrée dans l'industrie 4.0 et la dirige. Cela signifie également que les partenaires de la chaîne d'approvisionnement, tels que les fournisseurs, les clients, les distributeurs, les détaillants et les transporteurs, doivent être connectés à la nouvelle chaîne d'approvisionnement, qui doit inclure l'achat numérique, la logistique numérique, l'entreposage intelligent, les analyses prescriptives de la chaîne d'approvisionnement, la gestion efficace des pièces de rechange et la planification et l'exécution intégrées. Une analyse de la transformation digitale des chaînes d'approvisionnement, un sujet de grand intérêt pour les praticiens et les chercheurs, a été présentée dans ce document. Les principaux objectifs de la recherche sont de répondre à des questions telles que l'état actuel de la TD des chaînes d'approvisionnement dans les études universitaires, et l'importance actuelle du numérique dans la chaîne d'approvisionnement.

Cet article a tenté de mettre en évidence certaines des questions liées à la TD de la chaîne d'approvisionnement en utilisant une approche de revue systématique de la littérature. Les résultats de notre étude contribuent à la recherche actuelle et visent à mieux décrire et analyser le domaine. Ils nous ont amenés à mettre en évidence la compréhension de la transformation digitale dans la chaîne d'approvisionnement et à identifier des lignes directrices de recherche utiles pour de futurs projets universitaires.

5 REFERENCES :

Adhiatma, A. (2020). Digital Knowledge Supply Chain for Creative Industry: A Conceptual Framework. *Advances in Intelligent Systems and Computing*, 993, 914-924, ISSN 21945357, doi:10.1007/978-3-030-22354-0_85

Agrawal, P. (2018). Digital supply chain management: An Overview. *IOP Conference Series: Materials Science and Engineering*, 455(1), ISSN 17578981, doi:10.1088/1757-899X/455/1/012074

Agrawal, P., Narain, R., & Ullah, I. (2019). Analysis of barriers in implementation of digital transformation of supply chain using interpretive structural modelling approach. *Journal of Modelling in Management*.

Ali, S. (2016). Supply-Chain Security of Digital Microfluidic Biochips. *Computer*, 49(8), 36-43, ISSN 00189162, doi:10.1109/MC.2016.224

Ângelo, A. (2017). Digital transformation in the pharmaceutical compounds supply chain: Design of a service ecosystem with e-labeling. *Lecture Notes in Business Information Processing*, 299, 307-323, ISSN 18651348, doi:10.1007/978-3-319-65930-5_26

Ardito, L. (2019). Towards Industry 4.0: Mapping digital technologies for supply chain management-marketing integration. *Business Process Management Journal*, 25(2), 323-346, ISSN 14637154, doi:10.1108/BPMJ-04-2017-0088

Arribas V., José A., (2018). 3D technology in fashion: from concept to consumer, *Journal of Fashion Marketing and Management: An International Journal*, Vol. 22 Issue: 2, pp.240-251, <https://doi.org/10.1108/JFMM-10-2017-0114>

Ayers, J. B. (2006). *Handbook of supply chain management*. Auerbach publications.

Ayers, J.B. (2001). *Handbook of Supply Chain Management*. Boca Raton: The St. Lucie Press/APICS Series

Baranowski, K. (2018). Chemical supply chains go digital. *Chemical Engineering (United States)*, 125(2), 50-52, ISSN 00092460

Bechtsis, D. (2018). Intelligent Autonomous Vehicles in digital supply chains: A framework for integrating innovations towards sustainable value networks. *Journal of Cleaner Production*, 181, 60-71, ISSN 09596526, doi:10.1016/j.jclepro.2018.01.173

Beilin, I. (2019). Digital modeling of economic processes and supply chain management in the formation of cooperative relations in the petrochemical cluster of the region. *International Journal of Supply Chain Management*, 8(4), 532-537, ISSN 20513771

Bicocchi, N. (2019). Dynamic digital factories for agile supply chains: An architectural approach. *Journal of Industrial Information Integration*, 15, 111-121, ISSN 2452414X, doi:10.1016/j.jii.2019.02.001

Büyükoçkan, G. (2017). An extension of MOORA approach for group decision-making based on interval valued intuitionistic fuzzy numbers in digital supply chain. *IFSA-SCIS 2017 - Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems*, doi:10.1109/IFSA-SCIS.2017.8023358

Büyükoçkan, G. (2017). An extension of ARAS methodology based on interval valued intuitionistic fuzzy group decision making for digital supply chain. *IEEE International Conference on Fuzzy Systems*, ISSN 10987584, doi:10.1109/FUZZ-IEEE.2017.8015680

Büyükoçkan, G. (2018). Digital Supply Chain: Literature review and a proposed framework for future research. *Computers in Industry*, 97, 157-177, ISSN 01663615, doi:10.1016/j.compind.2018.02.010

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A. and N. Venkatraman (2013). "Digital business strategy: Toward a next generation of insights." *MIS Quarterly* 37 (2), 471-482.

Centobelli, P., Cerchione, R., & Ertz, M. (2020). Agile supply chain management: where did it come from and where will it go in the era of digital transformation?. *Industrial Marketing Management*, 90, 324-345.

Chanias, S. (2016). Understanding digital transformation strategy formation: Insights from Europe's automotive industry. *Pacific Asia Conference on Information Systems, PACIS 2016 - Proceedings*

Chen, I. J., & Paulraj, A. (2004). Towards a theory of supply chain management: the constructs and measurements. *Journal of operations management*, 22(2), 119-150.

Chopra, S. and Meindl, P., 2007. *Supply chain management: strategy, planning, and operation*. 3rd edition. Upper Saddle River: Pearson Prentice Hall.

Ciulli, F. (2019). Circularity Brokers: Digital Platform Organizations and Waste Recovery in Food Supply Chains. *Journal of Business Ethics*, ISSN 01674544, doi:10.1007/s10551-019-04160-5

Covaci, F. (2019). Modelling decision making in digital supply chains: insights from the petroleum industry. *Kybernetes*, ISSN 0368492X, doi:10.1108/K-10-2018-0565

Crawford, C. (2019). The smart supply chain: A digital revolution. *AATCC Review*, 19(3), 38-45, ISSN 15328813, doi:10.14504/ar.19.3.2

Das, A. (2019). Managing Disruptions and the Ripple Effect in Digital Supply Chains: Empirical Case Studies. *International*

- Series in Operations Research and Management Science, 276, 261-285, ISSN 08848289, doi:10.1007/978-3-030-14302-2_13
- Defraeye, T. (2019). Digital twins probe into food cooling and biochemical quality changes for reducing losses in refrigerated supply chains. *Resources, Conservation and Recycling*, 149, 778-794, ISSN 09213449, doi:10.1016/j.resconrec.2019.06.002
- Delina, R. (2016). Digital single market conceptual model for intelligent supply chains. *IDIMT 2016 - Information Technology, Society and Economy Strategic Cross-Influences - 24th Interdisciplinary Information Management Talks*, 269-278
- Djatna, T. (2015). An Analysis and Design of Responsive Supply Chain for Pineapple Multi Products SME Based on Digital Business Ecosystem (DBE). *Procedia Manufacturing*, 4, 155-162, ISSN 23519789, doi:10.1016/j.promfg.2015.11.026
- Dweekat, A.J., Hwang, G. and Park, J. (2017), "A supply chain performance measurement approach using the internet of things: toward more practical SCPMS", *Industrial Management & Data Systems*, Vol. 117 No. 2
- Dyatlov, S. (2018). The impact of network human capital on economic growth of supply chain in digital economy. *International Journal of Supply Chain Management*, 7(5), 877-885, ISSN 20513771
- Evtodieva, T. (2020). Business analytics of supply chains in the digital economy. *Advances in Intelligent Systems and Computing*, 908, 329-336, ISSN 21945357, doi:10.1007/978-3-030-11367-4_32
- Farahani, P. (2016). Digital supply chain management agenda for the automotive supplier industry. *Shaping the Digital Enterprise: Trends and Use Cases in Digital Innovation and Transformation*, 157-172, doi:10.1007/978-3-319-40967-2_8
- Faridi, M. R., & Malik, A. (2020). Digital transformation in supply chain, challenges and opportunities in SMEs: a case study of Al-Rumman Pharma. *Emerald Emerging Markets Case Studies*.
- Fedyk, A. (2016). How to Tell If Machine Learning Can Solve Your Business Problem. *Harvard Business Review*.
- Ferretti, M., & Schiavone, F. (2016). Internet of Things and business processes redesign in seaports: The case of Hamburg. *Business Process Management Journal*, 22(2), 271-284. <https://doi.org/10.1108/BPMJ-05-2015-0079>
- Fichman, R. G., Dos Santos, B. L. and Z. E. Zheng (2014). "Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum." *MIS Quarterly* 38 (2), 329-353.
- Gafurov, A. R., Skotarenko, O. V., Nikitin, Y. A., & Plotnikov, V. A. (2020, August). Digital transformation prospects for the offshore project supply chain in the Russian Arctic. In *IOP Conference Series: Earth and Environmental Science* (Vol. 554, No. 1, p. 012009). IOP Publishing.
- Gravili, G. (2018). The influence of the Digital Divide on Big Data generation within supply chain management. *International Journal of Logistics Management*, 29(2), 592-628, ISSN 09574093, doi:10.1108/IJLM-06-2017-0175
- Gustafsson, E. (2019). Digital product fitting in retail supply chains: maturity levels and potential outcomes. *Supply Chain Management*, 24(5), 574-589, ISSN 13598546, doi:10.1108/SCM-07-2018-0247
- Hanifan, G., Sharma, A., & Newberry, C. (2014). *The Digital Supply Network: A new paradigm for supply chain management*. Accenture. Retrieved May 8, 2018 from [https://www.accenture.com/t20150708T025455__w__/fr-fr_acnmedia/Accenture/Conversion-](https://www.accenture.com/t20150708T025455__w__/fr-fr_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Local/fr-fr/PDF_5/Accenture-Digital-Supply-Network-New-Standard-Modern-Supply-Chain-Management.pdf)
- Assets/DotCom/Documents/Local/fr-fr/PDF_5/Accenture-Digital-Supply-Network-New-Standard-Modern-Supply-Chain-Management.pdf
- Hartley, J. (2019). Tortoise, not the hare: Digital transformation of supply chain business processes. *Business Horizons*, ISSN 00076813, doi:10.1016/j.bushor.2019.07.006
- Henriette, Emily, Mondher Feki, and Imed Boughzala. "The Shape of Digital Transformation: A Systematic," n.d.
- Hess, T. (2016). Options for formulating a digital transformation strategy. *MIS Quarterly Executive*, 15(2), 123-139, ISSN 15401960
- Horlach, B., Drews, P., Schirmer, I. and T. Böhm (2017). "Increasing the Agility of IT Delivery: Five Types of Bimodal IT Organization." In: *Proceedings of the Hawaiian International Conference on System Sciences*. Hawaii: USA.
- Hu, M. (2015). Sharing supply chain data in the digital era. *MIT Sloan Management Review*, 57(1), 95-96, ISSN 15329194
- Huang, Y. (2017). Pricing and coordination with consideration of piracy for digital goods in supply chains. *Journal of Business Research*, 77, 30-40, ISSN 01482963, doi:10.1016/j.jbusres.2017.03.023
- Ivanov, D. (2019). Digital Supply Chain Twins: Managing the Ripple Effect, Resilience, and Disruption Risks by Data-Driven Optimization, Simulation, and Visibility. *International Series in Operations Research and Management Science*, 276, 309-332, ISSN 08848289, doi:10.1007/978-3-030-14302-2_15
- Ivanov, D. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research*, 57(3), 829-846, ISSN 00207543, doi:10.1080/00207543.2018.1488086
- Jacques, A. (2017). The digital supply chain: Seizing pharma's untapped opportunity. *Pharmaceutical Technology*, 2017(1), ISSN 15432521
- Junge, A. L., & Straube, F. (2020). Sustainable supply chains—digital transformation technologies' impact on the social and environmental dimension. *Procedia Manufacturing*, 43, 736-742.
- Kache, F. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations and Production Management*, 37(1), 10-36, ISSN 01443577, doi:10.1108/IJOPM-02-2015-0078
- Kasemsap, K. (2017). Mastering fashion supply chain management and new product development in the digital age. *Advanced Fashion Technology and Operations Management*, 65-91, doi:10.4018/978-1-5225-1865-5.ch003
- Kim, Y. (2018). Pricing of digital video supply chain: Free versus paid service on the direct distribution channel. *Sustainability (Switzerland)*, 11(1), ISSN 20711050
- Kodolova, I. (2019). Development of innovative activity and supply chain strategy of enterprises in the age of digital economy. *International Journal of Supply Chain Management*, 8(4), 525-531, ISSN 20513771
- Koshel, V. (2019). Generation specific patterns of communication based on the supply chain management in the digital society. *International Journal of Supply Chain Management*, 8(3), 449-459, ISSN 20513771
- Kosmol, T. (2019). You will never walk alone: Why we need a supply chain practice view on digital procurement. *Journal of Purchasing and Supply Management*, 25(4), ISSN 14784092, doi:10.1016/j.pursup.2019.100553
- Kubáč, Lukáš, and Oldřich Kodým. 2017. "The Impact of 3D Printing Technology on Supply Chain." Edited by Ondrej

- Stopka. MATEC Web of Conferences 134: 00027.
- Kumar, V., Amorim, M., Bhattacharya, A. and Garza, J. (2016) Managing reverse exchanges in service supply chains. *Supply Chain Management: An International Journal*, 21 (2). pp. 157-165. ISSN1359-854
- Kurz, D. (2018). Leading the digital supply chain. *The Internet of People, Things and Services: Workplace Transformations*, 242-257, doi:10.4324/9781315182407
- Lambert, D. M., Stock, J. R., & Ellram, L. M. (1998). *Fundamentals of logistics management*. McGraw-Hill/Irwin.
- Lavikka, R. (2017). Transforming a supply chain towards a digital business ecosystem. *IFIP Advances in Information and Communication Technology*, 514, 295-301, ISSN 18684238, doi:10.1007/978-3-319-66926-7_34
- Li, B., & Li, Y. (2017). Internet of Things drives supply chain innovation: a research framework. *International Journal of Organizational Innovation*, 9(3), 71–92.
- Mann, J. (2015). *The Internet of Things: Opportunities and applications across industries*. International Institute for Analytics, Enterprise Research Service.
- Manyika, J., Chui, M., Bisson, P., Woetzel, S., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The Internet of Things: Mapping the Value Beyond the Hype*.
- Mapokgole, J. (2018). Ways on how digital technology can streamline supply chain management of today. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 2018, 336, ISSN 21698767
- Massimino, B. (2018). On the Inattention to Digital Confidentiality in Operations and Supply Chain Research. *Production and Operations Management*, 27(8), 1492-1515, ISSN 10591478, doi:10.1111/poms.12879
- Masvosvere, D. (2015). A conceptual model for digital forensic readiness in e-supply chains. *European Conference on Information Warfare and Security, ECCWS, 2015*, 413-422, ISSN 20488602
- Masvosvere, D. (2015). A model for the design of next generation e-supply chain digital forensic readiness tools. *2015 Information Security for South Africa - Proceedings of the ISSA 2015 Conference*, doi:10.1109/ISSA.2015.7335075
- Masvosvere, D. (2016). Using a standard approach to the design of next generation e-supply chain digital forensic readiness systems. *SAIEE Africa Research Journal*, 107(2), 104-119, ISSN 00382221
- Matt, C. (2015). Digital Transformation Strategies. *Business and Information Systems Engineering*, 57(5), 339-343, ISSN 18670202, doi:10.1007/s12599-015-0401-5
- Meier, C. (2016). Digital supply chain management. *Digital Enterprise Transformation: A Business-Driven Approach to Leveraging Innovative IT*, 231-262, doi:10.4324/9781315577166
- Menon, S. (2018). An Overview of Smart Manufacturing for Competitive and Digital Global Supply Chains. *2018 IEEE International Conference on Technology Management, Operations and Decisions, ICTMOD 2018*, 178-183, doi:10.1109/ITMC.2018.8691224
- Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D., Zacharia, Z.G., (2001), *Defining Supply Chain Management*, *Journal of Business Logistics*, vol. 22, no 2, pp. 1-25
- Merlino, M. *Logistica Management*, Editrice Temi, Milano, Italy, Various editorials on these subjects 2007-2016
- MaryAnne M. Gobble (2018) *Digitalization, Digitization, and Innovation*, *Research-Technology Management*
- Mikhaylyk, M. (2019). Empirical markers in the concept of digital logistics of multichannel supply chains. *E3S Web of Conferences*, 91, ISSN 22671242, doi:10.1051/e3sconf/20199108056
- Mola, L. (2017). Who knows what? Reconfiguring the governance and the capabilities of the supply chain between physical and digital processes in the fashion industry. *Production Planning and Control*, 28(16), 1284-1297, ISSN 09537287, doi:10.1080/09537287.2017.1375147
- Morrissey, M. (2017). Online supply chain education in a digital world. *On the Line: Business Education in the Digital Age*, 309-319, doi:10.1007/978-3-319-62776-2_23
- Neumann, H. (2018). Efficiency and clarity in the delivery process: Digital container management ensures a transparent supply chain at Bell. *Fleischwirtschaft*, 98(10), 52-54, ISSN 0015363X
- Okoli, Chitu, and Kira Schabram. (2010). *A Guide to Conducting a Systematic Literature Review of Information Systems Research*
- Parker, D. (2019). A randomised controlled trial and cost-consequence analysis of traditional and digital foot orthoses supply chains in a National Health Service setting: Application to feet at risk of diabetic plantar ulceration. *Journal of Foot and Ankle Research*, 12(1), ISSN 17571146, doi: 10.1186/s13047-018-0311-0
- Patel, K., McCarthy, M.P. (2000). *Digital Transformation: The Essentials of E-Business Leadership 1st*, McGraw-Hill Professional 2000
- Perla, L. (2018). The 'supply chain' of teachers' digital skills training. *The TPACK traceability in the teachers' trainers. ACM International Conference Proceeding Series*, 604-612, doi:10.1145/3284179.3284284
- Perrin, A. (2015). *Social media usage*. Pew research center, 125, 52-68.
- Plant, L. (2017). Implications of open source blockchain for increasing efficiency and transparency of the digital content supply chain in the Australian telecommunications and media industry. *Australian Journal of Telecommunications and the Digital Economy*, 5(3), 15-29, ISSN 22031693, doi:10.18080/ajtde.v5n3.113
- Preindl, R., Nikolopoulos, K., & Litsiou, K. (2020, January). Transformation strategies for the supply chain: the impact of industry 4.0 and digital transformation. In *Supply Chain Forum: An International Journal (Vol. 21, No. 1, pp. 26-34)*. Taylor & Francis.
- Ready, P. J., Gunasekaran, A., & Spalanzani, A. (2015). Bottom-up approach based on Internet of Things for order fulfillment in a collaborative warehousing environment. *International Journal of Production Economics*, (159), 29–40. <https://doi.org/10.1016/j.ijpe.2014.02.017>
- Reis, João, Marlene Amorim, Nuno Melão, and Patrícia Matos (2018). “Digital Transformation: A Literature Review and Guidelines for Future Research.” In *Trends and Advances in Information Systems and Technologies*, edited by Álvaro Rocha, Hojjat Adeli, Luís Paulo Reis, and Sandra Costanzo, 745:411–21. Cham: Springer International Publishing.
- Sabri, Y. (2018). Supply Chain Digital Innovation in Complex Capital Projects. *2018 IEEE International Conference on Technology Management, Operations and Decisions, ICTMOD 2018*, 262-267, doi:10.1109/ITMC.2018.8691284
- Schmidt, B., Wang, L. 2016. "Cloud-enhanced predictive maintenance", *International Journal of Advanced Manufacturing Technology*, <http://dx.doi.org/10.1007/s00170-016-8983-8>
- Scholz, J. (2018). Digital Technologies for Forest Supply Chain Optimization: Existing Solutions and Future Trends. *Environmental Management*, 62(6), 1108-1133, ISSN

0364152X, doi:10.1007/s00267-018-1095-5

Schrauf, S. (2016). Industry 4.0: How digitization makes the supply chain more efficient, agile, and customer-focused.

Schwertner, K. (2017). Digital Transformation of Business. *Trakia Journal of Science* 15, no. Suppl.1: 388–93. <https://doi.org/10.15547/tjs.2017.s.01.065>.

Scrimshaw, J. (2016). Digital textile printing in the 'virtual' supply chain. *Digital Textile*, 2016(5), 4, ISSN 17421128

Scuotto, V. (2017). A multiple buyer–supplier relationship in the context of SMEs' digital supply chain management. *Production Planning and Control*, 28(16), 1378-1388, ISSN 09537287, doi:10.1080/09537287.2017.1375149

Stolterman E., Fors A.C. (2004) Information Technology and the Good Life. In: Kaplan B., Truex D.P., Wastell D., Wood-Harper A.T., DeGross J.I. (eds) *Information Systems Research*. IFIP International Federation for Information Processing, vol 143. Springer, Boston, MA

Siciliano, B., & Khatib, O. (Eds.). (2016). *Springer handbook of robotics*. Springer.

Sundaram, R., Sharma, R., & Shakya, A. (2020). Digital transformation of business models: a systematic review of impact on revenue and supply chain. *International Journal of Management*, 11(05).

Susilo, F. (2018). Digital supply chain development in blockchain technology using Rijndael algorithm 256. *IOP Conference Series: Materials Science and Engineering*, 453(1), ISSN 17578981, doi:10.1088/1757-899X/453/1/012075

Tao, R. (2018). Smart rural tourism supply chain cooperative technology based on digital footprint. *IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association*, 30(7), 710-717, ISSN 03795462

Trokanas, N. (2017). Towards an Ontological Backbone for Pharmaceutical Digital Supply Chains. *Computer Aided Chemical Engineering*, 40, 2329-2334, ISSN 15707946, doi:10.1016/B978-0-444-63965-3.50390-1

Tsolakis, N. (2019). Intelligent autonomous vehicles in digital supply chains: From conceptualisation, to simulation modelling, to real-world operations. *Business Process Management Journal*, 25(3), 414-437, ISSN 14637154, doi:10.1108/BPMJ-11-2017-0330

Tziantopoulos, K. (2019). Supply chain reconfiguration opportunities arising from additive manufacturing technologies in the digital era. *Production Planning and Control*, 30(7), 510-521, ISSN 09537287, doi:10.1080/09537287.2018.1540052

Vazquez-Martinez, G. (2018). CloudChain: A novel distribution model for digital products based on supply chain principles. *International Journal of Information Management*, 39, 90-103, ISSN 02684012, doi:10.1016/j.ijinfomgt.2017.12.006

Wang, B. (2018). Exploration of Digital Competency Requirements within the Fashion Supply Chain with an Anticipation of Industry 4.0. *International Journal of Fashion Design, Technology and Education*, 11(3), 333-342, ISSN 17543266, doi:10.1080/17543266.2018.1448459

Wei, F. (2019). Applications of digital technologies in sustainable logistics and supply chain management. *Innovative Logistics Services and Sustainable Lifestyles: Interdependencies, Transformation Strategies and Decision Making*, 235-263, doi:10.1007/978-3-319-98467-4_11

Wei, S. (2019). The Effects of Justice and Top Management Beliefs and Participation: An Exploratory Study in the Context of Digital Supply Chain Management. *Journal of Business Ethics*, ISSN 01674544, doi:10.1007/s10551-018-04100-9

Westerman, G. (2015). Revamping your business through digital transformation. *MIT Sloan Management Review*, 56(3), 2-5, ISSN 15329194

Westerman, G. (2016). Why digital transformation needs a heart. *MIT Sloan Management Review*, 58(1), ISSN 15329194

Wong, L. (2019). Time to seize the digital evolution: Adoption of blockchain in operations and supply chain management among Malaysian SMEs. *International Journal of Information Management*, ISSN 02684012, doi:10.1016/j.ijinfomgt.2019.08.005

Yegina, N. (2019). Applying consumer behavior theory integrated with supply chain strategy in the context of the digital transformation of the economy. *International Journal of Supply Chain Management*, 8(3), 341-347, ISSN 20513771

Zhang, L. (2018). Protection in DRM and pricing strategies in a two-echelon digital product supply chain. *International Journal of Systems Science: Operations and Logistics*, ISSN 23302674, doi:10.1080/23302674.2018.1471174

Segmentation et caractérisation de la demande des magasins de détail. Cas des produits cosmétiques

MOUHCINE LAAROUSHI¹, JULIO CESAR MONTECINOS², LOUBNA BENABBOU³, MUSTAPHA OUHIMMOU⁴,

^{1,2,4}Département du génie de la production automatisée, ÉTS Montréal
Québec, Canada

¹Mouhcine.laaroussi.1@ens.etsmtl.ca

²Julio.Montecinos@etsmtl.ca

⁴Mustapha.Ouhimmou@etsmtl.ca

³Département des sciences de la gestion, UQAR, Campus de Levis
Québec, Canada

³loubna_benabbou@uqar.ca

Résumé – Les chaînes d’approvisionnement des produits cosmétiques sont parmi les plus complexes dans l’industrie d’aujourd’hui. Ceci revient à la demande hétérogène, à la diversité des produits, et à une chaîne de distribution très compliquée basée sur un très grand nombre de détaillants, où chacun possède sa propre stratégie. De ce fait, les décideurs se trouvent incapables de comprendre la logique de leur chaîne d’approvisionnement. Et par conséquent, d’établir leur portefeuille de décisions devant une demande et une concurrence qui ne cessent d’augmenter. En conséquence, il est indispensable de développer de nouvelles approches pour supporter la prise de décision relative aux chaînes d’approvisionnement du cosmétique. Notre approche a pour but de faire une segmentation et une caractérisation de la demande des magasins de détail pour visualiser la complexité du problème et permettre une bonne et unique gestion pour chaque groupe. La segmentation sera faite en se basant sur des techniques d’exploration de données, notamment les techniques de regroupement, suivi d’une caractérisation des groupes obtenus afin d’aider les décideurs à bien comprendre et prioriser leurs groupes. L’étude de cas est portée sur une chaîne d’approvisionnement d’une grande compagnie du cosmétique au Canada.

Abstract – Cosmetic supply chains are among the most complex in the industry today. This is due to the heterogeneous demand, the diversity of products, and a very complicated distribution chain based on a very large number of retailers, each has its own strategy. As a result, decision makers find themselves unable to understand the logic of their supply chain, and therefore unable to establish their portfolio of decisions in the face of ever-increasing demand and competition. As a result, they are forced to develop new approaches to support cosmetic supply chain decision-making. Our approach aims to segment retail stores to reduce the complexity of the problem and allow the management of each group alone. The segmentation will be done based on data mining techniques, including clustering techniques followed by a characterization of the groups obtained to help decision makers to understand and prioritize their groups. The case study is based on a supply chain of a major Canadian company of cosmetics.

Mots-clés – Chaîne d’approvisionnement ; Segmentation ; Magasins de détail ; Produits cosmétiques.

Keywords – Supply Chain; Clustering; Retailer stores; Cosmetics products.

1 INTRODUCTION

De nos jours, l’industrie des produits cosmétiques est en perpétuelle évolution. Toutes les personnes indépendamment de leur classe sociale, de leur genre ou de leur race achètent des produits de soin de santé et de maquillage. De ce fait, les compagnies du cosmétique essaient de s’adapter à cette diversité de consommateurs sur le plan de la tarification, de variété et de marques. Par conséquent, la mission des décideurs de la chaîne d’approvisionnement devient de plus en plus complexe.

La demande dans les chaînes d’approvisionnement des produits cosmétiques est parmi les plus complexes dans l’industrie. Un premier défi consiste à la périssabilité très rapide des produits proposés : un produit en retard est un produit qui n’est plus à la mode. En revanche, les décideurs de la chaîne d’approvisionnement des produits cosmétiques se trouvent face à des centaines de produits à gérer. Ceci revient à

la variété des consommateurs dans le marché, soit en ce qui concerne : le genre, la classe sociale, la race, la nature de peau, la religion [Helmi et al., 2020] et bien d’autres facteurs. Par ailleurs, la distribution des produits cosmétiques se fait soit en détail par produit, ou avec des promotions sous forme de plusieurs collections de produits différents, qui sont destinés à un grand nombre de détaillants distribués sur un large horizon géographique. Alors que chacun d’entre eux demande ses propres promotions avec ses propres collections, qui changent assez fréquemment. Ainsi, les décideurs sont dans l’obligation de gérer un large nombre de combinaisons dans leurs stocks, et par conséquent de gérer un nombre très élevé de « Stock Keeping Units ».

Tous ces facteurs réunis empêchent la compréhension de la logique derrière la demande de ces chaînes d’approvements. Surtout que les détaillants proposent

leurs propres promotions avec un caractère imprédictible et très dynamique.

La prise de conscience de l'augmentation de la demande des produits cosmétiques accompagnée de l'augmentation de la concurrence, ainsi que la complexité d'une telle chaîne d'approvisionnement, a donné intérêt à développer de nouvelles approches afin d'aider à la prise de décision dans les chaînes d'approvisionnements complexes telle que celles des produits cosmétiques.

Le but de ce papier est d'utiliser les méthodes d'exploration de données dans la prise de décision relative aux chaînes d'approvisionnements. L'objectif sera d'utiliser des techniques de regroupement sur les magasins de vente des produits cosmétiques afin de permettre une bonne compréhension de la demande et une hiérarchisation des magasins en caractérisant chaque groupe en fonction des attributs. L'étude de cas va être faite sur des données provenant d'un leader des produits cosmétiques au Canada.

Le reste de ce papier est organisé comme suit : la deuxième section va être consacrée à la définition du problème et des motivations derrière la réalisation de ce travail. La troisième section présente une revue de la littérature des différents algorithmes utilisés et des techniques d'évaluation. La quatrième section définit les étapes de notre approche, suivie d'une étude de cas réelle présentée dans la cinquième section. La dernière section résume le papier et présente nos travaux futurs.

2 DEFINITION DU PROBLEME :

Afin de répondre aux exigences de la demande hétérogène, de la diversité des promotions, et du besoin diversifié des consommateurs, les chaînes d'approvisionnement des produits cosmétiques doivent être flexibles tout en étant efficaces.

En plus de la diversité des produits et des promotions, la distribution des produits cosmétiques se fait à travers plusieurs échelons de plusieurs chaînes logistiques. Chacune commence par un grand détaillant qui est le premier client direct du producteur. Chaque grand détaillant favorise sa demande des produits cosmétiques en fonction de l'espace disponible dans ses magasins, avec très peu de marge de manœuvre pour le producteur. Par conséquent, chaque magasin devient objet de concurrence avec les autres compagnies, plus la concurrence existante entre les différents magasins d'une même compagnie.

Ce niveau très élevé de concurrence perturbe la production et donne naissance à un autre problème au niveau de la logistique en générant des stocks excédentaires et des retours au producteur. En plus, chaque détaillant utilise ses propres stratégies pour retourner ces produits non vendus, ou ceux retournés par les clients. Ce qui génère un grand décalage entre la date d'envoi et de retour des produits. Par conséquent, les décideurs trouvent des difficultés dans la gestion et la compréhension de la logique de ces retours vu le grand nombre des magasins. L'idée la plus intuitive serait de retirer certains produits qui ne se vendent pas dans les grands magasins. Une telle décision réduira le nombre de produits présentés par la compagnie dans le marché, ce qui peut nuire à son image à cause de la concurrence très sévère. De plus, les produits cosmétiques présentent des catégories de prix : du moins cher à des produits extrêmement chers. Alors, les décideurs sont

obligés de concevoir des stratégies de distribution plus flexible et robuste.

Dans cet article, notre proposition est basée sur une segmentation des magasins de détail en fonction des produits envoyés, vendus et retournés, suivie d'une caractérisation de chaque groupe ou segment. Cette solution permettra aux décideurs : (1) de retrouver une compréhension du problème de marchandage pour décider la quantité à produire et minimiser les produits retournés. (2) De comprendre et construire leur portefeuille de décision pour gérer la distribution et les produits de retours pour chaque groupe. (3) de hiérarchiser leurs groupes de magasins afin de définir leurs magasins de luxe.

Le regroupement va être fait en utilisant des outils d'exploration de données, en raison du grand volume de données et de dimensions qu'on ne peut pas gérer avec des analyses classiques.

3 REVUE DE LITTERATURE

Nous vivons actuellement dans l'ère des données, où un très grand volume de données est collecté et stocké chaque jour [Han, et al., 2012]. L'exploration de données est un processus qui transforme cette vaste collection de données à des informations potentiellement utiles [Han, et al., 2012]. En utilisant plusieurs techniques telles que la caractérisation et la discrimination des données, l'exploration des patrons fréquents, analyse d'association et de corrélation, classification, régression, *clustering* ou regroupement et analyse des valeurs aberrantes [Han, et al., 2012]. Le regroupement est l'une des techniques fondamentales dans l'exploration des données. Il sert à diviser les données en plusieurs groupes contenant des objets similaires [Berkhin, 2006]. Chacun de ces groupes est appelé 'cluster', et qui ont des caractéristiques intragroupes (les éléments des mêmes groupes) similaires, et intergroupes (de groupes différents) différentes. Le regroupement couvre un large spectre d'applications qu'on peut les diviser en deux objectifs principaux : la préparation de données et l'analyse exploratoire de données [Guyon et al., 2009]. Notamment dans la gestion des chaînes d'approvisionnement, le regroupement a été largement utilisé : pour l'évaluation de la performance des collaborations des chaînes d'approvisionnements [Derrouiche et Leksakul, 2011], le regroupement des consommateurs, des détaillants, des distributeurs, des manufacturiers et des fournisseurs [Danish et al., 2007].

Dans la littérature, les techniques de regroupement utilisent des caractéristiques des consommateurs, du revenu, du volume de la demande, des données géographiques et d'autres attributs qui caractérisent la fonction de vente des compagnies [Murray et al., 2015], [Murray et al., 2017], [A. A. Khalmarov et al., 2020]. Or, dans cet article, la segmentation est faite sur les volumes vendus, envoyés et retournés. L'ajout des volumes retournés va permettre de créer des groupes avec des caractéristiques de retour de produits différentes. Ce qui permettra aux décideurs d'adopter des politiques de retour bien ciblées pour chaque groupe.

Dans cette partie, nous présenterons une petite revue de littérature des différents algorithmes utilisés dans l'étude, ainsi que les techniques d'évaluation pour le choix du nombre optimal de groupes et du meilleur modèle à utiliser.

3.1 Les techniques de regroupement :

Selon la méthode adoptée, le regroupement peut être classé en quatre approches principales : (1) méthode de partitionnement, où initialement le nombre de groupes voulu est fixé, puis ils sont construits itérativement en optimisant un critère déterminé. (2) Méthodes hiérarchiques qui cherchent à construire une hiérarchie de groupes à partir d'une division d'un seul groupe contenant tous les points des données, ou à partir d'une fusion de plusieurs groupes initiaux où chaque groupe contient un seul point des données. Cette dernière approche est dite : regroupement hiérarchique ascendant qu'on va utiliser dans ce papier. (3) Méthodes basées sur la densité qui regroupent les points selon leurs densités et voisinages, ou à partir d'une fonction de densité. (4) des méthodes basées sur les grilles qui commencent par un ordonnancement de tous les points des données sur une grille de cellules. Par la suite, un regroupement est fait sur ces cellules en fonction des distances ou des poids de ces derniers. S'ajoutent à celles-ci d'autres modèles de regroupement probabilistes, des données de grandes dimensions, regroupement graphique, et des regroupements avec des contraintes. Dans cette partie, une brève revue de littérature a été présentée, les lecteurs intéressés peuvent consulter les travaux de Berkhin [Berkhin, 2006], et Han, Kamber et Pei [Han, et al., 2012]

Dans ce papier, nous avons utilisé cinq algorithmes principaux. (1) regroupement hiérarchique ascendant [Johnson, 1967]. (2) K-means (3) Partitioning-Around-Medoids (PAM) (4) Gaussian-Mixture (GM) et (5) fuzzy-c-means. K-means [Hartigan et al., 1997] est une technique de regroupement basée sur l'approche de partitionnement. Elle commence par un nombre de points initiaux considérés comme les centres des groupes. Puis elle assigne chaque objet de l'ensemble de données au groupe dont le centre est le plus proche. Par la suite, elle recalcule les centres des groupes, en utilisant la moyenne, qui vont être changés itérativement en optimisant la somme des carrés des distances intragroupes jusqu'à ce qu'ils restent inchangeables. L'inconvénient de cette méthode réside au fait de l'utilisation de la moyenne fictive pour définir les centres des groupes, ce qui les rend très sensibles aux valeurs aberrantes. Partitioning-Around-Medoids (PAM) [Kaufman, 1990] est une version plus robuste du k-means, vu qu'elle est basée sur le même principe sauf qu'elle utilise un objet représentatif (point réel de la base de données) pour chaque groupe au lieu de la moyenne. Et par conséquent, il évite l'attraction des centres vers les données aberrantes. Gaussian Mixture [Dempster et al., 1977] est un modèle probabiliste qui peut être considéré comme une généralisation du k-means, où il prend en considération aussi la forme de la covariance des données, ainsi qu'il utilise la maximisation de la fonction d'expectation comme critère à optimiser. Généralement, le regroupement peut être divisé en deux autres grandes parties : un regroupement solide où on retourne directement la classe de chaque objet. Et un regroupement flou qui consiste à calculer des coefficients d'appartenance des objets à chaque groupe. Fuzzy-c-means [Bezdek et al., 1984] : est un modèle de regroupement flou, basé sur le même concept que celui du k-means à l'exception qu'il retourne des coefficients d'appartenance. L'avantage de cet algorithme c'est qu'il permet de construire des frontières floues entre les groupes.

3.2 Métriques d'évaluation de la qualité du regroupement

Dans l'analyse du regroupement, il n'existe pas de modèle universel, chaque modèle excelle différemment en fonction des

données traitées. Par conséquent, l'une des questions les plus importantes à laquelle on doit répondre lors du regroupement est l'évaluation des groupes obtenus afin de trouver le modèle qui convient le plus aux données [Halkidi, 2001]. En général, il existe trois types d'indices d'évaluation des modèles de regroupement : (1) des indices internes : ils évaluent la qualité du regroupement en utilisant les mêmes données et variables dont on a fait le regroupement sans aucune information supplémentaire ; (2) des indices externes : ils utilisent une information supplémentaire sous forme de variable cible qui reflète notre intuition sur la structure des données ; (3) des indices relatifs : utilisés pour l'optimisation des paramètres d'un algorithme. On évalue les résultats d'un même modèle, mais avec différents paramètres. Dans cet article, on se contentera d'une évaluation interne puisque nous n'avons pas d'information supplémentaire bien claire pour établir une variable cible pour la validation externe.

La validation interne couvre l'évaluation de quatre critères très importants. (1) Mesure de compacité : Elle évalue la proximité des objets de même groupe ; (2) Mesure de séparation : Elle évalue la distance d'un groupe par rapport aux autres groupes qui doivent être bien espacés ; (3) Mesure de stabilité : Elle évalue la force du modèle vis-à-vis aux changements de la base de données. En d'autres termes, elle reflète la cohérence des résultats obtenus face aux perturbations des données, ou à un ré-échantillonnage à partir des données d'origines ; (4) Mesure de connectivité : Elle reflète le fait qu'une même densité de point dans la base de données appartient au même groupe.

L'indice de Davies-Bouldin, l'indice de Calinski-Harabasz, Silhouette score, et le gap statistique sont les mesures d'évaluation interne utilisées dans cette étude. L'indice de Davies-Bouldin est défini comme le ratio de la distance moyenne intragroupe sur la distance entre les centres des groupes. Il varie entre $[0, +\infty[$ et on vise toujours à le minimiser. L'indice de Calinski-Harabasz est défini comme le ratio de la dispersion intergroupe sur celle intragroupe. Il varie entre $[0, +\infty[$ et on cherche toujours à le maximiser. La silhouette score mesure la similarité entre un point de données et les points de son propre groupe par rapport à ceux du groupe voisin. Il prend des valeurs dans l'intervalle $[-1, 1]$ et doit être maximisé. Ces trois indices sont des indices combinés qu'évaluent la compacité et la séparation des groupes en même temps. [Han, et al., 2012]

Gap statistique est une méthode d'évaluation proposée par [Tibshirani et al, 2001]. Elle a pour but de standardiser le graphe du logarithme de la somme des carrés des distances entre les points d'un même groupe (on le note $\log(W_k)$). Ceci en le comparant avec une référence de distribution nulle, c'est-à-dire avec des échantillons aléatoires des données d'origine. Elle est calculée en utilisant la formule suivante :

$$\text{Gap statistique} = E(\log W_k) - \log W_k$$

Avec : $E(\log W_k)$ est l'espérance des $\log(W_k)$ calculés à partir des regroupements des échantillons. Et le $\log W_k$ est calculé à travers le regroupement des données d'origine. Cette méthode a tendance de performer mieux que les autres indices. Elle reflète la qualité des groupes ainsi que la stabilité du modèle utilisé.

Les différents indices présentés dans cette partie vont être utilisés : pour l'évaluation des modèles afin de choisir le

meilleur, et pour le choix du nombre de groupe k à utiliser. Nous allons ajouter aussi une autre approche de détermination de k qu'on appelle 'Elbow Method' [Sugar, 1998]. Initialement, le modèle choisi est exécuté en utilisant plusieurs valeurs de k . Puis on trace la somme des carrés des distances intragroupes en fonction de k . La valeur optimale sera le point qui donne naissance à un coude dans le graphe.

3.3 Caractérisation des groupes :

La dernière étape dans notre approche sera de caractériser chaque groupe en utilisant des valeurs typiques des variables d'entrées. Cette caractérisation aidera les décideurs à comprendre et bien positionner chaque groupe afin de savoir sa nature et ses spécificités par la suite. Une étude descriptive est une méthode utilisée pour la caractérisation des groupes en utilisant des statistiques telles que : les moyennes, les écarts types, et les histogrammes. Or, ceci ne donne aucune vision réelle sur les objets des groupes à raison de l'utilisation des valeurs statistiques fictives. Dans ce sens, l'utilisation des arbres de classification sera d'une grande utilité. En général, la classification est une technique d'apprentissage qui associe à chaque classe un ensemble de points de données. Il existe plusieurs modèles de classification dans la littérature [Han, et al., 2012], dont les arbres de classification qui nous intéressent. C'est une structure hiérarchique de plusieurs niveaux divisés en trois parties essentielles : un nœud racine, des nœuds internes et les nœuds terminaux qu'on appelle feuilles. La classification est définie en attribuant à chaque feuille une classe. Pour cela, l'algorithme construit ce qu'on appelle une branche, on continue à diviser les nœuds en utilisant des conditions posées sur les attributs en partant du nœud racine jusqu'à l'obtention d'un nœud homogène qui va être défini comme feuille. On s'intéresse aux conditions des différentes branches qui construisent l'arbre et qui donnent sa force d'interprétation. L'un des algorithmes les plus populaires pour construire des arbres de classification est l'algorithme des arbres de classification et de régression (CART) [Breiman, 1984]. Il construit des arbres binaires : il identifie récursivement les attributs les plus significatifs et divise la population en utilisant des critères de remorquage.

4 APPROCHE ET METHODOLOGIE

Dans cette partie, nous présentons notre approche pour le regroupement et la caractérisation de la demande des détaillants. L'approche proposée est initiée par un stade descriptif des données, suivi par la préparation de la base de données pour la modélisation, ensuite l'analyse de regroupement, puis le choix du meilleur modèle de regroupement qu'on va caractériser à la fin (Figure 1).

4.1 Analyse descriptive et graphique des données

La première étape de notre approche consiste à faire une analyse descriptive des données en utilisant des valeurs statistiques, telles que les moyennes, les écarts types, les quantiles, le maximum et le minimum sur toutes les variables afin de les caractériser et voir comment elles sont distribuées.

Puis nous analyserons graphiquement nos données en utilisant des tracés des variables, des boîtes à moustache pour voir la distribution des variables et des valeurs aberrantes, des graphes en paire afin de comprendre l'évolution des variables l'un par rapport aux autres et leurs corrélations. Enfin, on peut faire des matrices de corrélation afin de s'assurer des corrélations qui existent entre les variables.

4.2 Préparation des données

Généralement, la préparation des données exige trois traitements essentiels : (1) traitement des valeurs manquantes : en fonction du volume de données et du problème, on peut supprimer les lignes contenant des données manquantes ou même des variables avec beaucoup de points manquants. Cette méthode peut générer un manque de données, alors le plus intéressant sera de les imputer avec des méthodes d'imputation uni-variées ou multi-variées, et d'une façon itérative ou pas [Pratama et al., 2016]. (2) Traitement des données aberrantes : en fonction du problème, les données aberrantes peuvent être supprimées ou remplacées en utilisant des méthodes d'imputation déjà mentionnées. (3) Transformation des données : l'utilisation des données telle qu'elles sont peut déséquilibrer les variables d'entrées pour les méthodes d'exploration des données. Alors, généralement on peut standardiser les valeurs des données d'entrées en même intervalle de valeurs.

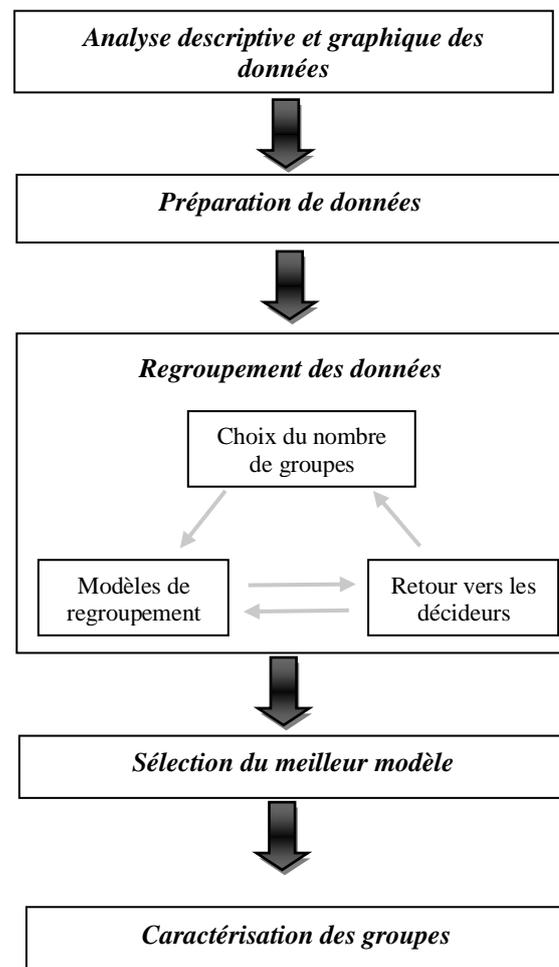


Figure 1. Approche de regroupement proposée

4.3 Regroupement des données

Une question très importante à répondre avant d'entamer le regroupement est de savoir le nombre optimal de groupes k à utiliser. Pour ce faire, nous allons exécuter le modèle de regroupement choisi en utilisant différentes valeurs de k . Puis, pour chaque valeur de k , on calcule les métriques d'évaluation. Par la suite, ses différentes métriques sont tracées en fonction de k , et le k qui les optimise sera choisi. Généralement, dans cette étape on ne doit pas se contenter d'une seule valeur.

D'une part, dans la plupart du temps, il est difficile de détecter la valeur optimale de k . D'autre part, la détermination du nombre de groupes à utiliser ne s'arrête pas à une valeur déterminée d'une façon théorique, mais l'usage final et l'intuition des décideurs doivent être prises en compte [Guyon et al., 2009].

Après avoir défini les nombres de groupes à utiliser, il est temps d'exécuter les modèles de regroupement choisis et les évaluer en utilisant les métriques d'évaluation. Dans notre cas, nous avons choisi : un regroupement hiérarchique ascendant, k -means, PAM, Gaussian Mixture et le fuzzy- c -means.

Par la suite, les résultats doivent être remis en cause avec les décideurs. Les décideurs évaluent les résultats avec un œil d'expert pour voir, d'un côté, la cohérence des regroupements, c'est-à-dire est ce que les magasins de détail d'un même groupe ont réellement des caractéristiques communes. Ceci est dû aux résultats différents qu'on peut obtenir en fonction du critère et de la méthode utilisée par le modèle. D'autre côté, le nombre de groupes est-il satisfaisant. Si le nombre est trop petit, la complexité du problème ne s'atténuera pas, ce qui est notre but initial. Cependant, s'il est très grand, les décideurs ne voient aucun intérêt à diviser le problème au nombre très élevé de groupes, vu qu'ils doivent être gérés chacun à part entière.

4.4 Sélection du meilleur modèle

À ce stade, nous aurons deux composantes très importantes et suffisantes pour la prise de décision sur le choix du modèle à utiliser : les métriques d'évaluation qui contrôlent la partie théorique des modèles, ainsi que les retours et les intuitions des décideurs de la chaîne d'approvisionnement qui contrôlent la partie pratique de l'utilisation des résultats. Donc, la combinaison de ses deux composantes donne une vision éclairée sur le choix du meilleur modèle.

4.5 Caractérisation des groupes

Les modèles de regroupement retournent des groupes de magasins de détail classés en fonction du volume de vente, d'envoyé et du retourné pour chaque produit. Cette dimension rend la compréhension des groupes très difficile. Par conséquent, une caractérisation des groupes en fonction des variables d'entrées va être très utile par la suite. Pour cela, nous allons transformer le problème en un problème de classification : on définit nos classes à l'aide des résultats de regroupement. Puis nous utiliserons des arbres de classification afin de retourner toutes les règles de classification, qui seront données sous forme de conditions sur la quantité des produits vendus, envoyés et retournés qui caractérisent chaque groupe.

5 ETUDE DE CAS DE LA DEMANDE DES MAGASINS DE DETAIL DES PRODUITS COSMETIQUES

L'étude de cas est portée sur la chaîne d'approvisionnement d'une compagnie du cosmétique Canadienne. La chaîne d'approvisionnement de cette compagnie est très complexe. Elle compte plus de 3000 magasins de détaillants à Montréal pour l'an 2019. La base de données, que nous avons reçu, contient les données de ces magasins concernant les produits envoyés, vendus, et retournés des quatre marques qu'ils possèdent. S'ajoute à ceci une autre variable qu'on va utiliser par la suite pour le choix du modèle final : pour chaque produit, les agents financiers de la compagnie notent les magasins de détail avec une échelle discriminatoire en fonction

de leurs quantités vendues. Nous appellerons cette variable 'Note' par la suite.

5.1 Analyse descriptive et graphique des données

Dans un premier temps, nous avons effectué une analyse descriptive des données. La remarque la plus importante à faire c'est qu'il y a des valeurs négatives. Ceci est dû au problème des stocks excédentaires de la compagnie. D'une part, les magasins de détail ont tendance à déplacer les produits non vendus vers d'autres magasins de vente, et donc ils sont obligés de saisir leurs produits avec des valeurs négatives pour compenser. D'autre part, la somme des produits vendus et retournés est différente de celle envoyée dans la plupart des cas. Ceci est expliqué par le décalage du retour des produits. Ensuite, une analyse graphique est utilisée pour voir les tendances, la distribution, et les évolutions des produits, les uns par rapport aux autres. Dans cette partie, nous avons constaté une seule corrélation forte qui existe entre les produits envoyés et vendus.

5.2 Préparation des données

Dans notre cas, nous avons dû faire avec des valeurs manquantes. Il existe des magasins où toutes les valeurs sont nulles. Nous avons procédé à une suppression de ses magasins vu que l'imputation n'était pas possible. Il n'existe aucune relation entre les magasins, ce qui ne donne aucune logique à l'imputation. Aussi, pour des fins de regroupement, nous avons essayé plusieurs formes de normalisation des valeurs des attributs afin de voir leur impact. Nous avons utilisé principalement le z -score, la normalisation min-max, et une mise en échelle décimale. À la fin, nous avons décidé de travailler avec les données sans normalisation vu leurs résultats plus satisfaisants.

5.3 Analyse de regroupement

Avant d'entamer le regroupement, nous avons procédé à la détermination du nombre de groupes à utiliser. La figure 2 résume le tracé des métriques en fonction du nombre de groupe k . Nous avons exécuté nos algorithmes (K -means en cas de la figure 2) avec des valeurs de k varie entre 2 et 10. Puis, nous avons utilisé la somme des carrés des distances (S - C - D) qui caractérise la 'Elbow Method', ainsi que les autres métriques : Silhouette score (S), le score Davies-Bouldin (D - B), Calinski-Harabasz (C - H), la valeur du gap (Gap) et la différence ($diff$) de deux valeurs de gap successives et la variance d'échantillonnage, où on prend la première valeur positive. Nous constatons que la valeur optimale de k change d'une métrique à l'autre. Le tableau 1 résume ces valeurs.

Tableau 1. Détermination du k optimale

	S - C - D	S	D - B	C - H	Gap	$diff$
k -optimale	3, 4 ou 5	2	2	3	8	4

Afin d'être plus flexible aux exigences des décideurs par la suite, nous avons décidé de travailler avec des valeurs de $k = 3, 4$ et 5 .

Après avoir regroupé les données en utilisant ces valeurs de k , nous avons procédé à une élimination successive afin de choisir une seule valeur à la fin. D'abord, les résultats des modèles avec un k égal à 5 retournent toujours un groupe

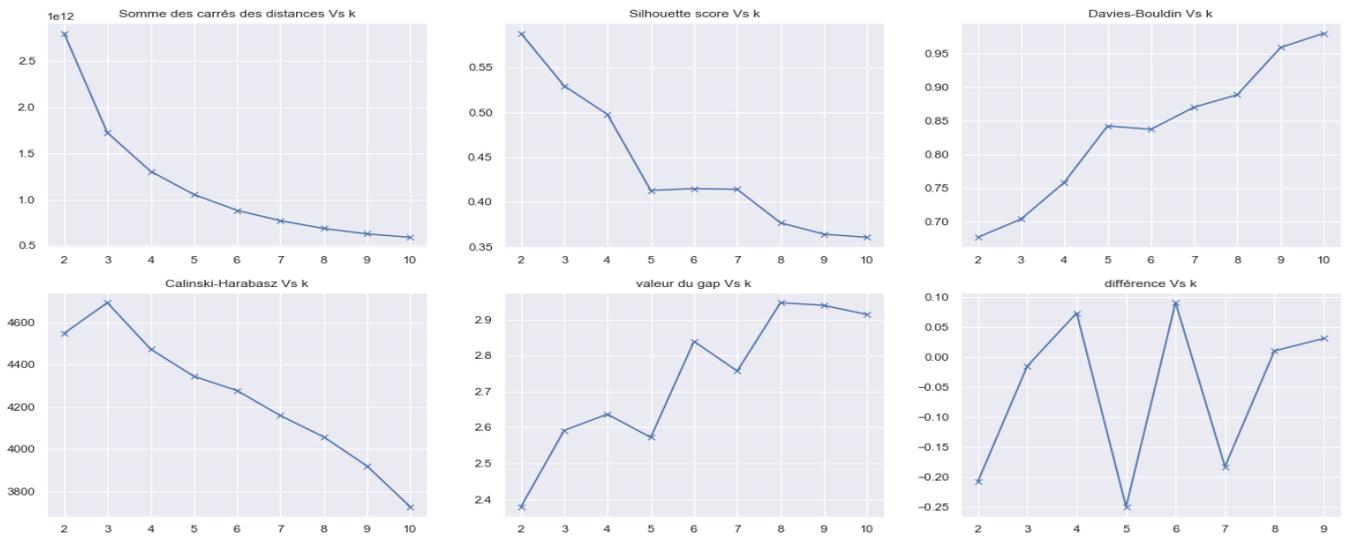


Figure 2. Les graphes des métriques pour la détermination du nombre de groupes

parmi les cinq, qui contient un très faible pourcentage de magasins. Ce qui est inutile pour les décideurs. Alors nous avons éliminé le cas de $k=5$. Par la suite, nous avons comparé les résultats des cas de $k=3$ et $k=4$, tout en consultant les décideurs. Les résultats d'un nombre de groupe $k=3$ n'étaient pas aussi satisfaisants que ceux pour le cas de $k=4$. À la fin, nous avons décidé de rester avec un nombre de groupes égal à 4. Dans ce qui suit, on se contente de présenter les résultats avec un nombre de groupe égal à 4.

Afin de trouver un regroupement idéal des magasins de détail, nous avons utilisé cinq modèles de regroupement différents pour trouver celui le mieux adapté à nos données : (1) un modèle hiérarchique ascendant (HA), (2) k-means (KMS), (3) PAM, (4) Gaussian Mixture (GM), et (5) le fuzzy-c-means (FCM). Pour chaque algorithme, une évaluation interne a été faite pour mesurer la qualité du regroupement en termes de compacité, séparation et stabilité. Les résultats obtenus sont illustrés dans le tableau 2 avec le pourcentage de magasins dans chaque groupe.

Tableau 2. Résultats des regroupements

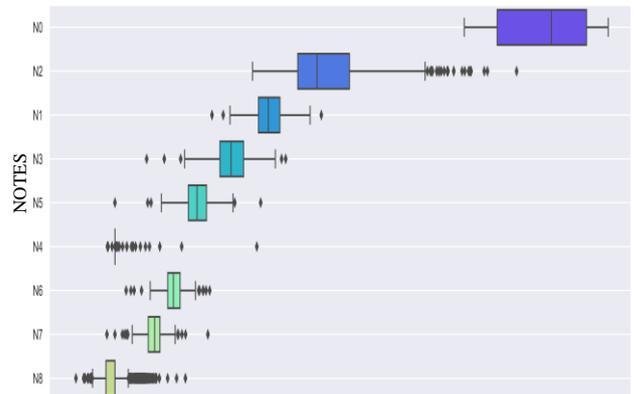
	S	D-B	C-H	Gap	% des magasins de chaque groupe
HA	0.46	0.82	4732	3.02	63 ; 17 ; 5 ; 15
KMS	0.49	0.78	4472	2.72	57 ; 27 ; 12 ; 4
PAM	0.37	0.82	3434	2.55	33 ; 27 ; 25 ; 15
GM	0.41	1.89	1437	1.6	79 ; 15 ; 5 ; 1
FCM	0.45	0.84	4277	2.65	53 ; 25 ; 16 ; 6

5.4 Sélection du meilleur modèle

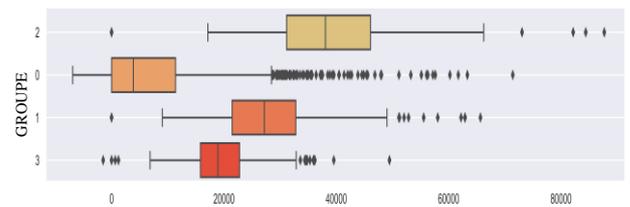
Selon les décideurs, le regroupement désiré doit regrouper les magasins de détail en respectant les volumes des produits vendus et retournés en ordre décroissant du volume. Afin de bien visualiser cette caractéristique, nous avons utilisé la variable 'Note' dans notre base de données. Pour chaque magasin, les décideurs attribuent aux quatre produits une note en fonction du volume vendu. Alors, après avoir fait le

regroupement, nous avons tracé des boîtes à moustache pour chaque produit : dans la figure 3-a nous avons tracé les boîtes à moustache du volume vendu du produit 1 ordonnées suivant la variable 'Note'. Et suivant les groupes construits par notre modèle de regroupement dans la figure 3-b (le cas du modèle hiérarchique). Ceci nous permet : de voir si les regroupements des modèles respectent l'ordre des volumes voulu. De comparer les chevauchements des groupes dans un même intervalle de volume en se référant au classement de la variable 'Note'. Et aussi de voir l'existence des valeurs aberrantes dans les groupes, qui signifie l'existence de magasins qui sortent du vrai intervalle du groupe dont ils appartiennent.

Alors, à partir des métriques, voir tableau 2, on peut conclure sur un bon regroupement donné par la méthode hiérarchique. Or, en consultant la figure 3 (C'est le cas du regroupement hiérarchique pour la quantité vendue du produit 1 qui nous



(b) Les boîtes à moustache en fonction des notes



(a) Les boîtes à moustache en fonction des

Figure 3. Les boîtes à moustache du produit 1 en fonction (a)des notes, et en fonction (b)des groupes hiérarchiques.

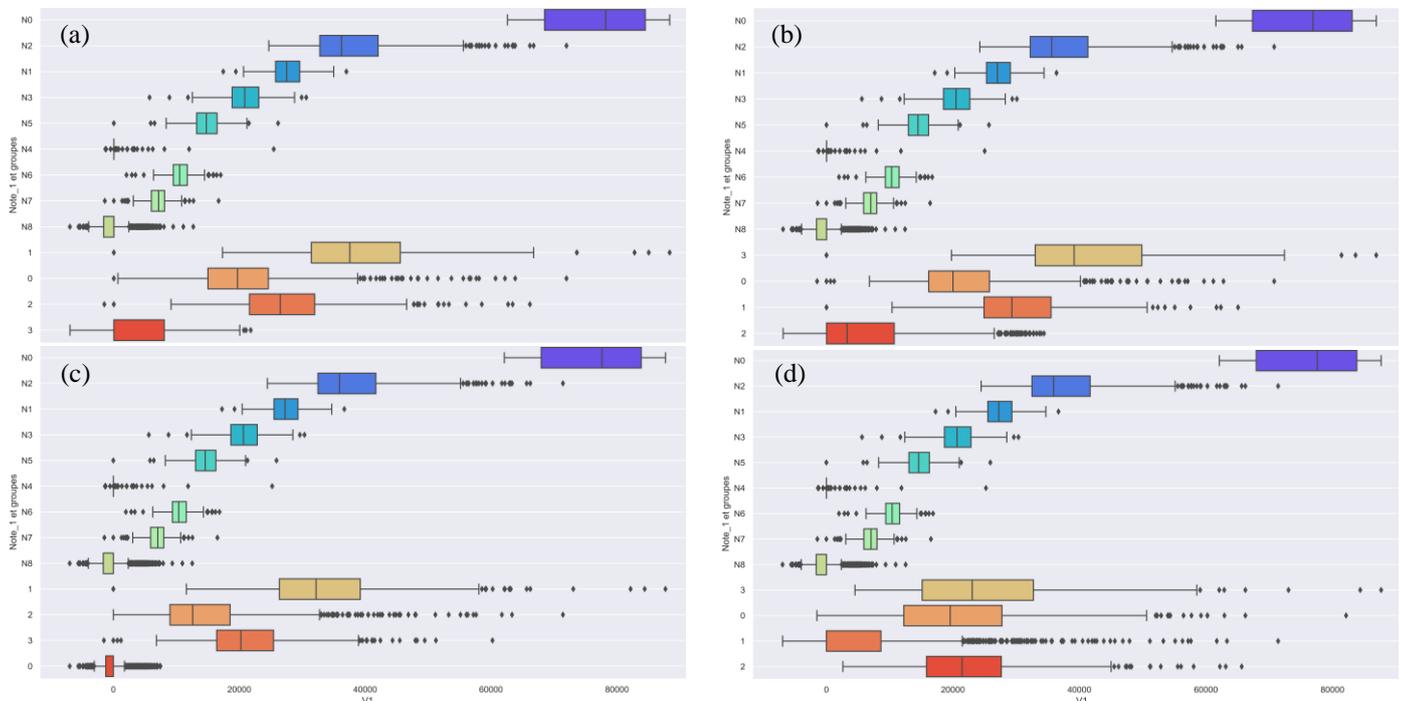


Figure 4. Les boîtes à moustache de la distribution du volume vendu du produit P1 en fonction des groupes des quatre autres algorithmes : (a)Fuzzy-c-means, (b)K-means, (c) PAM et (d) GaussianMixture.

servira d'exemple, mais les résultats sont similaires pour les autres produits 2,3 et 4 pour la quantité vendue, retournée et envoyée) : les intervalles des volumes des quatre groupes se chevauchent entre eux, avec beaucoup de valeurs aberrantes, en particulier pour le groupe 0. Ceci est dû à l'effet de chaînage fréquemment présent dans le modèle de regroupement hiérarchique ascendant [Bastin et Benzècri. 1982]. Donc, nous avons décidé d'éliminer ce modèle.

Par le même raisonnement, la figure 4 représente les boîtes à moustache du volume vendu du produit 1 en fonction des groupes et des notes pour les quatre autres modèles : Fuzzy-c-means (4-a), K-means (4-b), PAM (4-c), et GaussianMixture (4-d). On remarque que le GaussianMixture présente beaucoup de chevauchement entre les groupes et de valeurs aberrantes. Alors que le k-means et le PAM ont réussi à les réduire. Mais, le fuzzy-c-means reste le meilleur modèle qui minimise le chevauchement et les valeurs aberrantes. En plus, il a de bonnes valeurs des métriques d'évaluation, voir tableau 2. Par conséquent, nous avons choisi de rester avec le fuzzy-c-means pour le regroupement. La figure 5 est une représentation des volumes vendus du produit 1 en fonction du produit 2 avec le regroupement fuzzy-c-means. On remarque qu'il résulte à une bonne séparation du volume.

5.5 Caractérisation des groupes

La dernière étape consiste à faire une caractérisation de chaque groupe en fonction des produits envoyés, vendus et retournés, afin d'augmenter la compréhension et l'interprétation des groupes. La méthode utilisée est la construction d'un arbre de classification CART sur le regroupement choisi donné par le fuzzy-c-means. La figure 6 montre les quatre premiers niveaux de l'arbre obtenu. Par la suite, nous allons extraire toutes les

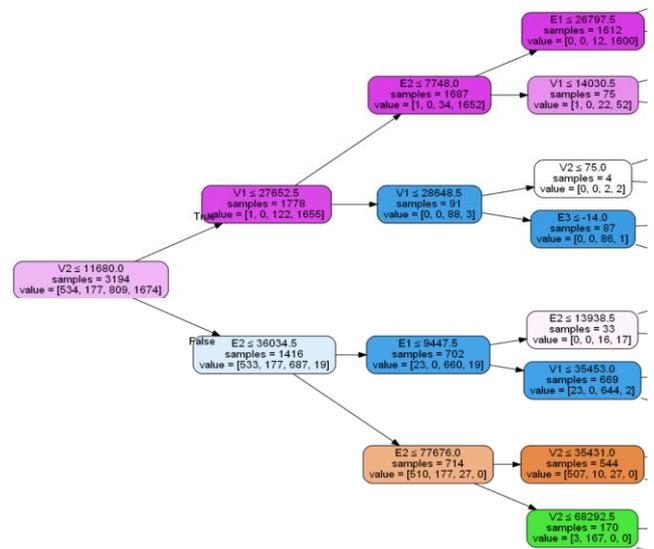


Figure 6. Illustration des quatre premiers niveaux de l'arbre de classification des groupes fuzzy-c-means

conditions liant le nœud racine avec chaque feuille. Chacune de ces branches contient des règles disjointes. Alors, les conditions d'une branche sont uniques et la caractérisation d'un groupe sera l'ensemble des règles représentant ce groupe. On servira d'exemple de la caractérisation du groupe 3 avec ses quatre conditions :

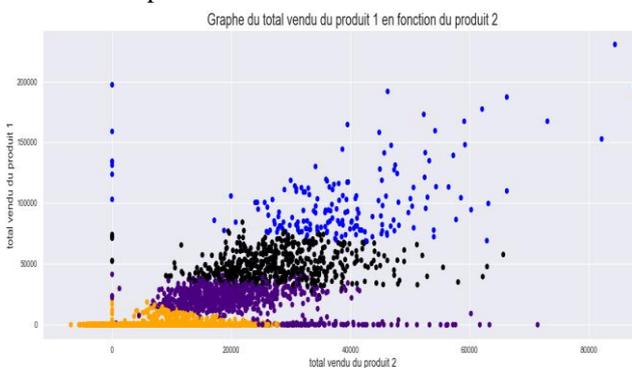


Figure 5. Volume vendu du produit 1, en fonction de celui du produit 2 en utilisant un regroupement fuzzy-c-

- (1) $(V2 > 35) \& (E2 \leq 73) \& (V1 > 15) \& (E1 > 63)$;
 (2) $(V2 > 12) \& (V2 > 35) \& (E2 > 74) \& (V1 > 31) \& (V3 \leq 17)$;
 (3) $(V2 > 88) \& (E2 > 36) \& (E1 \leq 10)$;
 (4) $(V2 > 68) \& (E2 > 36) \& (E1 > 10)$.

Avec V_i le volume vendu du produit i , E_i celui envoyé et avec des chiffres approximatifs en milliers. Tous les autres groupes ont été caractérisés de la même façon. Une telle caractérisation permet de bien comprendre les spécificités de chaque groupe. Par conséquent, de définir les stratégies de distribution pour chaque groupe de magasins.

Pour les décideurs, les résultats étaient assez cohérents et intéressants vu qu'il sépare très bien les magasins en fonction des volumes. En particulier au niveau des produits retournés, le FCM a réussi à bien séparer les produits en fonction des produits retournés. Dans cette période de données : 13% des produits retournés par les magasins de la classe 0. Les magasins de la classe 1 compensent leurs retours (il y a des valeurs négatives compensées entre les magasins) entre eux et ont 0% dans la globalité des retours. 45% des produits retournés par ceux de la classe 2. Et 42% des produits retournés par ceux de la classe 3. Ceci va permettre aux décideurs de cibler leur gestion de retours vers les classes qui retournent le plus.

6 CONCLUSION

Dans cet article, nous avons introduit une approche de segmentation et de caractérisation des magasins de détail des produits cosmétiques. Les chaînes d'approvisionnement des produits cosmétiques sont très complexes à raison de la demande hétérogène, de la variété des produits, et une distribution très compliquée. Sa chaîne logistique se compose de différents niveaux de distribution. Chacune commence par un grand détaillant qui gère sa demande en fonction de sa capacité de stockage sans aucune intervention externe. Ceci gêne les intérêts du producteur et ajoute plus de pression au niveau de la concurrence. Par conséquent, les décideurs se doivent faire face à une demande et une concurrence changeante.

L'approche proposée a utilisé les données historiques des produits envoyés, vendus et retournés des différents magasins de détail pour faire un regroupement en utilisant des techniques d'exploration des données, suivie d'une caractérisation des groupes avec des arbres de classification CART.

Le regroupement des magasins de détail obtenu permet une compréhension des possibles caractéristiques de la chaîne d'approvisionnement. En plus, les groupes constitués portent des caractéristiques similaires, qui peuvent donner une vision sur la politique à adopter. Cependant, une caractérisation du regroupement final en fonctions des attributs permet d'extraire des interprétations utiles.

Cette approche peut être étendue en ajoutant plusieurs autres attributs. Un travail futur consistera, d'une part, à ajouter le facteur géographique et démographique des magasins de détail, et le niveau de service de chacun. Ainsi l'utilisation d'un échantillonnage 'bootstrap' pour réduire la variance et contrer le surapprentissage des arbres de décision. D'autre part, nous allons faire des travaux de prédictions groupe par groupe pour aider la prise de décisions du producteur.

7 REFERENCES

- Helmi, R. A. A., Alkawaz, M.H. & Rahmat, S.N. (2020). Trusted Halal Application for Malaysian Cosmetic products. *16th IEEE International Colloquium on Signal Processing & its Applications*, 1-6.
- Han, J., Pei, J. & Kamber M. (2012). DATA MINING: concepts and techniques (3rd éd). Tokyo. MORGAN KAUFMANN
- Berkhin, P. (2006). A Survey of Clustering Data Mining Techniques. Dans Jacob, K., Charles, N., Marc, T. (Ed) *Grouping Multidimensional Data*. (pp. 25-71). Berlin/Heidelberg. Springer-Verlag
- Guyon, I., Luxberg, U., V., & Williamson R., C., (2009). *Clustering: Science or Art?*. 11(1)
- Derrouiche, R., Pongsak, H. & Leksakul P., H., K. (2011). Predictive performance model in collaborative supply chain using decision tree and clustering technique. 4th International Conference on Logistics. pp. 412-417. Italie
- Danish, I., Xu, X., Deng, S. & Imran, A. K. (2007). Clustering Framework for Supply Chain Management (SCM) System. *Second Workshop on Digital Media and its Application in Museum & Heritages*. pp. 422-426
- Johnson, S.C. (1967). Hierarchical clustering schemes. *Psychometrika*. 32, 241-254.
- Dempster A.P, Laird N.M, & Rubin D.B. (1977). "Maximum Likelihood from Incomplete Data via the EM algorithm". *Journal of the Royal Statistical Society, Series B*, 39(1): 1-38.
- Bezdek J.C., Ehrlich, R. & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. 10(2-3). 191-203.
- Hartigan, J. A., & Manchek A. (1979). "Algorithm AS 136: A k-means clustering algorithm." *Journal of the Royal Statistical Society*. 28(1). pp. 100-108.
- Kaufman, L. & Rousseeuw, P.J. (1990). *Fitting Groups in Data. An Introduction to Cluster Analysis*. Wiley, New York.
- Halkidi, M., Yannis B., & Michalis, V. (2001) "On clustering validation techniques." *Journal of intelligent information systems*. 17(2). 107-145.
- Halkidi, M., Yannis B., & Michalis, V. (2001) "On clustering validation techniques." *Journal of intelligent information systems*. 17(2). 107-145.
- Tibshirani, R. Walther, G. & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 63(2). pp. 411-423.
- Sugar, C. (1998). *Techniques for clustering and classification with applications to medical problems*. PhD Dissertation. Stanford University, Stanford.
- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont: Wadsworth.
- Pratama, I., Permanasari, A.E., Ardiyanto, I. & Indrayani, R. (2016). A review of missing values handling methods on time-series data. *International Conference on Information Technology Systems and Innovation*. pp. 1-6
- Bastin, C. & Benzècri, J.P. (1982). Exemple d'effet de chaînage complet autour d'un centre en classification ascendante hiérarchique. *Les cahiers de l'analyse des données (tome 7, no 2)*. pp. 189-208.
- Murray, P. W., Aguard, B. & M.B. Barajas. (2015) Forecasting Supply Chain Demand by Clustering Customers. *International Federation of Automatic Control*. 48-3
- Alexander A. Kharlamov, Luis Miguel D. F. Ferreira & Janet Godsell (2020) Developing a framework to support strategic supply chain segmentation decisions: a case study, *Production Planning & Control*, 31:16, 1349-1362, DOI: 10.1080/09537287.2019.1707896
- Murray, P. W., Aguard, B. & M.B. Barajas. (2017) Market Segmentation through data mining: A Method to extract behaviors from a noisy data set.

Session 4.5 : Industrie du futur - Transition numérique

Chairman : Bernard Grabot - Jean-Luc Maire

A Semantic Model for Cyber-physical production system inspired from Current developments

PUVIYARASU S.A.¹, CATHERINE DA CUNHA¹

¹ LS2N, Ecole centrale de Nantes
1 rue de la Noe, 44321, Nantes, France
puviyarasu.sa@ls2n.fr
Catherine.da-cunha@ec-nantes.fr

Abstract – Cyber-Physical system (CPS) is a coupling of physical and cyber components with a networked intertwined connection. It involves computational elements able to operate on different scales and interact with each other when a change in context occurs. CPS applications for production are termed Cyber-physical production systems (CPPS). It is a new and foreseeable development, whose degrees of integration has reached a self-deciding function possible in production and reflects the industry 4.0 production paradigm. There is a need of modeling approaches to manage complexities, structural opacities and provide a common understanding of CPPS. This paper addresses the semantic understandings of CPPS inspired by the current developments; it classifies the involved entities and illustrates the relations on different aspects. This will serve as a foundation for CPPS, in particular, by helping to understand the structural ambiguity. The real LS2N-CPPS is a supporting use-case of this research work.

Keywords – Cyber-physical production system (CPPS), Semantic models, Smart machines and products.

1 INTRODUCTION

Industries rely on digital transformation and emergent technologies to have flexible and efficient processes. Cyber-Physical System (CPS) is an enabler for the revolution 4.0. CPS is a coupling of physical and cyber components with networked connections. It involves computational elements able to operate on different scales and interact with each other when a change in context occurs [Chen, 2017]. The main capabilities involve computation, communication, and control. CPS applications for production are termed Cyber-Physical production systems (CPPS). [Monostori et al., 2016] defined that the CPPS systems are autonomous and cooperative elements connecting in situation-dependent ways, on and across all production levels, from machines up to production and logistics networks. It enhances the decision-making processes, response to unforeseen conditions in real-time. The basic functionalities of CPPS are Connectedness, Responsiveness, and Intelligence [Cardin, 2019] (See figure 1).

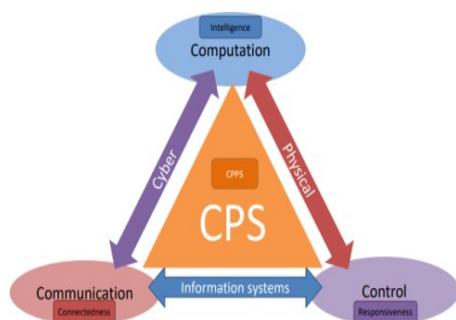


Fig.1 Capabilities of CPS and analogy to CPPS [Cardin, 2019]

The key characteristics of CPPS are modular, reconfigurable and networked system. It enhances the communication between the machines, people, and products in the ecosystem.

The key capabilities are, 1) Self-sensing: they are equipped to capture data and critical information from the environment involving product, quality, materials, machines etc. 2) Self-deciding: they can take data-driven decision in manufacturing, including the identification, collection, communication, analysis and learning the system by themselves. 3) Self-adapting: they adapt to changes in real market demands.

In addition, the system has modernized controlling capabilities, which are decentralized and centralized system [Weyrich et al., 2017]. The centralized level implies the availability of a single or few entities that have control over the entire network. It includes common PLC's, drive controllers, Industrial PC able to provide the highest performance on the top level. The decentralized level implies that single entities may control themselves. It includes plug and plays principles, HMI enable the decentralized control [Gronau et Theuer, 2016].

From the different dimensions of current developments, the CPPS learning system illustrates the smart production paradigm and thereby redefines numerous areas of modular production system.

From the previous CPPS studies, a lot of CPPS research challenges, capabilities, classification, and characteristics are discussed but there is a lack of studies on common understandings, their involved entities and relations. There is a need of semantic understandings on CPPS, their involved entities and the relations that enable them to address the system structural ambiguity and manage system complexity.

This paper is structured as follows: section 2 describes the

related works on CPPS and their current developments, section 3 describes the proposition: the identified entities and semantic models. Section 4 presents an application of semantic models on LS2N- Cyber-physical Production system, which is the use- case of this research work. Conclusions are drawn in section 5.

2 STATE OF THE ART

CPPS Semantic architectures related works and current developments are specified in this section. [Darwish et Hassanien, 2018] specifies the key aspects of CPPS and architecture for CPPS. The study doesn't illustrate the entities and the interconnections. [Tomiyama et Moyon, 2018] presents a resilient architecture to handle the event-driven process. The architecture concerns only limited entities and their relationships. [Sánchez et al., 2016] presented a CPS-based process control solution for smart manufacturing by focusing on underlying architectures and services. The study does not stress all aspects involved in the system. [Agostino et al, 2018] develops CPPS architectures, which assist in integrating data collection and feedback systems. The study focuses on sensors and data processing elements leaving out the entities. [Gronau et Theuer, 2016] proposed a hybrid architecture approach to stimulate the various degree of autonomy of CPPS. They stress that common CPPS modeling approaches are necessary to depict a wide variety of production processes with different production. [Cardin, 2019], specified that current Cyber-physical production system (CPPS) applications are in different categories which include, 1) Cyber-Physical production (Learning) system 2) Lab XP and 3) Industrial CPPS system. Next section shows the identified involved elements of CPPS.

3 IDENTIFIED ELEMENTS AND PROPOSED SEMANTIC MODEL OF CPPS

3.1 Identified involved Elements

In this section, we showed the identified levels and elements of CPPS. [Zamfirescu et al, 2014] shows the distinguished guide and exclusive nature of cyber-physical system and their abstraction levels. The levels include "cyber", "physical" and "feedback computations". It poses an exclusive nature and structural view of the Cyber-Physical system. His Cyber-Physical system reference framework model helps us to differentiate abstract levels of the system. By using the current CPPS classification 1) CPPS- learning system, 2) Lab XP CPPS system, 3) Industrial CPPS System on structural basis, we classify each abstraction levels of CPPS [Puviyarasu S. A et al, 2020].

The different levels of abstraction include "cyber level", "physical level" and "link level" as a natural abstraction. This abstraction poses the standard structure and functionalities of the system. Furthermore, the system has its additional abstraction: humans intervene to make informed decisions and perform production tasks. So, we define the additional abstraction levels as "Human-machine agents" and "Operators". In addition, the CPPS system supports the external system to coordinate and influence the external system service to perform. So, we add the level "External system infrastructure". As of the currently developed systems, the CPPS is influenced directly or indirectly by other systems like transportation, miscellaneous activities etc. So, we define the abstraction

level "Support system". This abstraction is a supplement from the current development. These above abstraction levels are considered generic categories, which helps structure the CPPS and identify the entities involved in each level.

The next section shows the account of each of the 7 abstraction levels and their descriptions. Then, the relation between the identified entities and layers is illustrated.

3.2 Abstraction levels description

3.2.1 Cyber parts/layer

We define "Cyber part/layer" as the level that refers to intangible components. It encompasses the hardware, software that helps store data, analyze, process, collect, control, and actuate the information within the CPPS ecosystem. The Cyber layer acts as a central medium in the system that pushes each connected entity to enable the task. It connects people, product and production system.

Other authors may use "Cyber layer", "Cyber and network layer", "Cyber stack", "Cyber technologies", "Cyber components", "Virtual world", Network world", "Cyber world".

3.2.2 Physical parts/layer

The Physical Parts/layer refers to all tangible components that actively or passively participate in the production process to add value. In CPPS, the machine components and physical product components are referred to in the physical layer. It comprises raw materials and machines physical object that forms an integral part to produce tangible asset.

Other authors may use "Physical layer", "Physical Stack", "Physical components", "Real-world", "Physical part" [Bocciarelli et al, 2017].

3.2.3 Linking parts/layer

The Linking Parts/layer refers to the elements connecting or intersecting two things, in particular to combine the real and virtual world of production. Linking components, "Interfacing components", "combined layer", "Link layer" are the other terms used in the literature [Akanmu et al, 2012].

3.2.4 Human-machine agent parts/layer

Human-machine agent layer refers to the elements that enable operators to manage industrial and process control machinery via a computer-based user interface. From the current developments, the human-machine agents feature a component of a certain device, or software application that enables humans to engage and interact with machines [Weyrich et al., 2017].

3.2.5 Operator's layer

An operator is an individual who operates the equipment or machine in the factory in order to perform a global task. In currently developed systems, there is a need for human intervention to make decisions. We define this layer to encompass the human in the loop with production entities. There are three different modes of operation where the interaction between humans and the inner system (CPPS) occurs. The modes are centralized, decentralized, and hybrid scenarios. In current developments, this abstraction provides the necessary freedom to act in the ecosystem. In the next section, the semantic model for CPPS is illustrated.

3.2.6 Support system layer

The support system defines the system inside the system of interest (SOI) boundaries like transportation, miscellaneous system, etc., directly or indirectly influencing the CPPS.

3.2.7 External system and Infrastructure layer

External system and infrastructure layer specifies the external elements that interact with the Cyber-physical production system. As of the current developments, CPPS connects with external system such as other partner facility, ERP, transportation system and other Enabling system.

3.3 Proposed generic semantic model for CPPS

Semantics is the study of meaning in programming languages, formal logic, and semiotics. It focuses on the various representational theories of coherence and correspondence meaning [Lappin et Fox, 2015].

Semantic models are used to represent a knowledge model with the core content of the domain and set of relations. It contributes to understanding the whole problem of interest. It also supports the process of interrelating information and systems from diverse sources. It facilitates communication between the interested parties. The commonly used modeling languages are unified modeling language (UML), System Modelling Language (SYSML), XML, and Automation ML.

UML modeling is appropriate for modeling the structure of the system [Rumbaugh et al, 2000]. The class diagram is a container for many attributes, objects that share specific semantics, relations, and behaviors. In this work, UML model language is used to represent the semantics of the cyber-physical production system, involved entities, and their relations.

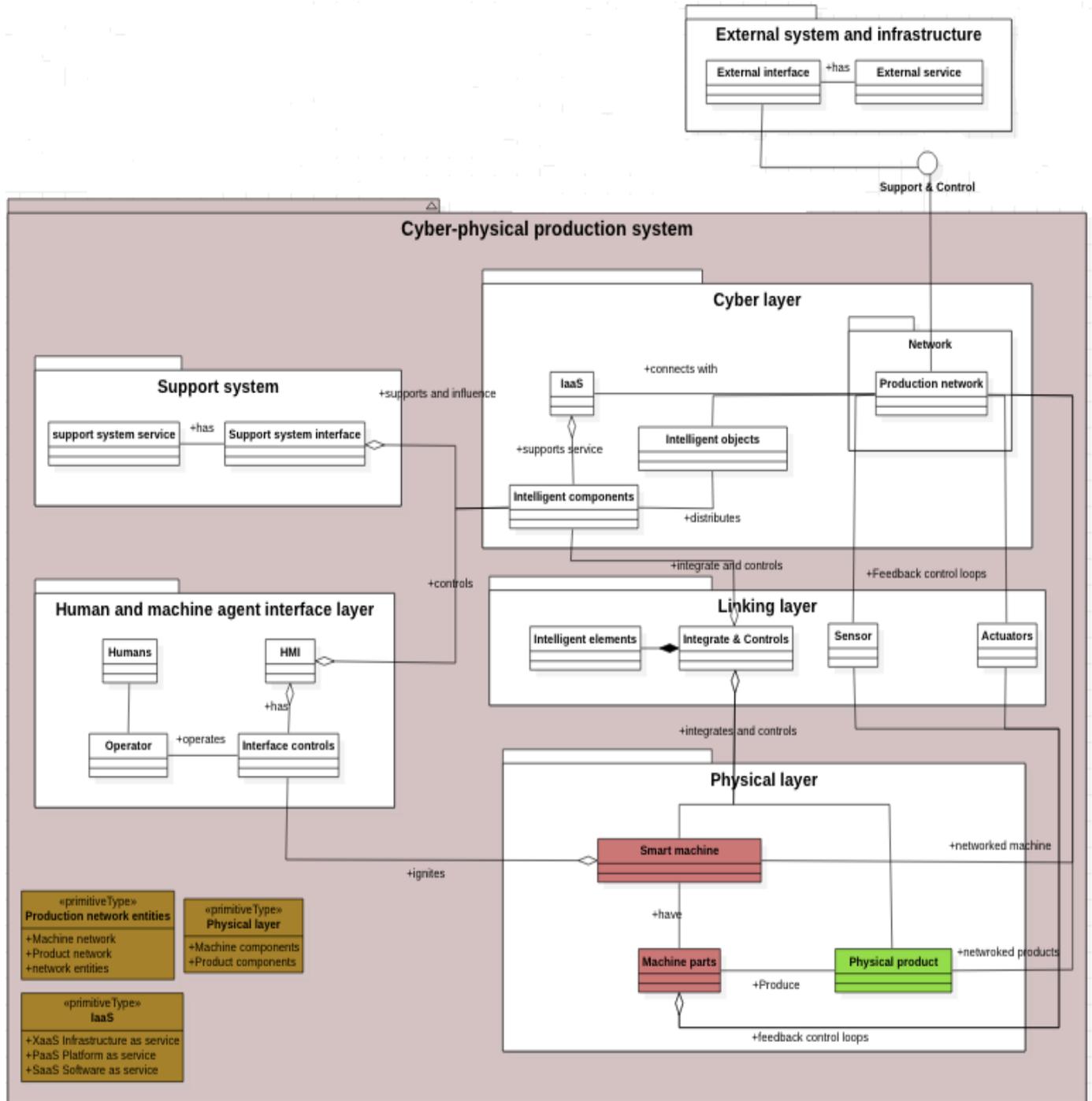


Fig.2. Proposed Generic Semantic model for CPPS



Fig 4.LS2N- Application Magazine machine (CPPS)

In this section, we instantiate the generic model described in section 3 on the magazine machine (See figure. 5). In the cyber layer, the “IaaS” is a cyber-infrastructure component that provides the services in the cyber layer. Only the main cyber-infrastructure “IaaS” component is specified in this section. For Magazine machine, the main infrastructure component are

- 1) MES (Manufacturing execution system) software which provides the service of assembly machine to track and document the transformation of raw material to finished product.

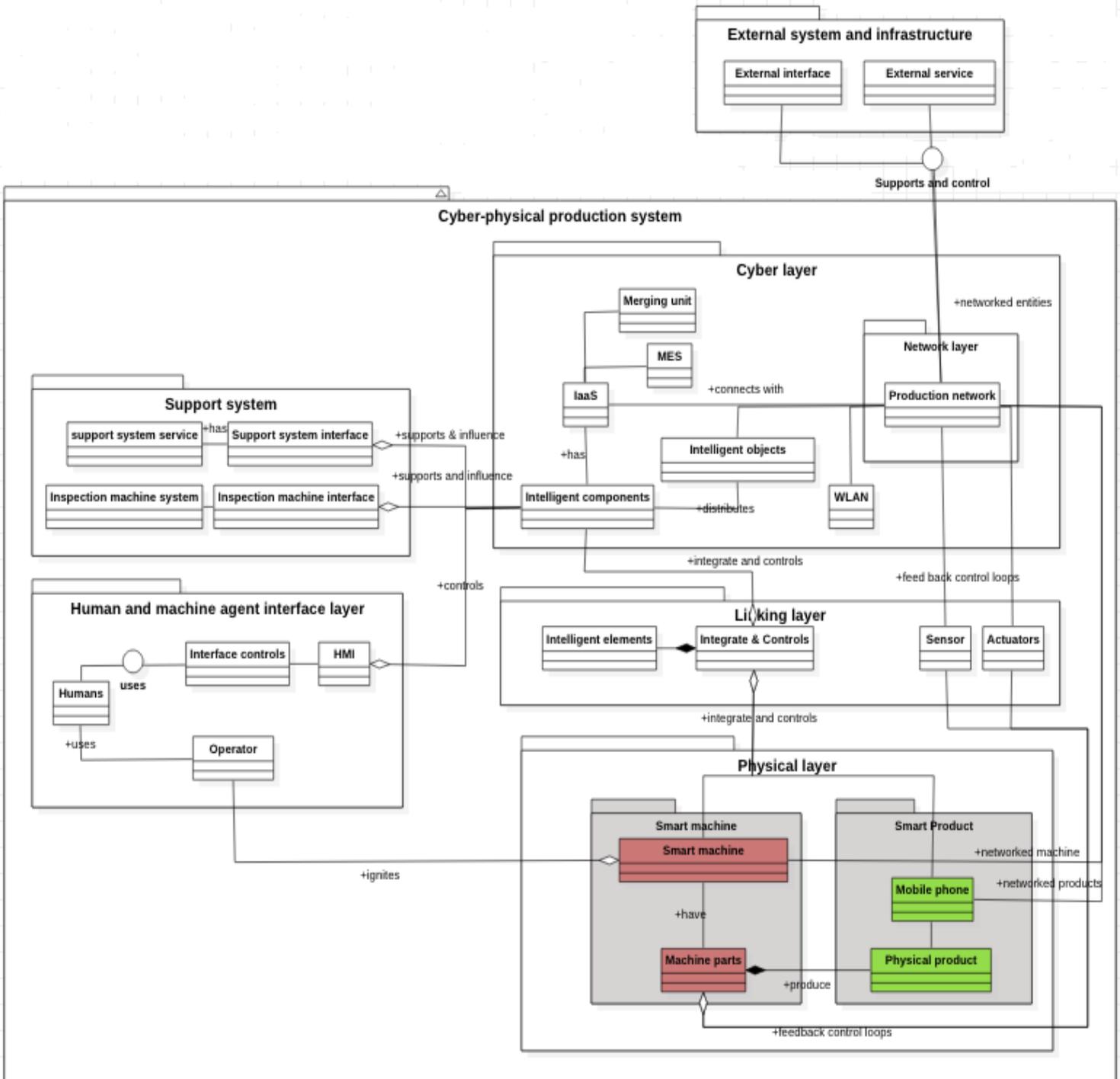


Fig 5. LS2N-Application magazine machine Semantic model (CPPS)

2) Cyber Merging Units- It is an integral unit in the cyber-infrastructure which is fulfilling the task of controlling and monitoring the top cover assembly part and collecting the data via the Ethernet network.

The Syslink connectors help to control its behavior and adapts to the changing environments. It acts both autonomously and together with humans.

The router defines the global networking of hardware and software components of the LS2N-CPP Learning system with help of WLAN automation network.

In the linking layer, the “intertwined magazine objects”- valve In/Op terminal, connect the real and virtual world of the magazine machine. It actuates the machine to perform a task based on magazine machine operation mode (manual or automatic).

In the physical layer, we find the physical tangible machine parts. In this layer is the RFID reader/writer tag which retrieves the information of the product and then writes the data describing the jobs done on each workpiece.

The Magazine Cylinders are “Machine parts” that help to store and separate the workpiece. Value terminal and magazine slot physical part help to assembly the “Top cover part”- Mobile phone. The Magazine machine transforms the semi-finished mobile phone into top-cover part assembled mobile phone.

The “machine parts” of magazine machine has a composition link with “Top cover part” workpiece.

The human-machine interface agent layer is composed of the control panel and power. When individual operators use the machine, they use power switches to ignite, and control panel to control and set the parameters of the Magazine machine.

The support system of the mobile assembly machine contains the transportation devices: the conveyor belt and a “mobile robot system”- which helps to transport the product from one station to another.

External system infrastructure layer- The external system of the Magazine machine is the Warehouse facilities –This Warehouse facilities allows the Assembled finished product to be stores. The Magazine machines provide the service of storing the top cover mobile case product directly to the external system.

5 CONCLUSION

The paper proposes a semantic model for the Cyber-Physical production system (CPPS) inspired by the current developments. It fills literature gaps. It clarifies the semantic understandings of CPPS, involved entities, and their relations aspects. The work lays a foundation for managing system complexity and understanding the structural ambiguity. The instantiation of the generic model on a real test-bed prove its applicability. Further research should consider application on industrial systems.

6 REFERENCES

Abele, E., Metternich, J., & Tisch, M. (2019). Learning Factories. In Concepts, Guidelines, Best-Practice Examples, Springer charm, Switzerland.

Agostinho, C., Ferreira, J., Ghimire, S., Zacharewicz, G., Pirayesh, A., & Doumeings, G. (2018). A Comprehensive Architecture to Integrate Modeling and Simulation Solutions in CPPS. *Enterprise*

Interoperability: Smart Services and Business Impact of Enterprise Interoperability, 349-355.

Akanmu, A. A., Anumba, C.J., & Messner, J.I., (2012) An RTLS based approach to cyber-physical systems integration in design and construction, *International journal of distributed sensor Networks*, 8(12), 596845

Bocciarelli, P., D’Ambrogio, A., Gigilio, A., Paglia, E. (2017), A BPMN Extension for modelling Cyber-physical production systems in the context of Industry 4.0, 14TH IEEE International conference on Networking, Sensing and Control (ICNSC)(pp.599-604), Calabria, Italy.

Cardin, O (2019), Classification of cyber-physical production systems applications: Proposition of an analysis framework. *Computer in Industry*, 104, 11-21.

Chen, H. (2017). Theoretical foundations for cyber-physical systems: a literature review. *Journal of Industrial Integration and Management*, 2(03)175.

Darwish, A, & Hassanien, A.E (2018). Cyber physical system design, methodology, and integration: the current status and future outlook. *Journal of Ambient intelligence and Humanized computing*, 9(5), 1541-1556.

Festo didactic cyber-physical factory, A universal Industry 4.0 training factory. A guided tour on the learning industry 4.0 system, doi 1058@2017/12.

Gronau, N., & Theuer, H. (2016). Determination of the optimal degree of autonomy in a cyber-physical production system, *Procedia CIRP*, 57, 110-115.

Lappin, S., & Fox, C. (2015), *The handbook of contemporary semantic theory*, John Wiley & sons.

Monostori, L, Kadar, B., Bauernhnsi, T., Kondoh, S., Kumara, S., Reinhart, G.,... & Ueda, K. (2016). Cyber-physical systems in manufacturing .*CIRP Annals*, 65(2), 621-641.

Puviyarasu S.A, Belakdi, F, da Cunha C., Chriette, A., Bernard, A. (2020) A semantic model interface model to support the integration of drones in cyber-physical factory. *International conference on interoperability for enterprise systems and applications*, Tarbes, France 2020.

Rumbaugh, J, Booch, F & Jacobson, I. (2000) *Le guide de l'utilisateur UML (Vol.3)*. Eyrolles.

Sanchez, B.B., Alcarria, R., de Rivera, D.S., & Sanchez Picot, A. (2016). Enhancing process control in Industry 4.0 Scenarios using Cyber-physical systems. *J. Wirel, Mob. Networks Ubiquitous Computing Applications* 7(4), 41-64.

Tomiyaama, T., & Moyon, F. (2018). Resilient architecture for cyber-physical production systems. *CIRP Annals*, 67 (1), 161-164.

Weyrich, M., Klein, M., Schmidt, J. P., Jazdi, N., Bettenhausen, K.D., Buschmann, F.,... & Wurm, K. (2017), Evaluation model for assessment of Cyber-physical production systems. *Industrial internet of things* (pp.169-199). Springer, Charm.

Zamfirescu, C.B., Pirvu, B.C., Gorecky, D. & Chakravarth, H. (2014). Human-centered assembly: A case study for an anthropocentric cyber-physical system, *Procedia technology* 15, 90-98.

Intégration verticale pour le pilotage des systèmes de production : un état de l'art

THOMAS CADORET^{1,2}, CHRISTOPHE DANJOU¹, JULIEN LE DUIGOU²

¹ Polytechnique Montréal
2500 Chemin de Polytechnique, Montréal, QC H3T 1J4, Canada
thomas.cadoret@polymtl.ca, christophe.danjou@polymtl.ca

² Université de technologie de Compiègne, Roberval (Mechanics energy and electricity),
Centre de recherche Royallieu - CS 60 319 - 60 203 Compiègne Cedex, France
thomas.cadoret@utc.fr, julien.le-duigou@utc.fr

Résumé – A l'ère de l'industrie 4.0, l'accent est mis, au sein des sites de production, sur la capitalisation de la donnée. La tendance actuelle tend vers un site de production agile, une conception personnalisée. Cela ne sera possible qu'en impliquant totalement les différents acteurs, de la conception à la production, de la gestion de production à la fabrication. Dès lors se posera la question de l'interopérabilité entre les différents systèmes de production, mais au-delà entre les différents systèmes de l'entreprise. L'importance de relever ces challenges d'interopérabilité a été identifié dans le rapport Insight du World Economic Forum, mettant l'accent sur la nécessité de régler ces problèmes d'interopérabilité non seulement à une échelle nationale, mais globale afin de parvenir à l'Industrie 4.0 (Burns et al., 2019). Ce posera également la question du temps de réactivité de ces systèmes de production. Ainsi la notion de temps-réel concernant le pilotage de ces systèmes devra être exploitée. Ce document va effectuer un état de l'art des propositions de solutions actuelles afin d'assurer une interopérabilité entre les différents organes de l'entreprise, afin de capitaliser pour réutiliser efficacement les données produites par ces différents organes.

Abstract – In the era of Industry 4.0, the emphasis concerning production sites is on data capitalization. The current trend is on agile production sites, smart production and conception. This will only be possible by fully involving the various players, from design to production, management to manufacturing. The question then arises of interoperability between the different production systems, and beyond, those between the different organs of the company. The importance of solving these interoperability challenges have been identified in the Insight report by the World Economic Forum, which outlined the need to resolve these issues not solely at a national level but globally, in order to implement Industry 4.0 (Burns et al., 2019). Another question that need to be solved is about the reactivity time of these production systems. Thus, the notion of real time concerning driving these systems will have to be approached. This document will carry out a state of the art of current solution proposals in order to ensure interoperability between the various organs of the company, in order to capitalize the data, to have an efficient reuse of the data produced by these various organs.

Mots clés - interopérabilité, industrie 4.0, production intelligente, intégration verticale, ontologies

Keywords – interoperability, industry 4.0, smart manufacturing, vertical integration, ontology

1 INDUSTRIE 4.0

Depuis une décennie nous observons une transformation des moyens de production laissant penser une quatrième révolution industrielle : Industrie 4.0 (Drath et Horch, 2014). Tandis que la troisième pouvait être définie comme l'automatisation et la robotisation des chaînes de production, la quatrième, quant à elle, vise à développer et améliorer la connectivité au sein des sites de production. L'Industrie 4.0 a donc pour but d'ajouter une capacité d'intelligence aux sites de productions, dans un monde où l'agilité et la personnalisation de masse est désormais de mise. Cela passe notamment par l'utilisation de technologies telles que les CPS (Cyber-Physical Systems, ou systèmes cyber-physiques), l'IoT (Internet of Things, ou Internet des objets), et de standards afin de permettre à ces différents outils de communiquer entre eux, mais également avec les couches supérieures tel que les systèmes de contrôle et

d'acquisition de données (SCADA), les logiciels de pilotage de production (MES) ou encore les logiciels de gestion (ERP). Les systèmes de fabrication modernes doivent être plus flexibles et plus réactifs pour répondre aux changements de l'industrie et des clients. La mise en place d'une infrastructure numérique plus riche en informations ouvre la voie à cette flexibilité. Afin de mettre en œuvre ces changements, l'industrie 4.0 a des objectifs tels que la mise en place d'une architecture de référence, une gestion efficace et l'amélioration de l'efficacité de l'utilisation des ressources.

Une personnalisation de masse ne peut se faire qu'en exploitant un maximum de données provenant des organes de production. Plus le niveau de détails est accru, plus il est possible de réagir rapidement à des aléas de production. Ce niveau de détails sert également, à terme, à l'amélioration et à l'optimisation de la production (comme il a déjà été envisagé

d'exploiter les données de la production pour guider et optimiser la FAO) (Ostermeyer et al., 2019). Cela constitue une première étape. La seconde est un retour d'information provenant directement des systèmes de décision, afin de planifier toujours plus rapidement et précisément la production de lots uniques. Un pilotage en temps réel faisant intervenir efficacement les systèmes de gestion permet en effet une optimisation de cette production (permettant l'anticipation de phénomènes propres à l'entreprise, et une utilisation intelligente des systèmes de production). Il est question ici de production globale donc le but est d'encourager l'interopérabilité en ce qui concerne la fabrication, mais appliqué au processus et élargi à des standards définissant les moyens de communication (Newman et al., 2008).

L'industrie du futur voit naître de plus en plus de données, que ce soit grâce à l'avènement de l'IIoT (L'utilisation de l'Internet of Things dans l'industrie) ou des systèmes de production intelligents. Il faut la traiter, et traiter efficacement (c'est-à-dire sans perte de précision, et sans erreur) afin de permettre non seulement une collecte de cette dernière mais au-delà une capitalisation et ainsi permettre une valorisation de la donnée capitalisée. Il est possible de supporter une intelligence décentralisée afin de permettre l'analyse et la prise de décision. (Goncalves et al., 2011). Cependant, le problème qui survient lorsque l'on parle de capitalisation de données provient de la donnée elle-même. Celle-ci est souvent hétérogène, que ça soit au niveau du fond (pas toujours le même niveau de détails) ou de la forme (nombres de formats et de normes existents). Comment assurer une interopérabilité verticale au sein du site de production pour permettre un pilotage optimisé des systèmes de production ? Il s'agit, dans cette étude, de faire un état de l'art, puis relever les manques qui pourraient subsister en ce qui concerne l'interopérabilité entre systèmes de production, mais également entre systèmes de production, outils de pilotage et outils de décision.

Dans la section 2, une définition approfondie de l'interopérabilité est apportée. Il est particulièrement question de comprendre les enjeux et défis concernant l'intégration verticale, dont la notion est définie dans cette partie. La troisième partie définit précisément les architectures de références, fournissant les outils nécessaires à une interopérabilité organisationnelle, afin d'aborder au mieux la transition vers l'Industrie 4.0, et développant ainsi la communication entre systèmes au sein du site de production, ou entre ceux-ci et les acteurs extérieurs (ERP, clients, ...). Dans une quatrième partie, l'accent est mis sur les différents standards existants à l'heure actuelle pouvant apporter un réel avantage quant à l'interopérabilité technique, qui peut être définie comme les composantes logicielles ou matérielles, les systèmes et les plateformes qui permettent d'assurer la communication d'une machine à l'autre, et quant à l'interopérabilité sémantique, c'est-à-dire la capacité à interpréter automatiquement l'information échangée avec justesse et sens. Une cinquième partie se concentre davantage sur le concept d'ontologies, et en quoi il apporte une réponse lorsque l'on cherche à obtenir une interopérabilité sémantique. Le concept d'ontologies de références, ainsi que le processus de réconciliation sémantique sont abordés. Puis dans une sixième partie et dernière partie, une conclusion est apportée concernant les possibilités actuelles concernant le partage d'informations dans l'entreprise ainsi que les enjeux pour le futur de l'industrie.

2 L'INTEROPERABILITE : UNE SOLUTION

La première étape pour définir l'interopérabilité, c'est définir le périmètre, c'est-à-dire les différents systèmes concernés par l'étude, pour comprendre leur façon de fonctionner, et leur capacité à communiquer avec l'extérieur, dans quel langage. L'interopérabilité, d'après l'Institute of Electrical and Electronic Engineer, c'est la capacité de deux systèmes, de deux composants ou plus à échanger des informations et à utiliser ces informations venant d'être échangées (IEEE, 1990). Au-delà d'une capacité de communication, il est crucial pour assurer une interopérabilité entre deux systèmes, que ces deux systèmes se comprennent. Il est alors question d'interopérabilité sémantique. Un langage commun peut être envisagé, ou une couche de traduction peut être incluse à la phase de communication. Se pose dès lors la question de l'intégration des systèmes. Réaliser une intégration, c'est rendre deux systèmes interdépendants et non séparables. C'est dans ce contexte qu'un cadre est fixé, afin de savoir s'il est judicieux d'aller jusqu'à une intégration. A contrario, la compatibilité est définie comme l'étape située avant l'interopérabilité.

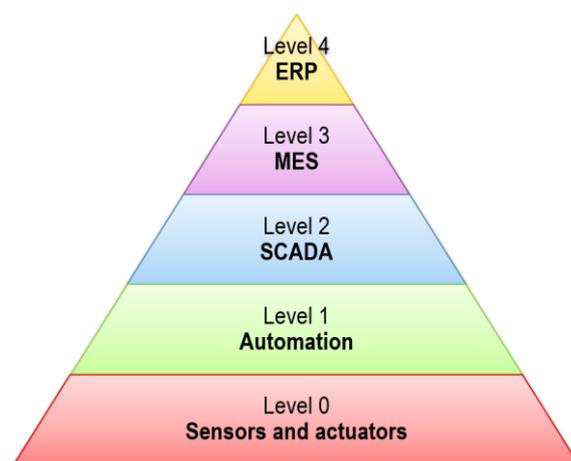


Figure 1. Pyramide de l'automatisation (ISA, 1995)

L'interopérabilité entre systèmes permet, à terme, une intégration horizontale, c'est-à-dire entre les différents systèmes d'une même couche de la pyramide de l'automatisation. Cela permet également une intégration verticale, permettant une remontée d'informations provenant des capteurs des systèmes des sites de production pour l'analyse, l'anticipation de pannes, mais également une descente des choix de production effectués en amont, afin de permettre un site de production agile.

Quatre types d'interopérabilités peuvent être observés : l'interopérabilité juridique, l'interopérabilité organisationnelle, l'interopérabilité technique et l'interopérabilité sémantique. (EIF, 2017)

L'interopérabilité juridique concerne le processus consistant à faire coopérer les règles juridiques entre les juridictions, à différents niveaux subsidiaires au sein d'un même État ou entre deux ou plusieurs États. La question de savoir si de nouvelles lois doivent être mises en œuvre ou des lois existantes ajustées ou réinterprétées pour réaliser cette interopérabilité dépend des circonstances données. L'interopérabilité juridique vise à garantir que des organisations opérant au sein de cadres juridiques, politiques et stratégies différentes puissent travailler ensemble.

(Scholl et al., 2011) définissent l'interopérabilité organisationnelle comme la capacité à définir les processus métiers et à apporter de la collaboration au sein des structures

qui souhaitent organiser l'échange d'information et, ce, malgré des différences d'environnements, et ainsi réconcilier les processus.

(Van Der Veer & Wiles, 2008) définissent l'interopérabilité technique comme les composantes logicielles ou matérielles, les systèmes et les plateformes qui permettent d'assurer la communication d'une machine à l'autre. L'interopérabilité technique est centrée sur les protocoles (de communication) et l'infrastructure nécessaire à ces protocoles. Nous pouvons ici citer deux types d'architectures d'intégration de systèmes à un réseau : l'architecture point à point, plus simple mais plus coûteuse, et l'architecture faisant intervenir un médiateur. (Pavot, 2010)

Au-delà des protocoles de communication, l'interopérabilité technique s'intéresse également à tout moyen qu'il est possible de mettre en œuvre afin de rendre deux entités interopérables, communicantes et permettant de s'intégrer facilement à un réseau de machines. Par exemple, (Vladimír et al., 2015) proposent l'utilisation de postes de travail modulaires pour l'assemblage par l'utilisation de blocs et accessoires afin de rendre ce poste de travail flexible.

(Özacar et al., 2011) définissent l'interopérabilité sémantique comme la capacité à interpréter automatiquement l'information échangée avec justesse et sens, dans le but de produire des résultats utiles aux yeux des utilisateurs finaux du système en question. La norme (ISO 14258 1998) définit trois façons de réaliser une interopérabilité sémantique :

- L'intégration : Revient à utiliser un standard commun pour tous les systèmes et composants. Pour revenir à l'analogie des langues, cela revient à imposer l'anglais comme unique langue parlée.
- L'unification : Un méta-modèle de données est utilisé entre les composants contenant des correspondances sémantiques.
- La fédération : Des modèles distincts sont utilisés, mais des outils sont utilisés afin d'apporter une correspondance sémantique entre les différents modèles.

Il va falloir intégrer l'ensemble de ces types d'interopérabilité pour une intégration verticale réalisable techniquement (grâce à une interopérabilité technique), efficace et sans perte de précision (grâce à une interopérabilité sémantique), et la plus optimisée possible selon le secteur d'activité (grâce à une interopérabilité organisationnelle).

Une intégration verticale, possible en développant ces différents niveaux d'interopérabilité, est nécessaire afin de rendre l'industrie 4.0 et les CPS plus productifs et plus économiques (Lu, 2017). A ce jour, plusieurs cadres et modèles ont été instaurés afin d'encourager cette interopérabilité.

3 MODELES ARCHITECTURAUX DE REFERENCE

Si l'on souhaite développer une intégration verticale, il est important de définir les acteurs qui sont mis en jeu, ainsi que les pratiques à adopter, afin d'avoir un cadre commun. Pour cela, il est possible de s'appuyer sur des architectures de référence. Une architecture de référence contient tout un ensemble de conseils et recommandations pour l'intégration de produits ou de services afin de permettre sa bonne implantation. Ces architectures de références aident à sélectionner, selon la situation et les acteurs mis en jeu, les standards adéquats à adopter. Par exemple, (Sowell, 2006) présentent un cadre d'architecture nommé C4ISR, développé par le département de la défense américaine, dont le but est d'assurer l'intégration et l'interopérabilité entre trois points de

vues : organisationnel, systémique et technique. L'EIF implémente une interopérabilité par le biais de dimensions organisationnelles, techniques et sémantiques pour les communications informatiques entre les administrations, entreprises et citoyens dans l'Union européenne (Synergy, 2005). L'AIF, présenté par (Berre et al., 2007), s'occupe des applications d'entreprise et des systèmes logiciels, et propose une intégration conceptuelle, une intégration applicative et une intégration technique.

Au sein de l'industrie 4.0, une telle architecture consiste en un guide de réalisation d'un système afin de pouvoir l'intégrer au mieux, dans un univers communiquant où la donnée transmise et perçue joue un rôle clé. Deux architectures permettent d'établir les bases de la forme que prendra l'information lors de cette intégration : IIRA et RAMI 4.0. Le but de ces architectures de références est de fournir une marche à suivre pour l'utilisation de standards relatifs au concept de « l'usine intelligente », applicables par la suite dans diverses entreprises appartenant à des secteurs différents.

3.1 IIRA

IIRA, ou Industrial Internet Reference Architecture, fut publié en 2015 par l'Industrial Internet Consortium Architecture Task Group. Le but était alors de fournir toutes les indications nécessaires afin de développer au mieux, et de façon standardisée, toutes architectures système axés IIoT afin qu'ils puissent respecter leurs promesses d'interopérabilité. Elle permet d'orienter dans le développement de ces systèmes, en y apportant une définition précise et commune, mais également un vocabulaire commun afin de pouvoir discuter des spécifications et implémentations.

IIRA a pu voir le jour grâce à l'Industrial Internet Architecture Framework (IIAF), un cadre d'architecture définissant des méthodes d'organisation sur le principe de points de vue différentes selon les différentes parties prenantes.

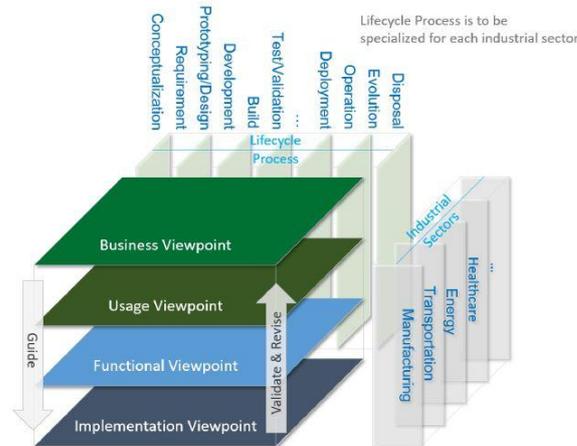


Figure 2. Représentation de l'architecture IIRA (Industrial Internet Consortium, 2019a)

Le principe central de cette architecture de référence est celui de points de vue. En effet, IIRA propose quatre points de vue définis en analysant les différents cas d'utilisation de l'IIoT, en identifiant les parties prenantes et en déterminant les principales préoccupations de ces parties prenantes.

Une dimension concerne le processus du cycle de vie, allant de la conception à la réalisation à la fabrication et à l'implémentation. Il ne s'agit pas ici d'une description du cycle de vie d'un système à proprement parler, mais une méthodologie permettant de mettre en lumière les différents problèmes du système qui peuvent affecter ce processus de cycle de vie.

Enfin, une dimension vise à définir les besoins selon différents secteurs industriels. L'Industrial Internet Consortium analyse également les retours provenant d'implémentations réelles au travers de ces différents secteurs industriels, afin de pouvoir juger de l'utilité de la version actuelle d'IIRA, et apporter les modifications qui paraissent nécessaires.

Cette architecture inclue des types de système, des configurations de système, des types de connexions, et des cas d'études provenant d'une variété d'industries différentes. Dans cette architecture de référence, les solutions proposées seront donc organisées selon différents points de vue, du plus abstrait (Business ViewPoint) au plus concret (Implementation ViewPoint). Une grande variété de secteurs industriels sont concernés par les solutions présentées par l'architecture, ainsi que l'entièreté du processus du cycle de vie, allant de la conceptualisation à la destruction, en passant par le déploiement. (Industrial Internet Consortium, 2019b)

3.2 RAMI 4.0

Le « Reference Architecture Model for Industry 4.0 » (RAMI 4.0) se concentre quant à lui davantage sur l'implémentation d'une stratégie Industrie 4.0 au travers de standards divers et de cas d'études. Tout comme l'IIRA, il possède une représentation en 3 dimensions, mais cette fois-ci les différentes couches sont plus détaillées dans leur représentation et partent d'un focus sur les actifs, c'est à dire une description des machines, données, câbles utilisés par exemple, pour ensuite dézoomer jusqu'à la couche « business », ou comment le composant industriel s'inscrit dans le business model choisi, dans les processus commerciaux. L'architecture RAMI 4.0 est une architecture orientée-service. Différents services sont mis en relations afin de répondre à la problématique du développement de systèmes orientés Industrie 4.0.

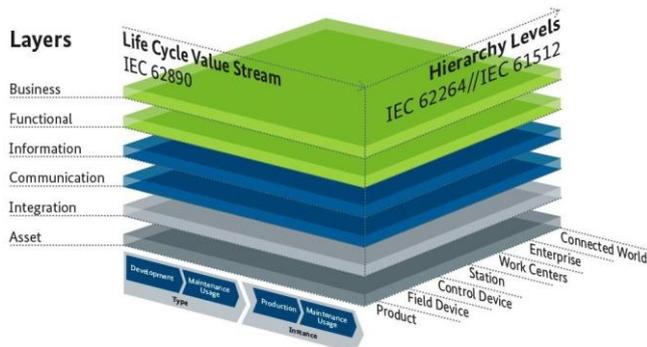


Figure 3. Représentation de l'architecture RAMI 4.0 (ZVEI, 2015)

Dans la dimension du cycle de vie du produit, différence est faite entre le type du produit et l'instance. Le type du produit se décide lors de la phase de développement. A la fin de cette phase, une version du produit est validée et la production est lancée. Ces produits ainsi fabriqués sont appelés les instances du type. L'amélioration des composants de ce produit peut requérir la modification de ce produit. Une boucle sera alors effectuée, repassant par une phase de développement puis une phase de production. Ainsi, cette architecture se préoccupe de la remontée d'informations et de la création de données dès la phase de conception.

La dimension concernant le niveau de hiérarchie décrit les différentes fonctionnalités au sein des sites de productions. Il est important de noter que, afin de répondre aux exigences de l'Industrie 4.0, de nouvelles fonctionnalités ont émergées

comme la catégorie « Product », ou encore la catégorie « Connected World » relative au domaine de l'IoT.

L'axe vertical, pour finir, décompose le composant industriel, ou la machine, couche après couche, allant du point de vue échange d'informations dans un contexte d'organisation business, jusqu'à l'intégration d'actifs (la couche assets).

3.3 Similitudes et complémentarités

RAMI 4.0 et IIRA furent développés afin de proposer un cadre pour le développement d'architectures pour la création de systèmes à l'ère de l'industrie 4.0. Pendant que IIRA se concentre sur la proposition de méthodes et modèles afin de concevoir des systèmes orientés IIoT, RAMI 4.0, quant à lui, se propose d'établir les bases de conceptions de standards respectant au mieux les exigences de cette quatrième révolution industrielle. Cela passe par exemple par une évaluation de l'intégration horizontale, entre machines, ou verticale, notamment à travers le Cloud. Avec cette architecture de référence, il est également possible d'évaluer les standards existant pour déterminer leur degré d'implication dans l'Industrie 4.0. Contrairement à RAMI 4.0, IIRA ne propose pas de recommandations précises quant aux standards qu'il est possible d'utiliser, et aux technologies disponibles. D'un autre côté, IIRA se focalise sur la conception de systèmes IIoT, en considérant qu'il s'agit là du concept qui transcendera l'industrie. IIRA élude, contrairement à RAMI 4.0, toute la partie fabrication.

Une solution fut proposée afin de joindre ces deux architectures de référence dans une publication intitulée « Architecture Alignment and Interoperability ». Cette publication vu le jour après que les créateurs de ces deux architectures de référence soient entrés en relation afin de comprendre les différences, similitudes et complémentarités entre RAMI 4.0 et IIRA (Lin et al., 2017).

Ces architectures de référence sont utiles lorsqu'il s'agit d'obtenir un guide à haut niveau de représentation. Elles fournissent des définitions communes et cohérentes pour le système d'intérêt, un vocabulaire commun avec lequel discuter (Industrial Internet Consortium, 2019b). Cependant, ces architectures manquent de flexibilités quant à la hiérarchie des systèmes (voir les 5 couches proposés par ISA 1995). IBM apporte une réponse avec la création de l'Industrie 4.0 Architecture décrivant l'architecture fonctionnelle d'un système de fabrication en deux couches, et proposant l'utilisation du cloud et du edge, c'est-à-dire des systèmes situés en périphérie permettant un traitement des données qui seront ensuite utilisées par l'entreprise (IBM, 2018)(Moghaddam et al., 2018)

Des architectures de référence sont d'ores et déjà proposées afin d'encourager, au sein des industries, une intégration horizontale, tout au long du cycle de vie, et verticale, entre les différentes couches de la pyramide de l'automatisation. Parmi les solutions, des standards permettent une interopérabilité sémantique entre certaines couches de cette pyramide

4 STANDARDISER LA COMMUNICATION

Le concept de standards afin d'unifier et de fournir une sorte de guide de bonne pratique n'est pas nouveau, tout comme le fait d'essayer d'unifier les systèmes au sein d'une entreprise. A titre informatif, dès le XIXème siècle, nous pouvons constater un très grand effort de normalisation et de standardisation dans le domaine industriel. Les standards simplifient le travail des parties prenantes en assurant une uniformité et en encourageant l'interopérabilité (Trappey et al., 2017).

Les architectures de référence, comme IIRA pour l'implémentation de systèmes IoT au sein de l'industrie, ou RAMI 4.0 pour l'accompagnement d'une transition vers l'Industrie 4.0, encouragent une interopérabilité organisationnelle. RAMI 4.0, par exemple, fournit tout un ensemble de bonne pratique selon le secteur d'activité de l'entreprise.

Afin d'assurer une interopérabilité entre systèmes de gestion et systèmes de production, il est également nécessaire d'assurer une interopérabilité à un niveau technique (lié à la normalisation des interfaces matérielles et logicielles) et sémantique (qui concerne la compréhension entre les différents acteurs). Pour ce faire, différents standards ont été mis en place afin de normaliser l'échange d'informations.

Concernant les systèmes d'information, ISA-95 définit les transactions d'informations entre systèmes de gestion de production et système de gestion de production. La norme MTConnect permet à un système de surveillance de collecter avec précision et cohérence les données de toute machine compatible MTConnect, quelle que soit sa marque ou son origine (Edrington et al., 2014). Elle établit un canal de communication pour une interconnectivité plug-and-play entre les dispositifs, les équipements et les systèmes, en fournissant des données en temps réel de toute l'usine. OPC UA offre un middleware pour intégrer des plates-formes hétérogènes.

OPC UA est, à l'heure actuelle, adopté par de plus en plus d'industriels et de consortiums afin de répandre la norme. OPC UA apporte à la fois une architecture de communication et une structuration de la donnée. Selon une étude comparative proposée par (Profanter et al., 2019) portant sur les caractéristiques et les performances de certains des protocoles les plus importants dans les domaines de l'automatisation industrielle et de l'IoT : OPC UA, ROS, DDS et MQTT. OPC UA apporte notamment une modélisation sémantique de l'information.

Tachons de voir maintenant comment une interopérabilité sémantique est possible, au travers deux standards : ISA95, standard permettant une caractérisation des flux d'informations entre les systèmes décisionnels et de production (entre les systèmes ERP et MES), mais également à l'origine de la pyramide d'automatisation (Fig. 1) et OPC-UA, standard définissant la remontée d'information depuis les capteurs, jusqu'au au Automates programmables, puis aux systèmes de gestion de production.

4.1 ISA95

Le standard ISA95 apporte des réponses quant au contrôle et à la gestion de la production. Tout d'abord, il serait plus exact ici de parler de la norme ISA95, étant reconnue internationalement en tant que norme IEC62264. Cette norme apporte des réponses quant au contrôle et à la gestion de la production. Elle passe notamment par une description terminologique de l'environnement et du monde de l'entreprise, ainsi que la description des relations gestion/production. Il s'agit, en définissant l'ensemble des flux d'information ainsi que leur structuration, d'encourager l'interopérabilité entre les systèmes décisionnels et opérationnels. Les possibilités d'intégration sont accrues, tandis que les risques et possibilités d'erreurs sont diminués.

ISA95 se décompose en 5 parties principales d'ores-et-déjà rédigées (d'autres ont vu le jour depuis, mais l'accent est mis ici sur les 5 parties initiales)(Schloten, 2007) :

- La première va s'attarder sur l'établissement de modèles et terminologies afin d'aider au mieux la préparation et le montage de projet où l'automatisation de l'information

entre la couche ERP et MES sera la problématique principale.

- La partie 2 proposera, pour chaque objet présent dans les modèles de la partie 1, une liste d'attributs permettant de décrire ces objets.
- La partie 3, quant à elle, ne concerne que la couche MES, et propose ainsi des modèles et terminologies afin d'analyser et décrire les activités au sein de cette couche. Cette partie se veut utile pour améliorer les systèmes de productions se trouvant déjà sur site.
- La partie 4 va compléter cette dernière en apportant des modèles de données afin de décrire les flux d'information au sein du niveau 3 (Entre systèmes MES, systèmes de gestion d'entrepôt, ou encore systèmes de maintenance), mais également entre les différents systèmes du niveau 3 et les systèmes ERP.
- La partie 5, pour finir, complète les deux premières parties. Tandis que les parties 1 et 2 définissaient clairement quelles informations allait être échangées entre les couches MES et ERP, ainsi que leur structure, la partie 5 se concentre sur la définition des transactions, répondant à la question : quoi faire de la donnée envoyée et reçue ?

4.2 OPC-UA

OPC-UA (Open Platform Communication Unified Architecture) est un standard d'échange de données développé par la Fondation OPC et dont la première version parût en 2006. OPC-UA est considéré comme une technologie de communication Machine to Machine permettant de connecter des capteurs entre eux, mais également les capteurs aux systèmes de supervision, permettant une automatisation à la fois simplifiée et solide (Mathias et al., 2020).

OPC UA fournit à la fois un protocole de communication et une méthode de modélisation des informations. A l'ère de l'industrie 4.0, OPC-UA semble être le seul protocole de communication à être recommandé par l'architecture RAMI 4.0, architecture synthétisant les orientations et standardisations nécessaires à l'Industrie 4.0, apportant à la fois un protocole de communication et une standardisation de la structuration de la donnée (Liu et al., 2019).

Le protocole OPC-UA peut fournir une communication de type « report by exception », c'est-à-dire que la donnée sera envoyée dans le réseau qu'à partir du moment où celle-ci se sera vu modifiée (The OPC Foundation, n.d.-b)(Lehnhoff et al., 2012). OPC-UA est donc léger en termes de consommation de bande passante. Malheureusement, malgré le fait que OPC-UA se veuille être ouvert, il n'en reste pas moins un standard « pay-to-play » (The OPC Foundation, n.d.-a). Ainsi, même si il est effectivement disponible pour tous ceux qui voudrait l'adopter, le fait de devoir payer pour rejoindre la communauté OPC-UA remet en question ce côté ouvert.

OPC-UA fonctionne sur le modèle client/serveur. Ainsi un serveur sera relié aux différents appareils dont on souhaite récolter des données. Ce serveur jouera le rôle de banque de données et de traducteur entre les différentes machines, PLC du réseau. Dès lors, n'importe quel OPC client pourra entrer en contact avec le serveur, via une communication en OPC-UA, afin de récolter les informations voulues, et envoyer des ordres aux différents PLC. Pour ce qui est de l'intégration verticale, c'est-à-dire entre le serveur OPC et les outils SCADA, MES, ERP, un bloc de traduction sera nécessaire, ces outils ne pouvant que rarement dialoguer au standard OPC-UA. Un exemple de traducteur existant s'appelle Iconics, un logiciel permettant de visualiser, archiver et analyser la donnée en

temps réel provenant des différents équipements de l'usine. (Iconics, n.d.)

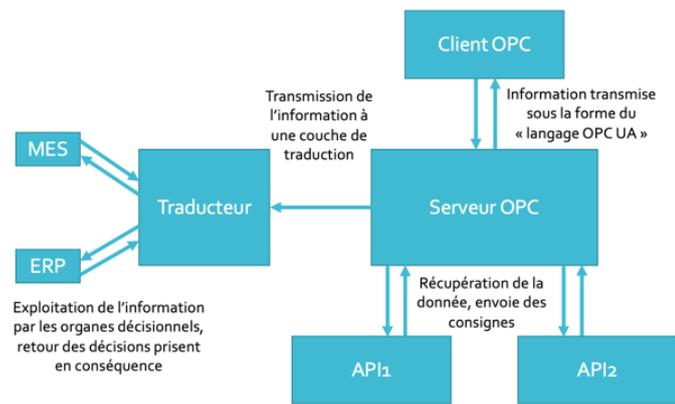


Figure 4. Échanges d'informations entre les différents acteurs d'une architecture OPC-UA

Une des forces d'OPC-UA réside dans son espace d'adressage. Celui-ci est en effet configurable afin de respecter le caractère Orienté Objet de OPC-UA. Par conséquent, il est possible de créer des relations complexes au sein même de la donnée, et il est ainsi possible de suivre des structures d'informations précises comme celle présentée par ISA95.

Concernant les applications d'OPC-UA, (Muller et al., 2017) a présenté une mise en œuvre libre et gratuite de l'OPC UA. Un serveur OPC UA personnalisable développé sur une carte Arduino permettait de contrôler la température de la buse d'une imprimante 3D en boucle fermée. (Cavaliere et al., 2019) présente plusieurs approches visant à introduire des améliorations en matière de TIC dans l'OPC UA afin d'améliorer sa convivialité dans les environnements industriels. OPC-UA permet l'appréhension d'une interopérabilité à la fois technique et sémantique. OPC-UA propose ainsi une réelle structuration de la donnée. Seulement, nous ne disposons ici que des moyens de structuration (avec notamment l'adressage permettant d'identifier la donnée), et il va falloir ensuite définir qu'elle axe prendre quant à la manière dont nous souhaitons la structurer. Le concept d'ontologie apporte une réponse.

OPC-UA et ISA95 sont complémentaires, et permettent, une fois adoptés, une communication plus facile, notamment, par exemple, grâce à la possibilité d'intégrer à un réseau de machines fonctionnant sous OPC-UA n'importe quelle machine ayant adopté le standard. Ces standards sont pour autant mal intégrés. Afin de permettre une interopérabilité entre ces standards, des ontologies pourraient être utilisées, pour pouvoir effectuer une réconciliation entre elles, principe qui sera détaillé dans la partie qui suit. Les ontologies offrent des représentations plus riches de la sémantique interprétable par les machines pour les systèmes et les bases de données, assurant ainsi une plus grande interopérabilité et intégration sémantique (Obst et al., 2007).

5 DES ONTOLOGIES POUR SE COMPRENDRE

5.1 Ontologies

Nous venons de voir en quoi les standards pouvaient permettre de structurer et conditionner la conception des systèmes, tout cela dans le but de structurer un ensemble de systèmes capable de communiquer entre eux.

Une question cruciale se posant lorsque l'on souhaite développer une interopérabilité verticale au sein d'une entreprise vient de la grande quantité de données qui sera générée lors de la production. Ces données sont bien souvent hétérogènes, se situant soit au même niveau (données de provenance de différents types de machines, ou de machines de différents fabricants), soit à un niveau différent (les informations générées par les systèmes de gestion n'auront par exemple pas la même forme que les données générées par la machine).

Une ontologie est définie comme la spécification formelle de conceptualisations partagées, contenant des concepts, des relations, des instances et des axiomes (Gruber, 1993)(Burns et al., 2019). Les ontologies, grâce à cet aspect formel, sont compréhensibles par un ordinateur, et donc exploitable par n'importe quel système informatique. Ce concept s'inscrit dans la volonté d'apporter au sein de l'entreprise une interopérabilité sémantique, et ainsi permettre l'échange de données sans qu'il y ait d'ambiguïté. Une ontologie va donc tenter de représenter, à l'aide de classes et de relations, toutes les relations qui pourraient exister au sein des données stockées. Les concepts seront représentés sous forme de classes auxquelles seront rattachées toutes les données relatives à cette classe, et ces classes seront reliées entre elles par des relations. Il est de ce fait par exemple possible de lier des données de provenance diverses sous un même concept. Une ontologie, c'est un modèle conceptuel consensuel et partageable d'un domaine (Gruber, 1993). Cela veut dire qu'au-delà d'un modèle de données, une Ontologie va s'intéresser au contexte et à la façon dont celui-ci est organisé. Dans le domaine de l'industrie, plusieurs ontologies peuvent ainsi être définies en rapport aux sous-domaines mentionnés ci-après : Les ontologies de processus, les ontologies de ressources et les ontologies de situation. Les Ontologies de ressources décrivent aussi bien les ressources humaines que les objets physiques. Au-delà sont définies des Ontologies de cœurs (*Core Ontology*). Les ontologies de cœur sont des conceptualisations universelles et indépendantes du domaine d'étude pouvant être réutilisées dans des domaines de connaissance variés. (Giustozzi et al., 2018)

Les ontologies de processus sont utilisées pour une représentation globale du contexte de l'entreprise, d'où le nom d'ontologies de contexte (*Context Ontology*). Nous pouvons ainsi considérer chaque élément du schéma de la figure 3 comme de petites ontologies faisant partie d'ontologies plus grandes, de grands groupes. La question va donc être la suivante : Comment créer des liens entre ces ontologies ? Il est fort possible que des similitudes apparaissent, et pouvoir comparer les ontologies pour relever ces similitudes et ainsi les ordonner serait d'un intérêt crucial pour une exploitation ultérieure de la donnée.

Concernant les ontologies tirées de standards, (Merdan et al., 2010) présentent un système de désassemblage automatisé basé sur une ontologie qui intègre la commande de machines distribuées à base d'agents multiples ainsi que la planification de trajectoire en temps réel basée sur la vision et l'automatisation des robots. (Seyedamir et al., 2018) propose le développement d'ontologies modulaires basées sur différentes parties de la norme ISA-95 pour décrire les informations sur les produits, les processus et les ressources des systèmes de fabrication. (Katti et al., 2020) proposent une automatisation de l'encodage d'une ontologie de production basée sur le code source du MES. À cette fin, les connaissances des collaborations OPC UA sont également exploitées lors de la création d'ontologies de ressources.

Afin d'obtenir une intégration verticale complète, il faut réussir à aligner ces différentes ontologies, créer des liens entre par exemple une ontologie de gestion, et une ontologie plus bas niveau, caractérisant des échanges d'informations entre système de production.

5.2 Réconciliation sémantique

L'utilisation des ontologies permet de formaliser la donnée concernant la conception ou la fabrication d'un objet. Les ontologies dans le domaine de l'industrie sont nombreuses, nous pouvons citer ici l'ontologie de modélisation des tolérances, l'ontologie de modélisation de la fabrication, ou encore l'ontologie de modélisation de conception. Toutes ces ontologies représentent ce que l'on appelle les ontologies de référence, et représentent la façon de modéliser la donnée, afin de représenter différentes perspectives du produit concerné. (Anderson Luis Szejka & Junior, 2017). Grâce à ces ontologies de référence, la donnée provenant de différents domaines concernant un même produit est structurée. Seulement, afin d'assurer une exploitation totale de cette donnée durant les différentes phases de conception et fabrication, il va être nécessaire de relier ces différentes ontologies. Pour ce faire, il va être possible d'effectuer une réconciliation sémantique. Il va s'agir de créer des règles sémantiques concernant les informations provenant des différentes ontologies devant être liées, et prenant en compte également les contraintes du produit ainsi que les informations renseignées par les différentes personnes chargées de la conception et de la fabrication. La création de ces règles va ainsi permettre par exemple d'alimenter, comme proposé par (Anderson Luis Szejka & Junior, 2017), une application de conception et de fabrication, stockant toute la donnée relative à un produit de façon structurée. Cette réconciliation s'effectue en 3 phases (Anderson L. Szejka et al., 2017) :

- L'ajustement du contexte, étape durant laquelle les concepts seront alignés selon un domaine spécifique, le contexte pourra notamment être aligné selon le produit à développer.
- L'intersection des ontologies : Première véritable étape de la réconciliation sémantique durant laquelle les ontologies seront comparées afin de définir des « nœuds communs » selon un contexte donné.
- L'alignement sémantique : Ici, des règles seront créées afin d'établir des liens entre les informations de différentes ontologies. Une cartographie sera ainsi mise en place, fonctionnant sur le principe d'équations mathématiques reliant ces informations.

(Chen et al., 2009) proposent une méthodologie qui intègre la connaissance du cycle de vie des produits basée sur l'ontologie, permettant à tous les participants de partager de la connaissance. (Adamczyk et al., 2020) proposent un système pour assurer une interopérabilité sémantique reposant sur l'utilisation des ontologies : le SISMS. Toutes les informations et connaissances essentielles sur le processus ont été modélisées dans des ontologies de référence et, plus tard, les ontologies de référence ont été complétées avec des données du cas réel.

Les ontologies vont donc permettre de caractériser la structuration des données de différents standards, en créant des liens entre les différents objets, les différentes classes constituant ces données. Une réconciliation sémantique sera effectuée, afin de faire le lien, trouver des similarités entre les données des différents standards, des différentes ontologies.

Cela permet, à défaut d'utiliser un standard commun, d'englober différents standards afin de permettre une compréhension entres systèmes.

6 CONCLUSION ET PERSPECTIVES

A travers cet article, nous avons vu en quoi les standards actuels pouvaient permettre une intégration verticale au sein de l'entreprise. Une interopérabilité à la fois sémantique, technique et organisationnelle est ainsi proposée, au travers notamment l'établissement de modèles représentant les échanges d'informations MES-ERP. Les standards, comme ISA95 permettant une unification de la communication MES-ERP, furent décriés. Une interopérabilité organisationnelle est encouragée par les architectures de référence proposant une architecture par secteurs d'activités notamment pour ce qui est du RAMI 4.0, ou tout au long du cycle de vie d'un produit. Une interopérabilité technique comme sémantique est proposée au travers l'utilisation des standards ISA95 et OPC UA. Nous avons également vu en quoi les ontologies pouvaient permettre une réconciliation entre ces différents standards. L'ensemble des solutions abordées dans cet article permettent la conception de systèmes qui s'intégreront et pourront échanger de l'information verticalement. Une remontée d'information est possible jusqu'aux systèmes de gestion de l'entreprise, et des décisions peuvent être prises afin de piloter les systèmes de production. Cependant, une intégration efficace entre les différents standards couvrant l'ensemble des systèmes de la pyramide de l'automatisation n'est pas encore d'actualité. Dans un premier temps, un alignement (entre ISA95 et OPC-UA par exemple) doit permettre de faire le lien, de donner un sens plus grand à la donnée (comme les objets de OPC UA). Dans un second temps, la création d'un modèle représentant l'ensemble des informations échangés, et reposant sur les différents standards, permettra une interopérabilité verticale complète. Il serait également intéressant, dans une prochaine étude, de voir en quoi ces solutions pourraient être efficace dans un environnement ou la hiérarchie proposé par ISA n'est plus aussi distincte, où la prise de décision décentralisée est de mise (avec l'utilisation des systèmes cyber-physiques), et vers lequel l'Industrie 4.0 nous amène.

Une communication verticale est donc possible, mais face au nombre grandissant de systèmes communiquant au sein de l'entreprise, une communication en temps réel parait de plus en plus difficile (problème de bande passante, de capacité de traitement de l'information, ...). Le Cloud Manufacturing se veut une réponse efficace, permettant la centralisation des informations, et son encapsulation au sein de services (Xu, 2012). Le Fog Computing semble aussi être une solution afin de faciliter le pilotage en temps réel des systèmes au sein d'une entreprise. Le Fog Computing permet ainsi un prétraitement des informations issues de ces systèmes, afin d'optimiser la communication avec le cloud. Il pourrait permettre une forte convergence, unification ou standardisation d'un point de vue réseau, sécurité, données, traitement de données, et contrôle. Cela mènera à améliorer l'interopérabilité, la sécurité, permettra un contrôle plus riche et efficace, et au-delà une fabrication plus efficace et flexible (Pop et al., 2018).

7 REFERENCES

- Adamczyk, B. S., Szejka, A. L., & Canciglieri, O. (2020). Knowledge-based expert system to support the semantic interoperability in smart manufacturing. *Computers in Industry*, 115(xxxx). <https://doi.org/10.1016/j.compind.2019.103161>
- Berre, A.-J., Elvesæter, B., Figay, N., Guglielmina, C., Johnsen, S.

- G., Karlsen, D., Knothe, T., & Lippe, S. (2007). The ATHENA interoperability framework. *Enterprise Interoperability II*, 560–580.
- Burns, T., Cosgrove, J., & Doyle, F. (2019). A Review of Interoperability Standards for Industry 4.0. *Procedia Manufacturing*, 38(2019), 646–653.
- Cavaliere, S., Salafia, M. G., & Scropo, M. S. (2019). Integrating OPC UA with web technologies to enhance interoperability. *Computer Standards and Interfaces*, 61(August 2017), 45–64.
- Chen, Y. J., Chen, Y. M., & Chu, H. C. (2009). Development of a mechanism for ontology-based product lifecycle knowledge integration. *Expert Systems with Applications*, 36(2 PART 2), 2759–2779.
- Edrington, B., Zhao, B., Hansel, A., Mori, M., & Fujishima, M. (2014). Machine monitoring system based on MTConnect technology. *Procedia CIRP*, 22(1), 92–97.
- EIF. (2017). New European Interoperability Framework. In *White Pages*. <https://doi.org/10.2799/78681>
- Giustozzi, F., Saunier, J., & Zanni-Merk, C. (2018). Context Modeling for Industry 4.0: An Ontology-Based Proposal. *Procedia Computer Science*, 126, 675–684.
- Goncalves, R., Agostinho, C., Sarraipa, J., De, A. R., Togores, A. R., De, Jardim-goncalves, R., & Sarraipa, J. (2011). *Standards Framework for Intelligent Manufacturing Systems Supply Chain*.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199–220.
- IBM. (2018). *Industrie 4.0 Architecture*.
- Iconics. (n.d.). <https://iconics.com>
- IEEE. (1990). *IEEE (Institute of Electrical and Electronics Engineers) Standard Computer Dictionary - A Compilation of IEEE Standard Computer Glossaries*.
- Industrial Internet Consortium. (2019a). *INDUSTRIAL INTERNET REFERENCE ARCHITECTURE v1.9*.
- Industrial Internet Consortium. (2019b). *The Industrial Internet of Things Volume G1: Reference Architecture*.
- Katti, B., Plociennik, C., Ruskowski, M., & Schweitzer, M. (2020). Bidirectional transformation of MES source code and ontologies. *Procedia Manufacturing*, 42, 197–204.
- Lehnhoff, S., Rohjans, S., Uslar, M., & Mahnke, W. (2012). OPC unified architecture: A service-oriented architecture for smart grids. *2012 1st International Workshop on Software Engineering Challenges for the Smart Grid, SE-SmartGrids 2012 - Proceedings*, 1–7.
- Lin, S.-W., Murphy, B., Clauser, E., Loewen, U., Neubert, R., Bachmann, G., Pai, M., & Hankel, M. (2017). Architecture Alignment and Interoperability: An Industrial Internet Consortium and Plattform Industrie 4.0 Joint Whitepaper. *Plattform Industrie 4.0*, 19.
- Liu, C., Vengayil, H., Lu, Y., & Xu, X. (2019). A Cyber-Physical Machine Tools Platform using OPC UA and MTConnect. *Journal of Manufacturing Systems*, 51(April), 61–74.
- Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration*, 6, 1–10.
- Mathias, S. G., Schmied, S., & Grossmann, D. (2020). An Investigation on Database Connections in OPC UA Applications. *Procedia Computer Science*, 170(2019), 602–609.
- Merdan, M., Lepuschitz, W., Meurer, T., & Vincze, M. (2010). Towards ontology-based automated disassembly systems. *IECON Proceedings (Industrial Electronics Conference)*, 1392–1397.
- Moghaddam, M., Cadavid, M. N., Kenley, C. R., & Deshmukh, A. V. (2018). Reference architectures for smart manufacturing: A critical review. *Journal of Manufacturing Systems*, 49(June), 215–225.
- Muller, M., Wings, E., & Bergmann, L. (2017). Developing open source cyber-physical systems for service-oriented architectures using OPC UA. *Proceedings - 2017 IEEE 15th International Conference on Industrial Informatics, INDIN 2017*, 83–88. Newman, S. T., Nassehi, A., Xu, X. W., Rosso, R. S. U., Wang, L., Yusof, Y., Ali, L., Liu, R., Zheng, L. Y., Kumar, S., Vichare, P., & Dhokia, V. (2008). Strategic advantages of interoperability for global manufacturing using CNC technology. *Robotics and Computer-Integrated Manufacturing*, 24(6), 699–708. Obrst, L., Ceusters, W., Mani, I., Ray, S., & Smith, B. (2007). The Evaluation of Ontologies. *Semantic Web*, 139–158.
- Ostermeyer, E., Danjou, C., Durupt, A., & Le Duigou, J. (2019). Une méthodologie d'apprentissage automatique pour l'aide à la décision en contexte d'industrialisation. *Génie Industriel et Productique (2019) 2(2)*.
- Özacar, T., Öztürk, Ö., & Ünalir, M. O. (2011). ANEMONE: An environment for modular ontology development. *Data and Knowledge Engineering*, 70(6), 504–526. Paviot, T. (2010). *Méthodologie de résolution des problèmes d'interopérabilité dans le domaine du Product Lifecycle Management*.
- Pop, P., Raagaard, M. L., Gutierrez, M., & Steiner, W. (2018). Enabling Fog Computing for Industrial Automation Through Time-Sensitive Networking (TSN). *IEEE Communications Standards Magazine*, 2(2), 55–61.
- Profanter, S., Tekat, A., Dorofeev, K., Rickert, M., & Knoll, A. (2019). OPC UA versus ROS, DDS, and MQTT: Performance evaluation of industry 4.0 protocols. *Proceedings of the IEEE International Conference on Industrial Technology, 2019-Febru*, 955–962.
- Schloten, B. (2007). *The Road to Integration: A Guide to Applying the ISA-95 Standard in Manufacturing* (ISA (Ed.)). h
- Scholl, H. J., Kubicek, H., & Cimander, R. (2011). *Interoperability, Enterprise Architectures, and IT Governance in Government*. 345–354.
- Seyedamir, A., Ferrer, B. R., & Lastra, J. L. M. (2018). An ISA-95 based Ontology for Manufacturing Systems Knowledge Description Extended with Semantic Rules. *Proceedings - IEEE 16th International Conference on Industrial Informatics, INDIN 2018*, 374–380.
- Sowell, P. K. (2006). The C4ISR architecture framework: history, status, and plans for evolution. *The MITRE Corporation*.
- Synergy. (2005). European interoperability framework v 1.0. *The IDABC Q, 1*.
- Szejka, Anderson L., Junior, O. C., Panetto, H., Aubry, A., & Loures, E. F. R. (2017). A semantic reconciliation view to support the interoperable information relationships in product design and manufacturing. *IFAC-PapersOnLine*, 50(1), 15896–15903.
- Szejka, Anderson Luis, & Junior, O. C. (2017). The Application of Reference Ontologies for Semantic Interoperability in an Integrated Product Development Process in Smart Factories. *Procedia Manufacturing*, 11(June), 1375–1384. The OPC Foundation. (n.d.-a). *OPC Foundation : Become a member*. The OPC Foundation. (n.d.-b). *OPC Unified Architecture*. Trappey, A. J. C., Trappey, C. V., Hareesh Govindarajan, U., Chuang, A. C., & Sun, J. J. (2017). A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. *Advanced Engineering Informatics*, 33, 208–229. Van Der Veer, H., & Wiles, A. (2008). Achieving Technical Interoperability: the ETSI Approach. *European Telecommunications Standards Institute*, 3, 29. Vladimír, R., Andrea, L., & Norbert, S. (2015). Reconfiguring of Manual Workstations Designated for Customized Production. *American Journal of Mechanical Engineering*, 3(6), 257–260.
- Xu, X. (2012). From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, 28(1), 75–86.
- ZVEI. (2015). *Reference Architecture Model Industry 4.0*. https://www.zvei.org/fileadmin/user_upload/Themen/Industrie_4.0/Das_Referenzarchitekturmodell_RAMI_4.0_und_die_Industrie_4.0-

Un cadre modulaire d'utilisation de Jumeaux Numériques dans un Système Manufacturier Reconfigurable

HICHEM HADDOU BENDERBAL¹, ABDELKRIM R. YELLES-CHAOUCHE^{1,2}, ALEXANDRE DOLGUI¹

¹ IMT Atlantique, LS2N-CNRS,
Nantes, France

{hichem.haddou-ben-derbal, alexandre.dolgui}@imt-atlantique.fr

² IRT Jules Verne,
Bouguenais, France

abdelkrim-ramzi.yelles-chaouche@imt-atlantique.fr

Résumé – L'émergence de l'industrie 4.0 et de ses technologies associées, a transformé l'environnement de production en le rendant plus intelligent. Ceci est associé à l'évolution rapide des technologies d'acquisition de données et à l'énorme quantité de données générées. Parmi ces environnements de production modernes, le système manufacturier reconfigurable (RMS) est un concept capable de faire face aux conditions actuelles du marché, caractérisé par une demande de plus en plus personnalisée et volatile. Dans le même temps, le jumeau numérique (Digital Twin (DT)) est apparu comme un nouveau concept. DT représente une nouvelle vision axée sur les données (data-driven), qui combine l'analyse, l'optimisation et la simulation de systèmes en temps réel. Lors de la gestion des systèmes manufacturiers complexes, DT fournit de nouvelles perspectives et de nouveaux potentiels en matière d'aide aux processus décisionnels. Dans ce contexte, cet article tente de présenter un cadre modulaire intégrant RMS et Digital Twin (RMS-DT). RMS-DT est un modèle qui peut représenter l'état du système à tout moment, tout en permettant sa visibilité globale pour améliorer ses performances et permettre une prise de décision flexible. L'article se conclut par une discussion, les défis futurs et les perspectives afin d'améliorer le cadre proposé du RMS-DT.

Abstract – The emergence of Industry 4.0 and its related technologies transformed manufacturing environment by making them more intelligent and flexible. This is associated with the fast evolution of data acquisition technologies and the enormous amount of generated data. Among these modern manufacturing environment, Reconfigurable Manufacturing System (RMS) is a concept able to cope with the current market conditions, characterized by an increasingly personalized and volatile demand. At the same time, Digital Twin (DT) emerged as a new concept. DT represents a new data-driven vision that combines real time data analytics, optimization and simulation. When managing modern and complex manufacturing systems, DT provides new insights and potentials in decision-making process. In this context, this paper is an attempt to present an integrated RMS digital twin (RMS-DT) modular framework. RMS-DT is a model that can represent the system state at any moment in time while allowing a holistic system visibility to improve its performances and enable flexible decision-making. The paper is concluded with a discussion, future challenges and perspectives in order to enhance the proposed RMS digital twin framework.

Mots clés – Jumeaux numériques, systèmes manufacturiers reconfigurables, cadre modulaire, jumeau numérique RMS.

Keywords – Digital Twin, Reconfigurable Manufacturing System, Modular Framework, RMS digital twin.

1 INTRODUCTION

Les marchés actuels se caractérisent par un lancement fréquent des nouveaux produits, une demande de plus en plus personnalisée et volatile. En conséquence, la survie des entreprises dépend de leur reconfigurabilité, leur adaptabilité et flexibilité, leur capacité de répondre à l'exigence de personnalisation. Ainsi, de nombreuses recherches tentent d'apporter de nouvelles solutions pour répondre à ces défis. Dans ce contexte, le système manufacturier reconfigurable (RMS) est considéré comme l'une des solutions parmi les plus adaptées. Ce concept a été proposé par Koren (Koren *et al.*, 1999). Le lecteur trouvera une description des différences de ce système par rapport à tous les autres concepts les plus connus des systèmes manufacturiers dans (Dolgui et Proth, 2010). RMS se caractérise par la reconfigurabilité *c'est à dire* la capacité de remplacer ou de réorienter rapidement les composants (matériel et/ou logiciel) pour ajuster la production

(en termes de fonctionnalité et de capacité) afin de répondre à des changements demandés. Ceci est possible grâce à la structure modulaire et aux machines-outils reconfigurables (RMTs) qui sont des composantes majeures de RMS. Ces machines, elles aussi, ont une structure modulaire permettant d'être reconfigurées, afin de remplir différentes fonctionnalités et avoir différentes capacités.

RMS fait partie de l'industrie 4.0 (Bortolini *et al.*, 2018). Cette dernière est tirée par l'évolution rapide des technologies de l'information, de l'acquisition de données, l'intelligence artificielle (IA) et de la communication, qui sont le fondement principal de l'usine intelligente. On peut observer que l'IA est largement utilisée, mais son potentielle n'est pas encore pleinement exploité dans le cadre de l'industrie. Ceci s'explique par le fait que les modèles commerciaux et les besoins des fabricants sont très différents. De plus, ces derniers s'adressent aux petits marchés avec des produits dédiés, qui

nécessitent des outils spécialisés pour l'analyse prédictive et prescriptive dans un secteur donné. Néanmoins, le potentiel de l'IA est clair : « 50% des entreprises qui adopteront l'IA au cours des cinq à sept prochaines années peuvent doubler leur trésorerie » (Microsoft, 2019). L'une des étapes primordiales pour réussir de passer à l'échelle dans l'utilisation de l'IA est la numérisation (Digitalisation), où tous les composants d'une usine sont liés et peuvent communiquer en temps réel.

Dans ce contexte, les jumeaux numériques (DT) sont apparus comme un concept qui profite de ces technologies pour réaliser une production intelligente (Zheng *et al.*, 2019). En effet, DT fournit une représentation virtuelle haute-fidélité d'un système physique. Il comprend deux parties principales (physique et virtuelle) avec une transmission de données réelles entre elles. DT traitera, analysera et évaluera les données collectées (en temps réel et hors ligne). En conséquence, un large éventail d'informations peut être collecté et utilisé pour de nombreuses applications, telles que le suivi de l'état du système, la réalisation de prévisions, le diagnostic, la simulation et l'optimisation du système (planification, ordonnancement, sélection de configuration/reconfiguration, etc.).

Dans un environnement reconfigurable, où les changements de la demande mènent à ce que la flexibilité est un défi majeur, un DT peut s'avérer très utile dans un processus de prise de décision à tous les niveaux. En fait, la conception et la sélection d'une configuration RMS se révèlent complexes (Koren *et al.*, 2018), car le nombre de configurations augmente considérablement avec le nombre de d'équipements. D'autre part, afin d'anticiper l'adaptation des systèmes à de futurs scénarios de production en minimisant les besoins de remplacement de modules et en maximisant leur réutilisation, il faut considérer tout le cycle de vie des systèmes de production (Figge *et al.*, 2018). Ainsi, avec sa représentation virtuelle haute-fidélité de flux de données temps réel et du système lui-même, un DT peut aider à réaliser un processus de prise de décision flexible (par exemple en choisissant une configuration appropriée basée sur des données obtenues en temps réel, prédites ou historiques).

Les applications DT sont diverses et se sont avérées très utiles dans le contexte de production. Cependant, peu de travaux de recherche considèrent à la fois RMS et DT (Zheng *et al.*, 2019). De plus, à notre connaissance, la partie reconfiguration logicielle de RMS est rarement traitée par les chercheurs. Par conséquent, cet article tente de présenter un cadre générique et modulaire qui intègre à la fois les concepts RMS et DT. Ce dernier inclut la partie reconfiguration logicielle via un logiciel de type « Plug & Play ». De plus, l'idée est de montrer l'écart considérable entre les besoins et l'état de l'art en la matière et d'attirer l'attention sur l'importance de la recherche autour de ce sujet. Nous notons que l'idée d'un cadre modulaire d'utilisation de DT pour RMS a déjà été suggérée dans notre article (Haddou-Benderbal *et al.*, 2020), le présent article est une nouvelle version, revue, revisitée et étendue de présentation de cette idée où nous avons pris plus de soin pour présenter l'état de l'art et de discuter des applications possibles et des limites.

2 FONDEMENTS THEORIQUES

La littérature autour des RMS couvre divers domaines tels que la conception du système, la planification des processus, la planification et le contrôle reconfigurable, comme l'ont démontré plusieurs articles de l'état de l'art ces dernières années (Benyoucef *et al.*, 2020 ; Bortolini *et al.*, 2018 ; Brahimi *et al.*, 2019 ; Yelles-Chaouche *et al.*, 2020). De plus, l'Industrie 4.0 sous le focus de RMS couvre une multitude de nouveaux concepts et technologies émergents (Oztemel, 2020). Dans cette section on se focalisera sur une petite partie de ce domaine qui inclut respectivement : la conception de RMS, la numérisation (jumeau numérique) et la façon dont cette dernière est intégrée dans RMS.

L'un des principaux problèmes relevés dans la littérature sur les RMS est le problème de leur conception. L'objectif est d'utiliser leur caractéristique fondamentale qui est la reconfigurabilité pour concevoir des systèmes réactifs. Les problèmes de conception de RMS sont complexes, ceci est dû au fait que ces systèmes sont caractérisés par des capacités dynamiques et la nécessité d'intégrer le changement de capacité dans le processus de conception. En outre, les propriétés de RMS, y compris la reconfigurabilité, se révèlent après le déploiement du système de production (Andersen, 2017 ; Farid, 2016). RMS vise l'évolutivité (scalability) et la réactivité, soutenues par l'ajout ou la suppression de fonctionnalités. Cela permet à ces systèmes d'être facilement adaptables lors du cycle de vie du système de production. En plus d'être considérés comme une nouvelle forme de flexibilité, les RMS peuvent aussi être considérés comme une forme d'organisation de la production (Battaïa *et al.*, 2020). Par conséquent, les approches développées pour concevoir les systèmes de production conventionnelle doivent être revues, adaptées et/ou remplacées pour répondre aux nouveaux défis posés par les systèmes de fabrication reconfigurables (Koren *et al.*, 2018 ; Bortolini, 2018). Ainsi, pour répondre à ces exigences avec ces propriétés, la reconfigurabilité doit être intégrée dès le départ dans la conception du système afin d'avoir la meilleure réactivité face aux changements (Koren *et al.*, 2018). Par conséquent, tous les composants du système ainsi que ses différents niveaux doivent être pris en compte et préparés aux changements. Ces qualités sont favorisées par la reconfigurabilité de RMS (Elmaraghy *et al.*, 2009).

La reconfigurabilité représente une exigence non fonctionnelle du système, liée à son comportement à long terme (Andersen, 2017). Cela implique que les approches classiques, qui ne prennent en compte que les exigences immédiates du système, ne conduiront pas nécessairement à des systèmes reconfigurable/changeable dynamiquement. En outre, au lieu d'atteindre les objectifs de reconfigurabilité, ils font que le système conçu devient rapidement obsolète face aux variations du marché. Pour qu'un système soit facilement reconfigurable, certaines caractéristiques doivent être remplies telles que la modularité, l'intégrabilité, la convertibilité, le diagnostic, la personnalisation et l'évolutivité (Elmaraghy *et al.*, 2005, Koren *et al.*, 1999), ainsi que l'automatisation (Wiendahl *et al.*, 2007 ; Wiendahl *et al.*, 2004). Cette reconfiguration des systèmes

entraîne des impacts difficilement prévisibles qui doivent être anticipés (Plehn, 2017). Dans ce contexte, (Youssef et ElMaraghy, 2008) se sont orientés vers des approches de simulation pour optimiser et choisir la conception adéquate d'un RMS. D'autres auteurs ont étudié les problèmes d'ordonnancement dans les RMS. La spécificité de ce type de problème en RMS, c'est qu'il est possible d'avoir une multitude de gammes de fabrication, choisir entre elles et même changé entre elles (Dou *et al.*, 2016). On trouve aussi l'étude de problème de scalabilité (possibilité d'augmenter la capacité de production) dans le cas des RMSs pour un seul type de produit (Moghaddam *et al.*, 2018). Si différentes approches (Järvenäa 2016, Hammerstingl, 2018) ont été développées pour utiliser efficacement la flexibilité de RMS en automatisant la planification des processus (Battaia *et al.*, 2020), les problèmes de conception de lignes reconfigurables ne sont pas encore suffisamment étudiés.

En se basant sur ces observations, nous voyons qu'il est nécessaire d'adapter les approches existantes et/ou de développer de nouvelles méthodologies pour concevoir des RMS avec une capacité dynamique de changement. Ces méthodologies doivent considérer la reconfigurabilité comme une propriété dynamique du système de production. Ainsi, une caractéristique clé de l'industrie 4.0 est le développement de systèmes de production changeable et reconfigurables.

L'industrie 4.0 affecte les différentes dimensions de performance d'un système de production, elle doit donc être considérée comme un facteur lors de la conception de RMS. Ce dernier doit être réactif pour s'adapter rapidement aux conditions changeantes. Cela devient possible dans le contexte de l'industrie 4.0 grâce à l'émergence des jumeaux numériques (Digital Twin (DT)) (Grieves *et al.*, 2017). Des chercheurs comme Xu *et al.* (2018) ont souligné que la réalisation de l'industrie 4.0 est entravée par le manque d'outils puissants, en particulier les méthodes formelles et les méthodes systémiques. Les auteurs ont aussi montré le manque de préparation à soutenir la transformation numérique des infrastructures actuelles et le besoin de nouvelles infrastructures. Pour y remédier, ils ont argumenté qu'il était nécessaire de développer un DT avec des services décisionnels à valeur ajoutée. Le concept de DT a été étendu du produit au système de production (Ding *et al.*, 2019 ; Tao *et al.*, 2018 ; Tuegel *et al.*, 2011).

Le DT a pour objectif d'améliorer la production en effectuant une prise de décision autonome. Cette dernière est basée sur des indicateurs de performance, ainsi que l'utilisation de simulations et de données de capteurs dans la planification et l'exploitation (Boschert *et al.*, 2016). Zheng *et al.* (2019) ont mené une étude systématique sur la recherche et les applications des DT. Les auteurs ont souligné que la plupart des recherches autour des DT manquent de cas d'application référencables et sont uniquement au stade théorique. Ils ont également argumenté que les méthodes et les cadres d'application existants manquent de clarté. Sur cette base, les auteurs ont proposé un cadre d'application de DT pour la gestion du cycle de vie des produits.

Leng *et al.* (2020) ont proposé une reconfiguration rapide du système de production automatisé en utilisant une nouvelle approche orientée Digital Twin. Les auteurs ont tenté de trouver une solution de reconfiguration optimale grâce à un modèle de programmation à deux niveaux, le rééquilibrage de la productivité de niveau supérieur et la minimisation du coût de reconfiguration de niveau inférieur.

La complexité de la reconfiguration RMS augmente considérablement si l'on considère à la fois au niveau de la machine et au niveau du système (à la fois au niveau matériel et logiciel) (Huang *et al.*, 2018). Ainsi, le nombre de configurations augmente considérablement avec le nombre de machines et équipements. Cela rend le processus de conception puis de sélection d'une configuration RMS très complexe (Koren *et al.*, 2018). Développer des approches globales semble être très difficile en raison de la grande variété de problèmes. En conséquence, l'intégration de plusieurs DT dans RMS pourrait être plus efficace. Cela est dû au fait que chaque DT peut se concentrer sur des problèmes spécifiques tels que la diagnosticabilité, la maintenance prédictive, la formation, la sélection de reconfiguration ou la planification, etc.

En outre, souvent dans la pratique, les opérateurs s'appuient sur leurs expériences et jugements (Charalambous *et al.*, 2019) malgré un grand nombre de capteurs et des volumes de données, ce qui indique que les systèmes avancés de planification et de contrôle des processus de production souffrent de l'absence de solutions bien intégrées malgré le potentiel énorme de la numérisation. De plus, DT peuvent aider à optimiser les activités de la planification grâce à leur contrôle complet et en temps réel des flux d'informations. En outre, la disponibilité de quantités importantes d'information favorise l'adoption de la modélisation et la planification basées sur les données (data-driven), ainsi que l'utilisation de l'IoT et des technologies d'apprentissage automatique (Machine Learning (ML)). L'efficacité d'utilisation des techniques de ML pour les processus de production intelligents ont été démontrées dans le contrôle de l'atelier (Wang *et al.*, 2019), la maintenance prédictive (Yam *et al.*, 2001), la détection d'anomalies (Shang *et al.*, 2015) et la planification en ligne (Hammami *et al.*, 2017). Cette adoption et ces diverses utilisations de technologie DT sont largement considérées comme capables de révolutionner l'industrie manufacturière. Dans l'ensemble, on trouve que le ML et l'analyse des données n'ont pas été suffisamment explorés pour la conception du système de production, la planification de la production et l'ordonnancement (Shang *et al.*, 2019), en particulier en lien avec DT. Ceci est dû aux difficultés de sélection, d'agrégation et de l'analyse des données de manière générale.

Récemment, pour relever ces défis, des recherches sont faites pour acquérir des modèles à partir de données (Bessiere *et al.*, 2017 ; Kumar *et al.*, 2019 ; Beldiceanu *et al.*, 2016). Cependant, à notre connaissance, les travaux de recherche se concentrent davantage sur des parties spécifiques de couplage d'analyse de données, d'apprentissage machine et d'optimisation, et peu de travaux de recherche ont examiné les avantages d'un cadre global intégrant le DT avec ces techniques dans RMS tout en tenant compte des

caractéristiques fondamentales de RMS ainsi que de la reconfiguration logicielle possible. De plus, les travaux qui existent traitent seulement une partie spécifique qui touche juste une composante du RMS (Huang *et al.*, 2020 ; Kurniadi *et al.*, 2018) mais pas le système en entier.

Dans ce contexte, notre article présente une première tentative pour combler cette lacune. L'objectif est de profiter du concept de jumeau numérique et la conception générative pour proposer un cadre modulaire qui peut être utilisé pour assurer une reconfiguration facile de la partie logicielle du RMS en ajoutant de nouveaux blocs logiciels (ex: simulation, contrôle

...) selon le besoin et au bon moment du besoin de manière «Plug & Play», ce qui finira par aider à une meilleure sélection et donnera une meilleure compréhension et un processus de prise de décision flexible (par exemple pour la reconfiguration matérielle). De plus, Nous constatons qu'il y a une nécessité d'avoir un Framework global/générique. Ce dernier doit avoir une vue intégrée du système avec plusieurs modules qui peuvent être changé, adapté et/ou supprimé selon les tâches nécessaire (conception, reconfiguration, de planification et de conduite ...).

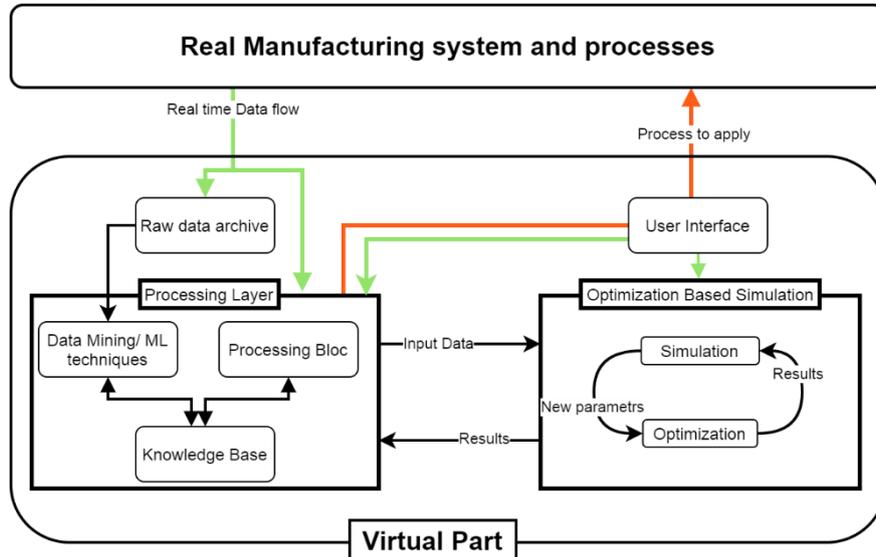


Figure 1. RMS-DT Framework

3 FRAMEWORK MODULAIRE RMS-DT

Dans cette section, nous présentons le cadre modulaire RMS-DT proposé pour un jumeau numérique dans RMS (Figure 1), expliquons ses différents blocs/modules ainsi que le flux d'informations. Nous referons à ce cadre par Framework RMS-DT dans la suite du document. Le Framework proposé se concentre sur l'espace virtuel. Il est basé sur une interprétation large du concept DT. Nous ne nous sommes pas concentrés sur toutes les technologies possibles liées à l'industrie 4.0. Les composants principaux de DT sont l'espace physique, l'espace virtuel et le flux de données/informations entre eux. Cet article se concentre uniquement sur l'espace virtuel.

Ceci va nous permettre de prendre des décisions de production de bas niveau telles que le réglage automatisé de la machine ou la qualité prédictive. De plus, Framework RMS-DT va nous permettre d'étendre l'utilisation de l'IA à des décisions de production de plus haut niveau (telles que la planification des processus, la planification de la production et l'ordonnancement) en s'appuyant sur des analyses prédictives et prescriptives (Gartner report, 2019).

Notre Framework RMS-DT est modulaire et décrit le processus de modélisation virtuelle et l'idée de construction des sous-systèmes d'application de DT. Il est introduit pour résoudre, entre autres, des problèmes de reconfiguration complexes. Comme nous avons déjà dit, quand le nombre de machines (RMT) augmente, le nombre de configurations RMS augmente plus rapidement qu'une fonction exponentielle naturelle (Koren *et al.*, 2018). Par conséquent, pour faire face à

la reconfiguration rapide des systèmes de production, il est nécessaire d'avoir une architecture de jumeau numérique comme base ainsi que des méthodes d'optimisations (Leng *et al.*, 2020). Dans ce contexte, RMS-DT représente une tentative de s'attaquer à ce type de problèmes et leur complexité de reconfiguration dans un environnement reconfigurable.

En utilisant le Framework RMS-DT, ce problème peut être résolu en suivant un processus en deux étapes. La première étape consiste à exécuter des expériences de reconfiguration sur l'espace virtuel haut-fidélité du RMS-DT. De plus, le flux de données en temps réel entre les deux espaces (physique et virtuel) est fluide et sans interruption. Cela crée et favorise des conditions qui permettent d'explorer toutes les reconfigurations possibles, de les optimiser, de les simuler et de les évaluer sur l'espace virtuel. Dans un deuxième temps, le retour d'expérience de l'espace virtuel (ex. : production optimisée, meilleure reconfiguration, maintenance ...) peut être utilisé pour être appliqué efficacement sur l'espace physique.

Ce Framework peut également être utilisé pour d'autres systèmes changeables (ex. Système de production Flexible). Ceci est possible grâce aux principes du RMS (dont la modularité et la reconfigurabilité) ainsi que l'intégration des avantages du DT et ML. En outre, au mieux de nos connaissances, la plupart des travaux de recherches sur RMS se concentrent plus sur la reconfiguration physique sans aborder la reconfiguration logicielle (software). Dans ce contexte, et afin de profiter des avantages de ce concept RMS, l'objectif de la structure modulaire RMS-DT est d'assurer une

reconfiguration aisée et facile du logiciel (ajout, modification et/ou suppression de blocs/modules) de manière « Plug & Play ». Cela nous permet d'avoir un Framework générique qui peut être spécifié en se basant sur l'option (Plug & Play). Par conséquent, le Framework RMS-DT est extensible et adaptable à différents besoins en intégrant de nouvelles technologies ou en adaptant/développant de nouveaux comportements/tâches. Par exemple le module de "Simulation basée sur l'optimisation" peut être remplacé par un solveur destiné à d'autres tâches comme trouver la meilleure stratégie de maintenance au lieu de trouver "la meilleure reconfiguration".

Etant donné que des décisions de haut niveau doivent impliquer la responsabilité d'un humain, le Framework RMS-DT s'appuiera sur le cadre de conception générative (Bahar et al., 2019) afin d'aider les utilisateurs à trouver les meilleures solution/plans. La conception générative (Figure 2) est un processus de conception itératif permettant aux utilisateurs d'exprimer leurs buts sous forme d'objectifs et de contraintes au Framework RMS-DT, ensuite ce dernier renvoie un ensemble de solutions de conception réalisables et / ou optimales.

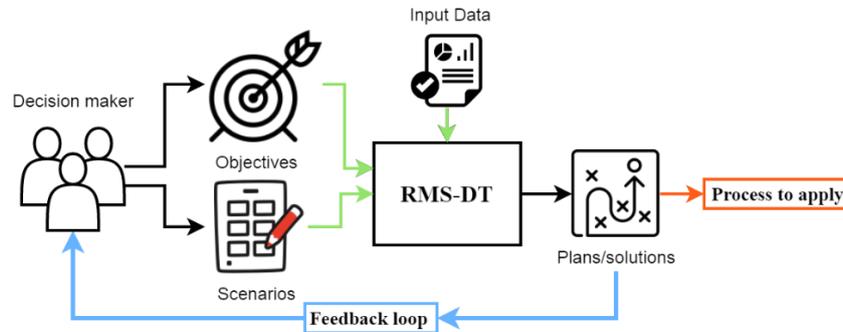


Figure 2. Conception générative

Pour garantir des conditions de haute-fidélité pour résoudre le problème, les données collectées à partir de la partie physique (à l'aide de technologies d'acquisition de données comme des capteurs par exemple) sont transmises en temps réel à la partie virtuelle (flèche verte sur la Figure 1). Les données sont collectées à partir de différentes sources (systèmes de fabrication, machines, AGV, PLC, etc.). Les informations brutes collectées en temps réel seront ensuite redirigées vers deux modules (i) "Raw data archive" et (ii) "Processing layer".

Le module "Raw Data Archive" stockera simplement les données brutes telles qu'elles sont collectées depuis l'espace physique. Cette archive de données brutes peut être utilisée plus tard comme source de données historiques, qui peuvent être une entrée pour le module "Data Mining/Machine Learning (ML) techniques" qui sera détaillé ci-après.

Le module "Processing Layer" : connecte le flux d'informations entre l'espace physique et l'espace virtuel via le mappage, le traitement et le stockage d'informations. L'interaction des données au sein de ce module permet l'interopérabilité des deux espaces (physique et virtuel). Ce module comprend trois modules fonctionnels principaux qui sont respectivement :

- **Le module "Processing Bloc" :** Ce module est responsable du traitement, de la cartographie et de l'analyse des données en temps réel de l'espace physique. Tout d'abord, les données brutes sont traitées par nettoyage (par exemple basé sur des règles logiques), structuration et clustering. Les données propres sont mappées pour correspondre et extraire les besoins des modèles de décideurs, des variables de décision et des informations. Les données obtenues sont ensuite analysées et comparées aux données stockées dans le module "Knowledge base" représentant la Base de connaissances. Après la

comparaison, nous nous retrouvons avec deux scénarios :

1. Les informations traitées n'existent pas dans la base de connaissances. Dans ce cas, les données sont stockées dans la base de connaissances puis redirigées vers le module "Optimization based simulation" qui est responsable de l'optimisation. L'objectif est de trouver une nouvelle solution correspondante aux nouvelles données.
2. La base de connaissances contient déjà ces données. Dans ce cas, la solution correspondante à appliquer est récupérée et redirigée vers l'espace physique.

- **Le module "Knowledge Base" :** représente la base de connaissances du système. Il stocke les données des deux espaces (physique et virtuel). Toutes les données stockées sont considérées comme des données traitées. Cela signifie que les données proviennent du module "processing bloc" ou qu'elles représentent de nouvelles données du "Data mining/Machine Learning (ML)". Les données de l'espace physique comprennent (mais sans s'y limiter) les données d'équipement (capacité, fonctionnalité, reconfigurations, outils, ...), les données d'environnement d'atelier, les données de production, etc. Pour les données d'espace virtuel, nous pouvons trouver (également sans s'y limiter) les données de décision, résultats d'évaluation de simulation/optimisation, données de prédiction et nouveaux modèles ou modèles inférés, etc.
- **Le module "Data Mining/Machine Learning (ML)" :** a pour rôle d'inférer/trouver de nouveaux modèles et corrélations entre diverses informations collectées collectées/stockées à la fois dans le module "Knowledge Base" et le module "Raw Data Archive".

Une fois le flux de données temps réel traité par le module "*Processing Layer*" et redirigé vers le module "*Optimization Based Simulation*", ce dernier tentera de trouver la meilleure solution. Ce module suit un processus en boucle en deux étapes. Tout d'abord, l'optimisation d'une fonction objective soumise à des contraintes (telle que définie par le décideur ou une équipe d'experts) est effectuée. Ensuite, une simulation stochastique est utilisée pour évaluer les résultats d'optimisation ainsi que la (les) fonction(s) objectif(s) et ses contraintes respectives. Le processus est ensuite répété jusqu'à la satisfaction d'un critère d'arrêt (ex. : une limite de temps, des solutions stables, ...).

Le module "*Optimization Based Simulation*" peut également être utilisé directement en mode asynchrone vers l'espace physique via l'Interface utilisateur ("**User interface**"). Cette dernière profite de la conception générative abordée dans la figure 2 en offrant à l'utilisateur (ex. : le décideur) la possibilité de tester et d'évaluer de nouveaux scénarios ou d'introduire de nouveaux paramètres du processus/système de production, ainsi que de choisir les solutions selon sa préférence. Dans ce contexte, l'utilisateur fixe des objectifs de production (indicateurs de performance, seuils, ...) et le Framework RMS-DT fournit des plans pour atteindre ces objectifs. Par la suite, le décideur peut affiner de manière itérative les objectifs exprimés jusqu'à atteindre son plan préféré. Le but final est d'obtenir des informations à leur sujet et de prédire, contrôler et optimiser le système en exploitation dans son environnement. Le Framework RMS-DT est considéré ici comme un cadre de prise de décision basées sur les données (data-driven decisions), en utilisant des techniques de conception générative pour enrichir l'apprentissage automatique.

4 SCENARIO ILLUSTRATIF D'UTILISATION

Pour illustrer comment ce Framework RMS-DT peut être exploité, prenant l'exemple de production d'une famille de produit en minimisant le coût et le temps. Les RMS sont généralement conçus autour d'une famille de produits définie. Dans l'industrie, la fabrication de tout produit implique de nombreuses machines et procédés. Donc pour ce problème de conception de notre RMS, nous devons connaître toutes les machines nécessaires pour réaliser notre produit ainsi que leurs modules à concevoir. Le deuxième problème qu'on rencontre, est que même la famille de produits est définie, son évolution dans le temps (quel produit sera fabriqué et quelle quantité à chaque période de temps dans le futur) ne l'est pas. Néanmoins, notre connaissance des produits de cette famille nous permet d'en déduire leurs éventuelles gammes de fabrication avec leurs coûts et temps respectives. Ceci implique que notre RMS doit avoir la capacité de produire chacun de ces produits en se reconfigurant par remplacement et/ou adaptation de certains modules tout en minimisant le coût et le temps de la production.

A ce stade, la décision à prendre au moment de l'exploitation c'est de choisir quelle est la bonne reconfiguration ainsi que les modules nécessaires à mettre en place (en remplaçant,

ajoutant ou supprimant) pour répondre aux exigences de fabrication du nouveau produit (en quantité, coût et temps). Si la une demande concerne un produit non prévu lors de la conception, il faut que le système soit capable de déterminer rapidement s'il peut être produit sur la ligne et le cas échéant quels nouveaux modules ou machine il faut ajouter à l'ensemble initialement conçus et comment reconfigurer le RMS pour ce produit (choisir les bons modules au bon moment pour assurer une reconfiguration du système). Dans ce cas, il est nécessaire de réévaluer les performances, coût et temps nécessaire pour cette nouvelle configuration, sachant que pour chaque produit il y a une multitude de possibilités de gammes de fabrication, ce qui implique une explosion des possibilités pour un ensemble de produits donné.

Pour résoudre ce problème avec le Framework RMS-DT nous suivent les étapes décrites ci-après. Nous commençons par collecter et transmettre l'information sur l'ensemble de modules et d'équipement à notre disposition dans la partie physique (« real time Data-flow ») vers la partie Virtual du RMS-DT. Dans un deuxième lieu, il faut transmettre les exigences ainsi que les informations nécessaires concernant la fabrication de l'ensemble de produits. A ce stade le module "*interface utilisateur*" peut être exploité pour injecter de nouvelles informations ou scénarios. Ces informations seront redirigées vers deux modules (i) "*Raw data archive*" pour le stockage et (ii) "*Processing layer*" pour le traitement en utilisant "*Processing Bloc*". A ce niveau, le module "*Processing layer*" va tenter tout d'abord, de trouver une solution (la reconfiguration qui minimise le coût et le temps) dans sa base de connaissance "*Knowledge Base*" à partir des solutions proposées pour l'ensemble de produits déjà traité. Pour rappel cette base de connaissance est mise à jour à chaque fois qu'une nouvelle solution est trouvée en utilisant le module "*Data Mining/Machine Learning (ML)*" ou bien le module "*Optimization Based Simulation*". Si aucune solution n'est trouvée, le module "*Optimization Based Simulation*" va être engagé pour utilisant des modèles d'optimisation guidés par la minimisation du coût et du temps de production. Ce dernier affine les résultats trouvés d'abord par le biais du bloc "*Simulation*". Dans le cas d'optimisation multicritères, plusieurs bons résultats sont possibles, donc on aura besoin d'un retour incluant les décideurs via le « *Feedback loop* » pour la validation de ces résultats. Une fois validés, ces derniers seront intégrés dans la base de connaissance. Le module "*Data Mining/Machine Learning (ML)*" peut aller encore plus loin en inférant des nouvelles solutions, relations entre les données et/ou modèles. Ceci en utilisant les données historiques du système (brutes stockées dans "*Raw data archive*" ou bien traitées dans "*Knowledge Base*").

5 CONCLUSION

Dans cet article, un Framework, qui décrit l'utilisation du jumeau numérique (DT) dans un RMS est fourni. Le Framework montre comment un DT peut être utilisé pour un RMS afin d'obtenir la flexibilité et la réactivité nécessaires pendant sa phase d'exploitation en fournissant un processus de

prise de décision flexible. Cela se fait en collectant en continu des données en temps réel à partir des composants RMS. Ensuite, ces dernières sont stockées, traitées et analysées à l'aide des modules d'analyse, de simulation et d'optimisation de l'information. Sur cette base, DT peut rapidement prendre des décisions critiques telles que la configuration RMS appropriée pour faire face efficacement aux changements soudains. A travers l'article nous nous sommes concentrés uniquement sur la partie virtuelle. Les détails sur la partie physique du DT ainsi que la manière dont les données sont collectées et échangées entre le RMS et le DT, sont hors de la portée de cet article mais rentrent bien évidemment dans nos perspectives futures. De plus notre modèle doit traiter plusieurs aspects de différentes lignes de production qui représentent généralement des domaines complexes et hétérogènes.

Pour les perspectives de recherche futures, un Framework complet, qui comprend non seulement la partie virtuelle, mais également une description détaillée de la partie physique et des protocoles de connexion entre eux sera fourni. De plus, nous envisageons de développer une approche de simulation basée sur l'optimisation en ligne pour la sélection de la configuration RMS basée sur ce qui a été fait dans (Haddou Benderbal *et al.*, 2016). Cela constituera un premier pas vers la réalisation du Framework proposé. En conséquence, un DT peut être utilisé dans un contexte RMS pour atteindre la diagnostiquabilité, qui est une caractéristique fondamentale de tels systèmes et est définie comme la capacité d'un RMS à être facilement diagnostiqué. À notre connaissance, peu d'articles étudient cette fonctionnalité. Étant donné que DT offre une transparence du système et un retour d'information en temps réel, cela peut être utilisé pour détecter et prévenir les causes profondes des pannes de machines et du système. De plus, la combinaison de l'IA dans notre modèle peut permettre d'améliorer les jumeaux numériques avec l'apprentissage automatique (ML) pour l'analyse prescriptive dans la production. Ceci facilitera la prise des décisions flexibles et robustes dans les étapes de conception, d'exploitation et en temps réel. Enfin, une direction de travail intéressante est d'étudier comment d'autres caractéristiques RMS peuvent être intégrées à l'aide de ce Framework RMS-DT ainsi que son utilisation générale pour satisfaire et répondre aux besoins de l'Industrie 4.0.

6 REMERCIEMENTS

Ces travaux ont été soutenus par le programme PERFORM, IRT Jules Verne, ainsi que par la Région Pays de la Loire.

7 REFERENCES

- Abbas, H., Alur, R., Mamouras, K., Mangharam, R., & Rodionova, A. Real-time decision policies with predictable performance. *Proceedings of the IEEE*, 106(9), 1593-1615. (2018).
- Andersen, A.L.: Development of changeable and Reconfigurable Manufacturing Systems: supporting context-specific design of changeability. Ph.D. thesis, Aalborg University (2017)
- Bahar A., Hyunmin C., Christopher B.J., Vers une ontologie pour la conception générative des assemblages mécaniques, Appliqué Ontologie, vol. 14, n ° 2, p. 127-153, (2019)
- Baltean-Lugoian, R., Bonami, P., Misener, R., & Tramontani, A. Scoring positive semidefinite cutting planes for quadratic optimization via trained neural networks, (2019).
- Battaïa O., Benyoucef L., Delorme X., Dolgui A. and Thevenin S. Sustainable and energy efficient reconfigurable manufacturing systems, Chapter 9 in: *Reconfigurable Manufacturing Systems: From the Design to Implementation*, (pp. 179-191), Springer. (2020)
- Battaïa O., Dolgui A., Guschinsky N. Optimal Cost Design of Flow Lines with Reconfigurable Machines for Batch Production, *International Journal of Production Research*, (2020).
- Beldiceanu, N., & Simonis, H. Modelseeker: Extracting global constraint models from positive examples. In *Data Mining and Constraint Programming* (pp. 77-95). Springer, Cham. (2016a).
- Beldiceanu, N., Carlsson, M., Douence, R., & Simonis, H. Using finite transducers for describing and synthesising structural time-series constraints. *Constraints*, 21(1), 22-40. (2016b).
- Benyoucef, L.: *Reconfigurable Manufacturing Systems: From Design to Implementation*. Springer (2020).
- Bessiere, C., Koriche, F., Lazaar, N., & O'Sullivan, B. Constraint acquisition. *Artificial Intelligence*, 244, 315-342. (2017).
- Bortolini, M., Galizia, F.G., Mora, C.: Reconfigurable manufacturing systems: Literature review and research trend. *Journal of Manufacturing Systems* 49, 93–106(2018)
- Boschert, S., Rosen R.: Digital Twin—The Simulation Aspect. In: Hehenberger, P. et al. (Hrsg.): *Mechatronic Futures: Springer International Publishing*, S. 59-74. ISBN: 978-3-319-32154-7, (2016).
- Brahimi, N., Dolgui, A., Gurevsky, E., Yelles-Chaouche, A.R.: A literature re-view of optimization problems for reconfigurable manufacturing systems. *IFAC-PapersOnLine*52(13), 433–438 (2019)
- Charalambous, E., Feldmann R., Richter G., and Christoph Schmitz C., AI in production: A game changer for manufacturers with heavy assets.
- Ding, K., Chan, F.T., Zhang, X., Zhou, G., Zhang, F.: Defining a digital twin-based cyber-physical production system for autonomous manufacturing in smart shop floors. *International Journal of Production Research*57(20), 6315–6334 (2019)
- Dolgui, A., Proth, J.M.: *Supply chain engineering: useful methods and techniques*. Springer Science & Business Media (2010)
- Dou J., Li J. and Su C. Bi-objective optimization of integrating configuration generation and scheduling for reconfigurable flow lines using NSGA-II. *Inter. J. of Adv. Manuf. Techn.*, 86(5-8), 1945–1962. (2016).
- ElMaraghy, H.A., Wiendahl, H.P.: Changeability - An introduction. In: *Changeable and Reconfigurable Manufacturing Systems*, pp. 3–24. Springer (2009)
- ElMaraghy, H.A.: Flexible and reconfigurable manufacturing systems paradigms. *International Journal of Flexible Manufacturing Systems*17(4), 261–276 (2005)
- Farid, A.M.: An engineering systems introduction to axiomatic design. In: *Axiomatic Design in Large Systems*, pp. 3–47. Springer (2016)
- Figge F., Thorpe AS., Givry P., Canning L., Franklin-Johnson E. Longevity and circularity as indicators of eco-efficient resource use in the circular economy. *Ecological Economics*, 150, pp. 297-306, (2018)

- Gartner report, Forecast Snapshot: Prescriptive Analytics Software, Worldwide, 2019
- Grieves, M., Vickers, J.: Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: *Transdisciplinary Perspectives on Complex Systems*, pp. 85–113. Springer (2017)
- Haddou-Benderbal, H., Dahane, M., Benyoucef, L.: Hybrid heuristic to minimize machine's unavailability impact on reconfigurable manufacturing system using re-configurable process plan. *IFAC-PapersOnLine*49(12), 1626–1631 (2016)
- Haddou-Benderbal, H., Yelles-Chaouche, A. R., & Dolgui, A. A Digital Twin Modular Framework for Reconfigurable Manufacturing Systems. In *IFIP International Conference on Advances in Production Management Systems* (pp. 493-500). Springer, Cham. (2020).
- Hammami, Z., Mouelhi, W., & Said, L. B. (2017). On-line self-adaptive framework for tailoring a neural-agent learning model addressing dynamic real-time scheduling problems. *Journal of Manufacturing Systems*, 45, 97-108.
- Hammerstingl V, Reinhart G. Skills in Assembly. Technical report (2018).
<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/ai-in-production-a-game-changer-for-manufacturers-with-heavy-assets> (March 7, 2019).
https://www.researchgate.net/publication/322580474_Skills_in_Assembly
- Huang, S., Wang, G., & Yan, Y. Building blocks for digital twin of reconfigurable machine tools from design perspective. *International Journal of Production Research*, 1-15. (2020).
- Huang, S., Wang, G., Shang, X., Yan, Y.: Reconfiguration point decision method based on dynamic complexity for reconfigurable manufacturing system (RMS). *Journal of Intelligent Manufacturing*29(5), 1031–1043 (2018)
- Järvenpää E., Siltala N., Lanz M. Formal resource and capability descriptions supporting rapid reconfiguration of assembly systems. In: *IEEE International Symposium on Assembly and Manufacturing (ISAM)*, 120-125, (2016).
- Koren, Y., Gu, X., Guo, W.: Reconfigurable manufacturing systems: Principles, design, and future trends. *Frontiers of Mechanical Engineering*13(2), 121–136(2018)
- Koren, Y., Heisel, U., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G., Van Brussel, H.: Reconfigurable manufacturing systems. In: *Manufacturing Technologies for Machines of the Future*, pp. 627–665. Springer Berlin Heidelberg (1999)
- Kumar, M., Teso, S., & De Raedt, L. Acquiring Integer Programs from Data. *IJCAI 2019*: 1130-1136, (2019).
- Kurniadi, K. A., Lee, S., & Ryu, K. Digital twin approach for solving reconfiguration planning problems in RMS. In *IFIP International Conference on Advances in Production Management Systems* (pp. 327-334). Springer, Cham. (2018).
- Leng, J., Liu, Q., Ye, S., Jing, J., Wang, Y., Zhang, C., & Chen, X.: Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. *Robotics and Computer-Integrated Manufacturing*, 63, 101895. (2020).
- Leng, J., Liu, Q., Ye, S., Jing, J., Wang, Y., Zhang, C., Zhang, D., Chen, X.: Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. *Robotics and Computer-Integrated Manufacturing* 63,101895 (2020)
- Microsoft, The Future Computed – AI & Manufacturing, May 2019
- Moghaddam S.K., Houshmand M. and Valilai O. F. Configuration design in scalable reconfigurable manufacturing systems (RMS); a case of single-product flow line (SPFL). *Int. J. of Prod. Res.*, 56 (11), 3932–3954. (2018).
- Oztemel, E., Gursev, S.: A taxonomy of industry 4.0 and related technologies. In: *Industry 4.0-Current Status and Future Trends*. IntechOpen (2020)
- Plehn C., A Method for Analyzing the Impact of Changes and their Propagation in Manufacturing Systems. Dissertation, TU Munich (2017).
- Shang, C., & You, F. Data analytics and machine learning for smart process manufacturing: recent advances and perspectives in the big data era. *Engineering*, 5(6), 1010-1016. (2019).
- Shang, C., Yang, F., Gao, X., Huang, X., Suykens, J. A., & Huang, D. Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis. *AIChE Journal*, 61(11), 3666-3682. (2015).
- Tao, F., Zhang, H., Liu, A., Nee, A.Y.: Digital twin in industry: State-of-the-art. *IEEE Transactions on Industrial Informatics*15(4), 2405–2415 (2018)
- Tuegel, E.J., Ingrassia, A.R., Eason, T.G., Spottswood, S.M.: Reengineering air-craft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*2011, 14 pages (2011)
- Wang, C., & Jiang, P. Deep neural networks based order completion time prediction by using real-time job shop RFID data. *Journal of Intelligent Manufacturing*, 30(3), 1303-1318. (2019).
- Wiendahl, H.P., Heger, C.: Justifying changeability. A methodical approach to achieving cost effectiveness. *Journal for Manufacturing Science and Production*6(1-2), 33–40 (2004)
- Wiendahl, H.P., ElMaraghy, H.A., Nyhuis, P., Z'ah, M.F., Wiendahl, H.H., Duffie, N., Brieke, M.: Changeable manufacturing: classification, design and operation. *CIRP Annals* 56(2), 783–809 (2007)
- Xu, L.D., Xu, E.L., Li, L.: Industry 4.0: state of the art and future trends. *Inter-national Journal of Production Research* 56(8), 2941–2962 (2018)
- Yam, R. C. M., Tse, P. W., Li, L., & Tu, P. Intelligent predictive decision support system for condition-based maintenance. *The International Journal of Advanced Manufacturing Technology*, 17(5), 383-391. (2001).
- Yelles-Chaouche, A. R., Gurevsky, E., Brahimi, N., & Dolgui, A. Reconfigurable manufacturing systems from an optimisation perspective: a focused review of literature. *International Journal of Production Research*, 1-19. (2020).
- Youssef A. and ElMaraghy H. Availability consideration in the optimal selection of multiple-aspect RMS configurations. *Inter. J. Prod. Res.*, 46, 5849–5882. (2008).
- Zheng, Y., Yang, S., Cheng, H.: An application framework of digital twin and its case study. *Journal of Ambient Intelligence and Humanized Computing*10(3),1141–1153 (2019)

Intégration d'objets connectés : Eléments d'aide au choix de plateformes IIoT

JUSSARA SANTANA¹, PIERS BARRIOS^{1,2}, BENOIT EYNARD^{1,2}

¹ UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE

Compiègne, France

jussara.santana@etu.utc.fr, benoit.eynard@utc.fr

² INETUM

Meudon-la-Forêt, France

piers.barrios@gfi.fr

Résumé – Les objets connectés industriels font partie de l'écosystème plus large des objets connectés. Ce dernier marque la quatrième révolution industrielle qui se concentre sur la communication entre les machines, les données massives et l'apprentissage automatique. Aujourd'hui en 2020, d'aucuns estiment qu'il y a environ 30 milliards d'objets connectés. D'ici 2025, plus de 75 milliards d'appareils seraient connectés. Les objets connectés industriels offrent aux industries de nombreuses opportunités d'amélioration. Cet article propose une méthodologie d'aide au choix de plateformes IoT open-source et propriétaires pour une ligne de production. A l'aide de cette méthodologie, deux plateformes ont été sélectionnées. Nous avons présenté leurs principales caractéristiques ainsi que projeté leur utilisation sur un démonstrateur : la ligne AMS.

Abstract – Industrial Internet of things are part of the wider ecosystem of Internet of Things. The latter marks the fourth industrial revolution which focuses on communication between machines, big data and machine learning. Today in 2020, some estimate around 30 billion connected objects. By 2025, more than 75 billion devices should be connected. Industrial internet of things offers to industries many improvement opportunities. This article presents a methodology to help industries to choose an IoT platforms for a production line. Using this methodology, two platforms were selected. We have presented their main characteristics as well as planned their use on a demonstrator: the AMS line.

Mots clés - Industrie 4.0, Objets Connectés Industriels, Plateformes IoT, Choix de Technologies

Keywords - Industry 4.0, Industrial Internet of things, IoT Platforms, Technology Choice

1 INTRODUCTION

La première référence à l'Industrie 4.0 (aussi appelée Industrie du Futur) a été introduite à la foire des technologies industrielles de Hanovre, Allemagne en 2011 [Bartevyan, 2015]. Depuis, de nombreuses entreprises ont commencé à développer des solutions conformes au concept de l'Industrie 4.0. De nombreux gouvernements soutiennent également le développement de telles solutions notamment les européens, états-unis et japonais confirmant que cette nouvelle industrie est considérée comme stratégique par les grandes puissances et acteurs industriels.

L'industrie 4.0 est innovante, compétitive, performante. Elle est également flexible, car sa chaîne de valeur peut être recomposé pour mieux s'adapter aux évolutions soit au niveau du marché, soit au niveau technologique [Alliance Industrie du futur, 2018].

Selon [Burke et al., 2017] l'Industrie 4.0 est aussi proactive (cherche à prédire les problèmes), transparente (permet une meilleure visibilité sur le fonctionnement global de l'industrie) et agile (peut être reconfigurée de façon rapide).

Les objets connectés (Internet of Things (IoT) en anglais) sont l'extension du réseau internet au monde physique [Julien et Martin, 2018]. Les objets connectés permettent aux entreprises de visualiser leur évolution, améliorer l'expérience client, économiser du temps et de l'argent, améliorer la productivité des employés, prendre des décisions et générer plus de

revenus. L'intérêt de cette étude est de pouvoir bénéficier de critères d'évaluation, d'analyse et de choix entre les multiples plateformes existantes et leur couverture fonctionnelle très hétérogène.

Les objets connectés industriels (Industrial Internet of Things (IIoT) en anglais) sont basés sur des systèmes cyber-physiques dont l'architecture minimale se compose de trois éléments : capteurs, actionneurs et contrôleurs [Hehenberger et al., 2016]. C'est grâce à ces éléments qu'on va pouvoir contrôler, surveiller les machines et transformer les données collectées en informations utiles à la prise de décision. Des technologies peuvent être considérées comme adjacentes à l'IIoT en raison de leur actionnabilité respectives. Celles-ci sont par exemple :

- L'informatique en nuage (cloud): modèle informatique qui permet un accès facile et à la demande par le réseau à un ensemble partagé de ressources informatiques configurables [Figer, 2012] ;
- Les systèmes cyber physiques (cyber physical system): des systèmes qui connectent le monde physique et le monde virtuel à travers l'internet [Rajkumar et al., 2010] ;
- Les données massives (big data) qui permettent d'identifier les briques numériques nécessaires à l'amélioration des procédés industriels et à la compréhension des besoins des clients.

Dans ce contexte global, l'objectif de cet article est de présenter une méthodologie d'aide au choix de plateformes

IoT, afin de combler des lacunes dans la littérature scientifique sur le choix des technologies de l'industrie du futur, tout en intégrant dans un même document une méthodologie de choix basée sur la pondération de critères propres à ce type de solution ainsi que l'intégration des solutions retenues dans un contexte industriel réel. Pour cela, nous vous présenterons dans un premier temps l'état de l'art couvrant le nombre d'objets connectés ainsi que la notion de plateforme IoT. Ensuite, nous présenterons notre méthodologie ainsi que quelques plateformes IoT sur lesquelles nous l'avons appliquée. Celles-ci, une quinzaine, sont divisées en deux groupes : plateformes open-source et plateformes propriétaires. Ensuite, nous présenterons une étude de cas sur la plateforme AMS. Enfin, la section discussion, propose une analyse de la méthodologie et la conclusion permettra de présenter les travaux futurs qui seront conduits.

2 ETAT DE L'ART

Alors que nous avons présentés en introduction quelques briques technologiques de l'industrie 4.0, nous souhaitons nous attarder sur deux aspects : le nombre d'objets connectés et son évolution d'une part et la notion d'objets connectés industriels et de plateforme IoT. En effet, la croissance du nombre d'objets connectés nous pousse à recourir à des solutions de gestion des IoT : les plateformes IoT.

2.1 Evolution du nombre d'objets connectés

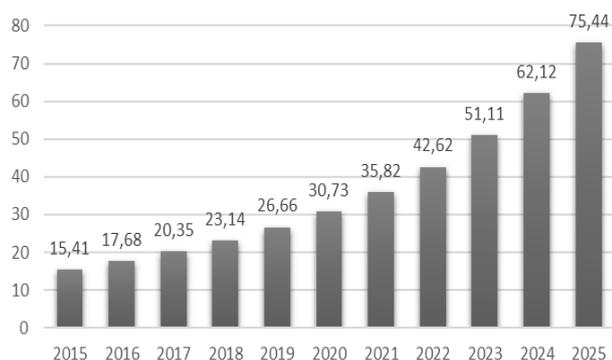


Figure 1. Nombre d'objets connectés estimés (en milliards) par année entre 2015 et 2025 selon [Statista Research Department, 2016]

Le nombre d'objets connectés et l'évolution de leur nombre fait actuellement débat en l'absence d'étude sérieuse sur le sujet. Les estimations sont diverses :

- Selon [Statista Research Department, 2016], et relayé par [Alam, 2018], le nombre d'objets connectés aurait été de 27 milliards en 2019. En 2025, le nombre total d'objets connectés dans le monde sera d'environ 75 milliards.
- Selon [The Internet of Things Report, 2020] qui fait une estimation entre 2019 et 2027, il y aurait eu 9 milliards d'IoT en 2019 et une croissance quasi-linéaire pour atteindre 41 milliard en 2027.
- Enfin [IoT Market - Forecasts at a Glance - IoT Analytics, 2014], relayé par [Odini, 2016], couvre la période 2014-2020 seulement mais mettait déjà en avant la disparité de 7 estimations provenant d'industriels : pour 2020, la plus basse projetant 18 milliards d'IoT et la plus importante 50 milliards d'IoT.

Aussi, aucune de ces références ne met en avant la méthodologie qui a conduit à ces estimations. De plus, que les références soient récentes ou plus anciennes, on remarque une

constance dans les écarts entre les différentes estimations pour une même année. Le seul point de consensus de ces études est qu'il y aura une croissance continue sans mise en exergue de quelconques facteurs limitants. Or, ceux-ci ont été mis en avant depuis de multiples décennies par [Meadows & Club of Rome, 1972] et sont désormais partagés par la communauté scientifique.

Ainsi, l'évolution du nombre d'objets connectés seraient potentiellement en croissance *a minima* pour la décennie à venir.

2.2 Objets connectés industriels

[Barrios et al., 2019] considère qu'en opposition avec l'IoT, l'IIoT se démarque par ses besoins en termes de mise à l'échelle, de gestion déportée des objets, de sécurité et de résilience ainsi que des cas d'utilisation en environnement parfois rude. Cet article portant sur l'interaction entre l'IIoT et la gestion du cycle de vie des produits (produit lifecycle management, PLM), suggère que l'imbrication de ces derniers est une clef dans l'émergence de nombreuses problématiques d'actualité.

Les objets connectés industriels sont l'utilisation des technologies d'objets connectés pour l'amélioration de la fabrication et des processus industriels. Ils sont présents sur les machines et les objets en cours de fabrication. Chaque automate, chaque moteur, chaque variateur ou chaque capteur peut ainsi informer directement les autres composants de la ligne de son état et leur fournir des données qui leur permettront de se paramétrer automatiquement, sans nécessiter de passer par un système central, réduisant considérablement les temps de réaction de l'ensemble [Julien et Martin, 2018].

Selon [Danjou et al., 2017], l'IIoT cherche à répondre deux principaux défis :

1. Imaginer, envisager, anticiper, de quelle façon ces technologies peuvent se combiner pour transformer les produits, les processus et les services offerts ;
2. Maîtriser ces technologies, souvent extérieures au cœur de métier de l'entreprise, afin d'être en mesure de créer ces nouveaux processus, produits ou services ;

Ainsi, la résolution de chacun de ces défis se fera à l'aide des technologies cités dans la première partie de ce document, ce qui permettrait aux chefs d'entreprises de mieux organiser et gérer leurs moyens de production, afin d'avoir une plus grande adaptabilité, une allocation plus efficace des ressources, le pilotage à distance et la simulation du fonctionnement de l'usine en temps réel.

Nous pouvons donc remarquer que de nombreuses études portant sur l'Industrial Internet of Things existent, et c'est pourquoi avant l'implémentation d'une solution IoT au sein d'une industrie, il faudrait analyser et comprendre les documents qui portent sur ce sujet, dès sa racine (les objets connectés et ses briques technologiques), jusqu'à son déploiement au sein d'une industrie, en passant par le choix de la plateforme la plus adaptée.

2.3 Plateformes IIoT

A présent, le marché des solutions IIoT est en pleine croissance et les solutions éligibles pour répondre aux besoins des entreprises sont de plus en plus nombreuses. S'il y avait, par le passé, une absence de solutions sur le marché, il y en a aujourd'hui de nombreuses solutions capables de répondre aux besoins des industriels.

[Ullah et al., 2020] propose un framework d'éléments clés d'une plateforme d'objets connectés. Il mentionne vingt et un éléments, tels que la scalabilité, la sécurité, le temps de mise

sur le marché, l'analyse et visualisation de données, les protocoles de communication et d'autres. Toutefois, son document ne propose pas des critères de sélection spécifique aux plateformes d'objets connectés industriels.

[Gartner, 2020] propose une évaluation du marché des IIoT. Dans cette évaluation, 17 fournisseurs ont été classés selon deux axes (capacité à exécuter et l'intégralité de la vision) et en quatre catégories (leaders, challengers, visionnaires et acteurs de niches). Nous trouvons PTC, Hitachi, et Microsoft comme étant les trois fournisseurs classés en tant que leaders du marché IIoT. PTC avec sa solution ThingWorx est, selon cette évaluation, la solution IIoT la plus complète du marché grâce à ses efforts d'innovation et sa capacité à offrir un modèle commercial qui simplifie la réussite de ses clients. Par ailleurs, l'étude présente l'analyse de données, la gestion d'appareils, la sécurité, la gestion des données (y compris le stockage des données) en tant que fonctionnalités de base des plateformes IIoT, ainsi que les forces, les faiblesses des plateformes classées et l'ensemble des critères d'évaluation utilisés. Cependant, la méthodologie qui conduit au choix de ces plateformes n'a pas été décrite.

[Choi et al., 2018] cite les éléments clés nécessaires à une plateforme IIoT, tels que la connectivité, la gestion d'appareils, la sécurité, la visualisation des données, l'environnement pour le développement de la plateforme ainsi que la performance. Il mentionne les défis techniques qui cherchent à être résolus par les plateformes IIoT dans le cadre de l'usine du futur : la gestion et l'extraction des données des machines, les mécanismes de protection du système, le déploiement d'alertes en cas de pannes et d'autres.

L'article [Zdravković et al., 2016] se centre sur l'étude de dix-sept plateformes IIoT basées sur le cloud. Ces dernières ont été sélectionnées selon : le nombre d'utilisateurs (base de clients) importants, l'existence de partenariats avec des fabricants d'appareils et des intégrateurs de systèmes. L'article donne donc un bref aperçu des principales fonctionnalités et caractéristiques des plateformes sélectionnées. L'article mentionne la connectivité en tant qu'élément clé d'une plateforme IIoT. D'autres éléments sont cités, tels que la gestion et la maintenance d'appareils, la visualisation et l'analyse de données, ainsi que l'existence d'applications pour la mise en place d'alertes.

[Kim et al., 2019] présente les facteurs qui encouragent les entreprises à adopter des plateformes IIoT open-source à l'instar des propriétaires. Tout comme [Choi et al., 2018], l'article mentionne la gestion d'appareils, l'analyse de données et l'aspect sécurité du système, en tant qu'éléments clés des plateformes IIoT. Enfin, ils décrivent et comparent 5 plateformes open-source, selon les critères suivants : l'analyse de données, les bases de données, le type de protocole et le langage de programmation supporté.

Ainsi, nous pouvons remarquer que sur le fond scientifique, la plupart des différents documents mentionnés dans cette section ont des points communs en termes de critères d'évaluation de plateformes IIoT (Tableau 1).

Tableau 1. Critères communs aux différents articles

	Connectivité	Gestion d'appareils	Visualisation de données	Sécurité	Base de données
[Choi et al., 2018]	x	x	x		
[Zdravković et al., 2016]	x	x	x	x	
[Kim et al., 2019]	x			x	x
[Ullah et al., 2020]	x	x	x	x	
[Gartner, 2020]	x	x	x	x	x

Enfin, nombreux de ces articles proposent des pistes d'implémentation pour une plateforme. Cependant, dans un contexte industriel, les acteurs recherchent une solution

installée sur le marché et, sauf besoin exceptionnels, seront réticents à utiliser une plateforme dont le développement n'est pas stabilisé. Aussi, selon les besoins spécifiques des acteurs, ceux-ci sont en recherche de solutions propriétaires et/ou open-source ; nous n'aborderons pas ici les éléments amenant à ce choix.

En conclusion, cet état de l'art a montré l'évolution *a priori* croissante du nombre d'IIoT et nous avons souligné les spécificités des objets connectés industriels. Ces deux éléments tendent à justifier la nécessité d'avoir une solution permettant de gérer les IIoT : les plateformes IIoT. A ce sujet, la littérature met en avant une hétérogénéité des facteurs permettant de définir une plateforme IIoT et fait la comparaison de certaines ; cependant aucune ne propose d'aide au choix.

3 METHODOLOGIE

À la suite de la présentation des différentes définitions et concepts, nous allons proposer une méthodologie d'analyse d'aide au choix des solutions IIoT. Le but de cette méthodologie est d'aider les entreprises à choisir les plateformes qui seraient les plus adaptées à leur cas d'usage.

Pour ce fait, trois étapes ont été définies :

1. Définition des critères de choix ;
2. Pondération de la grille de critères : Il s'agit d'attribuer à chaque critère un poids à chaque critère sélectionné ;
3. Choix de la plateforme : nous allons choisir les plateformes dont la pondération est la plus importante.

3.1 Critères de choix

En s'appuyant sur les travaux de cités précédemment, nous pouvons lister cinq critères généraux à prendre en compte lors du choix d'une plateforme IIoT :

- 1- Connectivité
- 2- Gestion d'appareils
- 3- Analyse de données historiques
- 4- Visualisation de données
- 5- Base de données

A cette liste, nous pouvons rajouter deux critères car ensemble ils prendront en compte des aspects tels que la disposition de ressources humains, logiciels et matériels au sein de l'entreprise, et le deuxième la prédiction d'événements de façon immédiate. Par ailleurs, les évolutions récentes tendent à confirmer la nécessité du critère suivant.

- 6- Type de service du cloud computing

Il est important de mettre en évidence qu'en plus des critères généraux, nous pouvons aussi définir des critères spécifiques à la situation rencontrée : Industrie, usine ou ligne de production. Cependant, ces derniers ne seront définis qu'après la réalisation d'une étude minutieuse de la ligne de production dans le but de répondre aux besoins les plus spécifiques de l'entreprise. D'autres critères indirects, mais aussi important à considérer sont : l'aspect financier (déploiement et financement) ainsi que l'aspect sécurité de la plateforme.

3.1.1 Connectivité

Le critère de connectivité va permettre d'identifier les principales caractéristiques du réseau, la communication machine-machine et également machine-plateforme. Nous allons ici nous focaliser sur les protocoles de communication suivants : OPC-UA, CoAP, et le MQTT. Par ailleurs, il est indispensable de recenser les protocoles que les machines de notre ligne de production supportent de façon directe ou indirecte (à l'aide d'une passerelle) sans quoi les machines et les plateformes pourraient ne pas communiquer. Selon

[Zezulka et al., 2018] l'OPC-UA (dont la première version a été lancée en 2006) est fortement recommandé en milieu industrielle car ce protocole de communication se concentre sur la communication avec les équipements industriels et les systèmes de collecte de données. Par ailleurs, ce protocole est librement disponible et implémentable sans restriction ni frais, avec une architecture orientée services. Cependant, des protocoles plus récents tels que MQTT et CoAP sont également de bonnes alternatives en raison de la forte croissance qu'ils connaissent et la part de marché significative qu'ils représentent.

3.1.2 Gestion d'appareils connectés

La gestion d'appareils connectés peut s'avérer une fonctionnalité importante à considérer, lors du choix d'une plateforme IIoT. En effet, indépendamment de la quantité d'appareils dû à être gérés, la plateforme choisie doit être capable de monitorer et contrôler tous les équipements connectés par groupe ou de manière individuelle, de fournir en temps réel une évaluation de la production. La plateforme doit aussi permettre de connecter, déconnecter et configurer les définitions des appareils des appareils connectés.

3.1.3 Analyse de données

Après l'étape de collecte de données, son analyse constitue l'étape qui permettra aux entreprises d'obtenir des informations importantes capables d'aider à la prise de décision, à l'optimisation de leur chaîne de valeur et à mieux répondre aux attentes des clients. Pour ce faire, deux types d'analyses sont disponibles :

- Analyse de données en temps réel : pour en permettre la surveillance et/ou permettre au système d'agir sur son environnement en temps réel et en cas d'anomalies.
- Analyse de données historisées : ce type d'analyse peut permettre aux acteurs industriels d'anticiper les dysfonctionnements du système, en planifiant les interventions nécessaires basées sur les prédictions.

3.1.4 Visualisation de données

La plateforme choisie doit permettre la visualisation des données analysées sur des tableaux de bords personnalisés. Cela permettra un meilleur management visuel des activités.

3.1.5 Base de données

Il serait intéressant de choisir une plateforme avec un cloud ou une base de données intégrée. En effet, les appareils faisant partie d'un réseau IIoT envoient des informations de données relativement petites vers le cloud ou la base de données. Ces données, qui peuvent être de différentes natures vont évoluer en termes de quantité on aura à contrôler des massives données. D'où l'importance d'avoir une base de données intégrée à la plateforme en d'assurer leur disponibilité à un moment donné.

3.1.6 Type de service du cloud computing

Parmi les principaux types de service, nous pouvons citer l'IaaS (infrastructure as a service), le PaaS (platform as a service) et le SaaS (software as a service). Chaque service mentionné possède sa spécificité et serait adapté à une situation particulière. Les plateformes SaaS, permettent par exemple, une mise en place rapide, simple et efficace. Les PaaS sont plus adaptés aux entreprises qui veulent construire sur un réseau existant. L'IaaS est la solution dont le déploiement requiert le plus de temps et d'expertise en informatique, car le fournisseur met seulement à la disposition du client l'infrastructure, libre à lui de gérer l'intégralité de

l'applicatif. Ainsi, le choix du type de service peut être à la fois considérée comme un critère spécifique ainsi qu'un critère général.

3.2 Pondération et choix de la plateforme

Cette étape consiste à attribuer à chaque critère sélectionné (critères généraux et spécifiques) un poids sur une échelle croissante (importance basse, moyenne, haute...) pour les critères dont nous avons plusieurs options de choix. Et également, attribuer un poids pour les critères qui sont présents dans la plateforme (1) ou 0 (pour signaler l'inexistence). A la fin de cette pondération, nous saurons quantifier les plateformes sélectionnées et donc choisir celle qui a le plus de poids, ça veut dire la plus adaptée au cas d'utilisation.

Le tableau ci-dessous, est un exemple de pondération des critères de choix. Dans le cas présent, nous avons considéré chacun des critères comme équivalent en pondération aux autres ; selon le besoin industriel, il conviendra d'ajuster celui-ci en fonction. Nous l'avons réalisé sous base des critères mentionnés dans la partie précédente. La pondération des deux premiers critères (type de service et connectivité) dépend du contexte, mais dans le cadre de notre exemple nous avons idéalisé que la plateforme idéale serait de type SaaS (importance haute : pondération = 3), supportant le protocole de communication OPC-UA (importance haute : pondération = 3) et ou le MQTT (importance moyenne : pondération= 2) et ayant les fonctionnalités basiques de plateformes IIoT (présence obligatoire : pondération = 1).

Tableau 2: Pondération des critères de choix

Critères	Pondération		
	SaaS	PaaS	IaaS
Type de service	3	2	1
Connectivité	OPC-UA 3	MQTT 2	CoAp 1
Analyse de données	OUI 2		NON -
Visualisation de données	1		-
Gestion d'appareils connectés	1		-
Base de données	1		-

3.3 Plateformes évaluées

Dans cette partie, nous allons donner un exemple du chemin à suivre pour l'évaluation et le choix d'une plateforme IIoT.

En raison du nombre de plateformes, avant de se lancer dans l'évaluation des plateformes, nous avons fait une pré-sélection des plateformes qui pourraient convenir à notre cas d'utilisation. Pour ce faire, une étude de marché a été réalisée. Dans cette étude, nous avons pu analyser de manière grossière une trentaine de plateformes, pour n'en retenir que 14.

Notre choix s'est porté vers les plateformes IIoT les plus populaires tels que ThingWorx, Mindsphere, Oracle IoT, Davra IoT, Bosch Suite, Cumulocity, Murano Exosite, Predix, Azure IoT et Thingsboard. En plus, nous en avons sélectionné d'autres moins populaires compte tenu de leur fonctionnalités qui semblaient adaptées aux installations industrielles mais

également de la facilité à trouver des informations sur internet. C'est par exemple le cas des plateformes telles que Nimbits Ixon IoT, Mobius, Device Hive et Eclipse OM2M.

Ensuite, nous les avons évaluées selon l'exemple de pondération et les critères de choix présentés dans le Tableau 2. Ainsi, nous avons pu évaluer toutes les plateformes choisies et en avons sélectionné deux : Thingsboard (en tant que solution open-source – Tableau 3) et Ixon (en tant que solution propriétaire – Tableau 4).

3.3.1 Thingsboard Open source IoT platform

D'après [Thingsboard, 2020], Thingsboard est une plateforme open-source sous licence Apache2.0. Nous pouvons donc l'utiliser gratuitement et l'héberger en tant que solution SaaS ou PaaS. En ce qui concerne ces principales fonctionnalités, la plateforme permet la collecte et le traitement de données, et également la gestion des appareils connectés. Dans cette plateforme, la connectivité des appareils s'effectue via les protocoles IoT standard (Figure 2) tels que MQTT, CoAP, et HTTP. Cette plateforme, dans sa version professionnelle, peut aussi supporter d'autres protocoles de communication telles que l'OPC-UA. Thingsboard combine la scalabilité, la tolérance aux pannes et la performance pour que les données soient bien enregistrées. Enfin, la plateforme prend également en charge les déploiements dans le cloud et sur site.

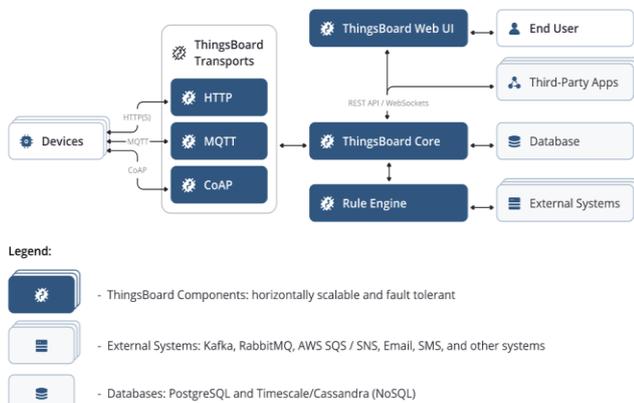


Figure 2. Diagramme de l'architecture de la plateforme Thingworx [Thingworx, 2020]

Tableau 3. Pondération des plateformes open-source

Critères	Plateformes				
	Eclipse OM2M	Device Hive	Thingsboard	Nimbits	Mobius
Type de services	2	2	2	2	2
Connectivité	1	2	3	2	2
Analyse de données	2	1	2	1	2
Visualisation de données	1		1		
Stockage/Base de données	1		1		
Gestion d'appareils connectés	1		1		
Total	8	5	10	5	6

En ce qui concerne le choix d'Ixon, le Tableau 4 montre que trois plateformes pourraient répondre aux critères que nous avons définis, ce sont : Ixon IoT plateforme et Davra IoT et ThingWorx. Notre choix pourrait se porter tant sur Davra que ThingWorx ou Ixon, car ces derniers ont obtenu la même pondération. De ce fait, nous avons deux possibilités : choisir librement l'une des trois plateformes ou rajouter des nouveaux critères pour casser le bris d'égalité. On pourrait par exemple penser aux critères économiques liés aux coûts (politiques de prix), à l'aspect sécurité (authentification de dispositifs, chiffrement de données et la certification de sécurité) ou à la scalabilité. Supposons que notre choix s'est porté sur la plateforme Ixon IoT.

3.3.2 Ixon IoT platform



Figure 3. Caractéristiques de la plateforme Ixon [Ixon, 2020]

Ixon est un SaaS d'objets industriels connectés (Figure. 3) qui permet la gestion d'appareils à distance, le stockage de données, la conversion de données brutes (collectés des appareils) en informations à valeur ajoutée et offre d'autres fonctionnalités aussi importantes.

En effet, l'Ixon IoT compte sur un router combiné à un IoT Gateway qui se connecte à un serveur VPN dans la plateforme Ixon, ce qui pourrait simplifier sa mise en œuvre.

Son intégration permet de gagner du temps, car on n'a pas besoin de savoir coder pour l'intégrer dans une ligne de production. Compte tenu de son type de service (SaaS), tous les aspects concernant la sécurisation, programmation et la mise à jour de la plateforme sont assurés par le constructeur.

4 ETUDE DE CAS : LIGNE AMS

Lorsqu'on veut mettre en place une plateforme IIoT dans une usine, il faut en amont définir les fonctionnalités qu'on veut atteindre. Aussi, en plus de comprendre les technologies existantes, il faut connaître la ligne de production ou usine dans laquelle la solution sera mise en place. Ceci peut paraître tautologique, force est de constater que de nombreuses implémentations de plateformes IoT *ex-nihilo* se soldent par une adéquation avec les actifs physiques.

4.1 Présentation de la ligne AMS

La plateforme AMS (Figure 4) est une ligne de production qui se trouve au sein du DigitaLAB de l'Université de Technologie de Compiègne. La ligne est composée d'un carrousel et son pont de rotation, un tapis roulant, un plateau et d'un robot Kuka.

Tableau 4. Pondération des plateformes propriétaires

Critères	Plateformes									
	Oracle IoT	Bosch IoT	Azure	Murano	Cumulocity	Predix	Ixon	Davra IoT	Thingworx	Mindsphere
Type de services	2	2	3	3	2	2	3	3	3	3
Connectivité	2	2	2	2	3	3	3	3	3	2
Analyse de données	2	2	2	2	2	2	2	2	2	2
Visualisation de données	1	1	1	1	1	1	1	1	1	1
Stockage/Base de données	1	1	1	1	1	1	1	1	1	1
Gestion d'appareils connectés	1	1	1	1	1	1	1	1	1	1
Total	9	9	10	10	10	10	11	11	11	10

4.1.1 Fonctionnement de la ligne

La première étape consiste à charger des pots dans le carrousel. La deuxième étape de la ligne consiste à peser et contrôler des pots dans le plateau étoile. En effet, le plateau étoile est équipé d'un système de pesage des pots qui seront ensuite contrôlés. Le contrôle est effectué en fonction du poids du pot. S'il est trop plein ou pas assez, il est éjecté de la ligne « car cela est la source d'un problème de non-conformité par rapport au critère de qualité défini ». S'il est conforme en termes de qualité, il continue son chemin sur la ligne. La dernière étape palettisation et packaging est effectué en fin de la ligne de production par le robot Kuka

4.1.2 Collecte des données

La plateforme AMS compte actuellement deux automates responsables de faire la collecte des données. D'un côté on a un automate Siemens (S7-1200) pour la collecte des données des équipements qui sont en aval du robot Kuka, et de l'autre on a un automate qui fonctionne en tant que capteur de données du robot Kuka.

Nous allons dans la partie suivante, s'en servir des solutions retenues dans la partie 3 afin d'expliquer comment l'intégration des plateformes choisies seraient faites.

4.2 Intégration de la plateforme Thingsboard dans la ligne

Thingsboard fournit des outils et des API prêts à l'emploi pour réduire considérablement les efforts nécessaires à la gestion d'une ligne de production tel que l'AMS, ce qui se traduit par l'amélioration de la production, ainsi qu'un suivi en temps réel. La connexion des appareils dans le cloud se fera à l'aide du protocole OPC UA (cf. Figure 5). L'automate Siemens et le robot Kuka seront reliés à la plateforme grâce à un processus qui permettra la conversion de données des appareils vers un format accepté par Thingsboard. Ces données seront ensuite stockées sur Cassandra (base de données NoSQL) et pourront être transférées vers des systèmes d'analyse, tels qu'Apache Spark ou Hadoop grâce à Thingsboard Rule Engine. Les données pourront aussi être visualisées en temps réel et les machines contrôlées à distance grâce à Thingsboard Real-Time dashboards.

4.3 Intégration de la plateforme Ixon dans la ligne AMS

Avec IXON on va pouvoir gérer à distance l'ensemble de la ligne AMS, dès la collecte et analyse de données jusqu'à la configuration d'alarmes. L'intégration de cette solution se fait à l'aide d'un router (IXrouter) qui va relier le contrôleur du robot Kuka et l'automate Siemens S7-1200 à l'ordinateur. Ces

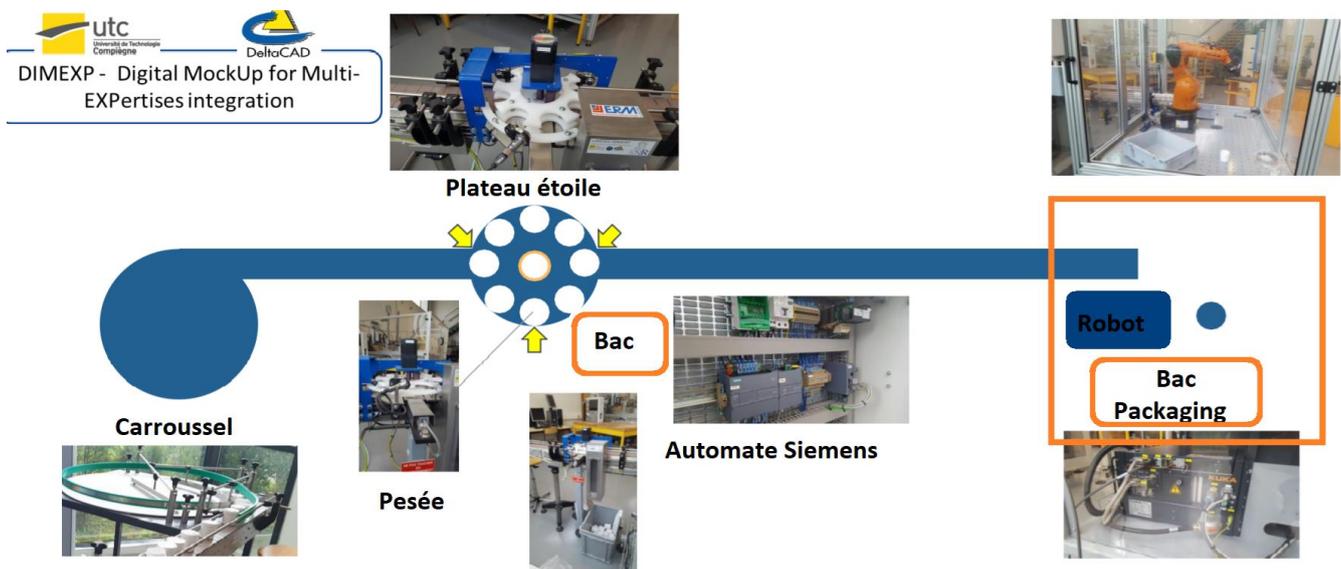


Figure 4: Ligne AMS

derniers seront connectés sur le côté LAN de l'IXrouter. Dès que l'IXrouter a accès à Internet, il va s'enregistrer sur Ixon Cloud (cf. Figure 6).

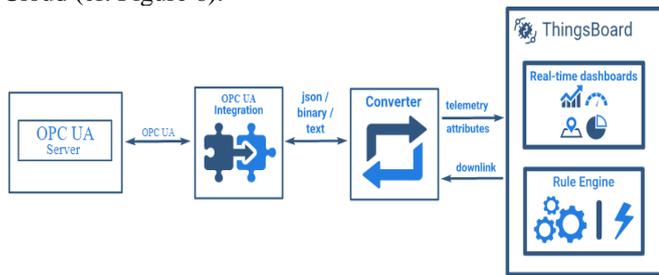


Figure 5. Intégration de la plateforme Thingsboard [Thingsboard, 2020]

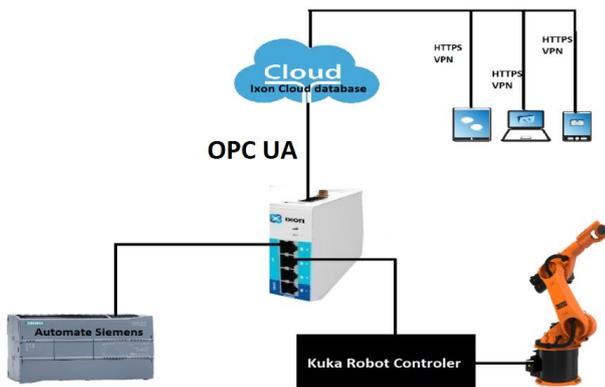


Figure 6. Intégration de la plateforme Ixon

5 DISCUSSION

La proposition ici présentée est un travail d'analyse qui s'adresse à la fois aux experts du domaine (industriels, scientifiques, consultants) et utilisateurs cibles. Ce travail pourrait faire l'objet de quelques améliorations. En effet, les critères choisis pourraient faire l'objet d'approfondissements et de revues croisées avec de multiples experts du domaine afin d'affiner leur description et pertinence. Enfin, le nombre de critères n'est pas suffisant pour différencier la multitude des plateformes existantes, il conviendrait ainsi de proposer et inclure dans les études futures des critères supplémentaires pour l'ensemble des plateformes, notamment la sécurité qui est un aspect important en milieu industriel.

Dans le but d'un travail d'analyse exhaustif, il faudrait recenser l'intégralité des plateformes existantes afin de pouvoir les analyser. Aussi, dans ce cas, le travail est limité par la documentation disponible de la part des éditeurs et de la véracité de ces informations, les éditeurs de logiciels ayant tendance à afficher comme disponible des fonctionnalités qui sont en développement, voire encore en projet.

6 CONCLUSION

A travers la mise en avant d'absence de méthodologie d'aide à la décision pour la sélection de plateforme IoT puis une proposition en ce sens, cet article souhaite contribuer à l'amélioration de l'industrie 4.0 et de ses technologies. Ceci montre que les industries peuvent améliorer leurs environnement grâce à un ensemble de technologies disponibles qui sont en constante évolution.

Vu qu'il existe plus d'une centaine de plateformes IoT sur le marché, il est important d'établir une liste de critères afin de faire un choix judicieux. Ces critères étant primordiaux dans le choix de la plateforme, on peut regretter qu'il n'existe pas

dans la littérature des stratégies et méthodologies pour aider les industries à basculer de l'industrie traditionnelle vers l'industrie du futur.

Nos choix de plateformes se sont portés sur Thingsboard (plateforme open-source) et Ixon (plateforme propriétaire) car les deux répondent bien à ce qu'on attend pour notre ligne de production AMS.

Enfin, il convient de noter que le but des deux plateformes (open-source et propriétaire) reste le même : transformer la façon de produire, améliorer la performance des entreprises, économiser de temps et d'argent.

7 REFERENCES

- Alliance Industrie du futur. (2018). *Guide des technologies de l'Industrie du Futur—Enjeux et panorama des solutions* (Vol. 3). http://www.industrie-dufutur.org/content/uploads/2018/03/Guide-des-Technologies_2018_V3.pdf
- Barrios, P., Danjou, C., & Eynard, B. (2019). *Product lifecycle management et Industrial Internet of Things : Revue systématique de la bibliographie*, CIGI-Qualita, 25-28 Juin 2019, Montréal, Canada
- Bartevyan, L. (2015). Industry 4.0 – Summary report. *DLG-Expert report*, 5/2015, 8.
- Burke, R., Mussomeli, A., Laaper, S., Hartigan, M., & Sniderman, B. (2017). *The smart factory*. https://www2.deloitte.com/content/dam/insights/us/articles/4051_The-smart-factory/DUP_The-smart-factory.pdf
- Choi, H., Song, J., & Yi, K. (2018). Brightics-IoT : Towards Effective Industrial IoT Platforms for Connected Smart Factories. *2018 IEEE International Conference on Industrial Internet (ICII)*, 146-152. <https://doi.org/10.1109/ICII.2018.00024>
- Danjou, C., Rivest, L., & Pellerin, R. (2017). *Douze positionnements stratégiques pour l'Industrie 4.0 : Entre processus, produit et service, de la surveillance à l'autonomie*, CIGI-Qualita, 3-5 Mai 2017, Compiègne, France
- Gartner. (2020). *Magic Quadrant for Industrial IoT Platforms*. Gartner. <https://www.gartner.com/doc/reprints?id=1-2434LPHV&ct=200903&st=sb>
- Hehenberger, P., Vogel-Heuser, B., Bradley, D., Eynard, B., Tomiyama, T., & Achiche, S. (2016). Design, modelling, simulation and integration of cyber physical systems: Methods and applications. *Computers in Industry*, 82, 273-289. <https://doi.org/10.1016/j.compind.2016.05.006>
- IoT Market—Forecasts at a glance—IoT Analytics*. (2014). <https://iot-analytics.com/iot-market-forecasts-overview/>
- Julien, N., & Martin, É. (2018). *L'usine du futur : Stratégies et déploiement : industrie 4.0, de l'IOT aux jumeaux numériques*.
- Kim, M., Lee, J., & Jeong, J. (2019). Open Source Based Industrial IoT Platforms for Smart Factory: Concept, Comparison and Challenges. In S. Misra, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, C. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, & E. Tarantino (Éds.), *Computational Science and Its Applications – ICCSA 2019* (Vol. 11624, p. 105-120). Springer International Publishing. https://doi.org/10.1007/978-3-030-24311-1_8
- Meadows, D. H., & Club of Rome (Éds.). (1972). *The Limits to growth: A report for the Club of Rome's project on the predicament of mankind*. Universe Books.
- Newman, P. (2020). *The Internet of Things Report*. Business Insider Intelligence Research Store.

<https://store.businessinsider.com/products/the-internet-of-things-report>

- ODINI, M. P. (2016). Objets connectés : Enjeu de la 5G Évolution des réseaux M2M () et IoT. *Techniques de l'ingénieur Internet des objets: Vol. base documentaire: TIB612DUO*. (Numéro ref. article : te8000). Editions T.I. <https://www.techniques-ingenieur.fr/base-documentaire/technologies-de-l-information-th9/internet-des-objets-42612210/objets-connectes-enjeu-de-la-5g-te8000/>
- Rajkumar, R., Lee, I., Sha, L. R., & Stankovic, J. (2010). Cyber-physical systems: The next computing revolution. *Proceedings of the 47th Design Automation Conference, DAC'10*, 731-736. <https://doi.org/10.1145/1837274.183741>
- Statista Research Department. (2016, Novembre 17). *IoT: Number of connected devices worldwide 2015-2025*. Statista. <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>
- Thingsboard. (2020). *ThingsBoard—Open-source IoT Platform*. ThingsBoard. <https://thingsboard.io/>
- Ullah, M., Nardelli, J., Molff, A., & Smolander, K. (2020). *Twenty-one key factors to choose an IoT platform: Theoretical framework and its applications* Mehar Ullah, Member, IEEE, Pedro Juliano Nardelli, Senior Member, IEEE, Annika Wolff, Kari Smolander. <https://arxiv.org/pdf/2004.04924.pdf>
- Zezulka, F., Marcon, P., Bradac, Z., Arm, J., Benesl, T., & Vesely, I. (2018). Communication Systems for Industry 4.0 and the IIoT. *IFAC-PapersOnLine*, 51(6), 150-155. <https://doi.org/10.1016/j.ifacol.2018.07.145>
- Ixon. (2020). All-in-one Industrial IoT solution. <https://www.ixon.cloud/all-in-one-industrial-iot-solution/>

CIGI 2021

Analyse de l'hybridation du QFD avec d'autres outils d'aide à la décision

AMENI ELOMRI¹, FATMA LEHYANI¹, ALAEDDINE ZOUARI¹, MICHEL TOLLENAERE²

¹UNIVERSITE DE SFAX, ISGIS, OLID, LR19ES21

3021, Sfax, Tunisie

omriameni3@gmail.com

fatma.lehyani@isgis.usf.tn

ala.zouari@isgis.usf.tn

²UNIVERSITE GRENOBLE ALPES, GRENOBLE INP, G-SCOP

michel.tollenaere@grenoble-inp.fr

Résumé – Depuis l'introduction du Déploiement des Fonctions Qualité (QFD) dans les industries, il a montré de nombreux avantages. Son utilisation permet d'élargir la chance de succès, de fournir des produits de meilleure qualité et de réduire le coût et le temps de développement. Malheureusement, cela ne veut pas dire qu'il n'y a pas beaucoup de difficultés rencontrées lors de son application. Par ailleurs, afin de réduire et traiter les lacunes du QFD, de nombreuses recherches l'ont intégré avec d'autres méthodes analytiques dans le but d'améliorer son déploiement et le rendre plus applicable. Ces intégrations ont entraîné plusieurs extensions et modification du QFD. Dans ce contexte, cet article décrit comment des techniques telles que l'aide à la décision multicritère (MCDM), la logique floue et d'autres méthodes ont pu être combinées avec QFD pour résoudre certains de ses inconvénients.

Abstract – Since the introduction of Quality Function Deployment (QFD) in industries, it has shown many advantages, its use would widen the chance of success, provide better quality products and reduce the cost and development time. But this does not prevent it from being free from difficulties while applying it. So, in order to reduce and address the shortcomings of QFD, a lot of research has integrated it with other analytical methods to improve it and make it more applicable, which led to several extensions and modifications of QFD. In this context, this article describes how techniques such as multi-criteria decision analysis (MCDM), fuzzy logic and others which can be combined with QFD to solve some of its drawbacks.

Mots clés – QFD, Méthodes hybrides, extensions QFD

Keywords – QFD, Hybrid methods, QFD extensions.

1 INTRODUCTION

En tant qu'une méthodologie populaire et pratique qui est centrée sur la satisfaction clients, en intégrant leurs besoins dans le processus de production, de la conception à la fabrication, le Déploiement de la Fonction Qualité (QFD) a attiré l'attention de nombreux chercheurs et a trouvé des diverses applications dans tous les domaines, de la fabrication à des services. En outre, plusieurs propositions de son hybridation avec d'autres outils d'aide à la décision ont été évoquées dans la littérature scientifique, pour l'améliorer et le rendre plus applicable, ce qui a généré des extensions du QFD. Pour répondre aux différents besoins des chercheurs et des praticiens du QFD, Une analyses des différentes méthodes hybridées avec cet outil est menée.

Dans ce cadre, nous avons effectué une revue de la littérature de 281 articles, publiés entre 2005 et 2020, en les classant sur la base de l'année de publication, les pays, les domaines industriels, les méthodes hybrides, etc. La classification et la catégorisation des différents outils ont été choisies sur la base de l'étude de [Abdolshah et Moradi, 2013] et la dernière revue

de la littérature sur les techniques de sélection des fournisseurs menée par [Chai et Ngai, 2020].

Ce document est structuré comme suit : Section 2 donne une présentation générale sur le QFD (historique, définitions, les objectifs, la maison de qualité (HOQ), etc.). La section 3 détaille la méthodologie de recherche considérée pour collecter, classer et analyser les articles révisés. Ensuite, une analyse de toutes les méthodes hybrides avec le QFD et un modèle de classification de ces méthodes sont présenté. Conséquemment, les finalités de cette hybridation sont discutées. Finalement, une conclusion et quelques perspectives sont proposées.

2 PRESENTATION GENERALE SUR LE QFD

2.1 Historique

Après la seconde guerre mondiale, les industries japonaises ont passé du site imitatif au site innovant. En d'autres termes, le concept de développement de produits a évolué de la copie et de l'imitation à un développement de produit basé sur l'originalité. À cette époque, deux principaux problèmes ont

motivé la naissance du déploiement de la fonction qualité (QFD), à savoir un manque de comment faire une qualité de conception et l'utilisation des diagrammes de processus de contrôle de la qualité après la fabrication du produit.

Le QFD a été introduit par Dr.Yoji Akao au Japon, à la fin des années 1960, comme un outil puissant de TQM pour le développement de nouveau produit. Cependant, cette approche reste inconnue et peu populaire par les entreprises jusqu'au 1978, avec la création du tableau de qualité qui relie les besoins des clients en termes de fonctions avec les caractéristiques de qualité dans les chantiers navals de Kobe de Mitsubishi Heavy Industrie. Tout au long de l'année suivante, le QFD a été largement appliqué dans les industries japonaises et il a été introduit aux USA et l'Europe en 1983 [Akao, 1997].

2.2 Définitions du QFD et ses objectifs

2.2.1 Définitions du QFD

Le déploiement de la fonction qualité est la traduction de six caractères japonais « hin shitsu ki no ten kai » où

- hin shitsu - Qualité, attributs ou caractéristiques
- ki no - Fonction ou mécanisation
- ten kai - Déploiement, diffusion, développement

Dans ce sens, le QFD signifie déployer les attributs d'un produit/services désiré par le client dans tous les composants appropriés d'une société [Revelle et al., 1998 ; Zheng et Pulli, 2007]. Cependant, il n'y a pas une définition unique pour le QFD, quatre-vingt et onze manières dont les chercheurs définissent le QFD ont été trouvées [Carnevalli et Miguel, 2008], où ses derniers se concentrent sur les désirs et les attentes des clients « voix de clients ». On peut résumer la définition du QFD comme étant une méthodologie pratiquée en équipe, où les attentes et les exigences des clients ont guidé tout le processus de développement de produits.

2.2.2 Les principaux objectifs du QFD

Le but ultime d'implémenter le QFD est de concevoir un produit de meilleure qualité pour répondre ou surpasser les besoins et les attentes des clients. D'après la littérature, on peut dire que les principaux objectifs d'implémenter le QFD sont :

- Comprendre les exigences des clients ;
- Concevoir un système qualité complet pour la satisfaction du client
- Générer des profits et ajouter de la valeur par la maximisation de la qualité et la satisfaction du client
- Développer des stratégies qui peuvent mettre une entreprise avant ses concurrents

2.3 Les domaines d'applications du QFD

L'industrie de construction navale et l'industrie automobile, telle que Toyota, Ford Motor, General Motors et Fiat Chrysler Automobiles sont les premiers utilisateurs du QFD. Aussi, l'industrie électronique et software sont parmi les premiers qui ont appliqué le QFD. Également, le QFD était largement appliqué dans les industries manufacturières et il était étendu aux industries de services tels que les soins de santé, l'éducation et la recherche, etc. On peut conclure que le champ d'application du QFD est illimité [Costa et al., 2000 ; Chan et Wu, 2002 ; Mehrjerdi, 2010].

2.4 Le processus du QFD

Il existe deux grandes organisations comme sources du QFD, à savoir l'American Supplier Institute (ASI) et GOAL / QPC, qui ont, toutes les deux, développé leurs propres approches pour le QFD [Shillito, 1994 ; Prasad, 1998 ; Garcia Martinez et al., 2007] :

- GOAL/QPC utilise une matrice développée par Akao,

c'est une approche générique basée sur un schéma de 30 matrices ou tableaux de qualité, où chaque matrice détaille un aspect spécifique du processus de développement. Ce modèle est flexible pour une conception complexe.

- ASI utilise une méthode de base à quatre matrices qui est la technique de mise en œuvre du QFD la plus courante. Chaque phase ou matrice représente un aspect plus spécifique des exigences du produit. Les aspects importants de chaque phase sont déployés dans la matrice suivante, comme le montre la figure 1.

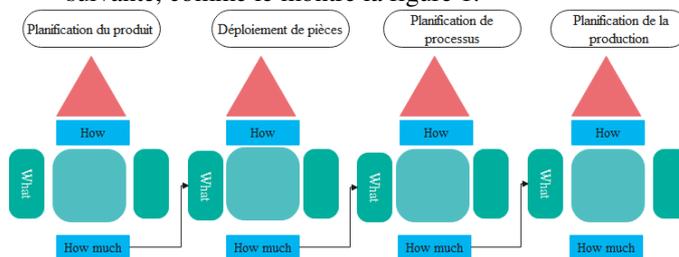


Figure 1 : Les quatre phases du QFD

2.5 Les éléments de HOQ

La maison de la qualité (House Of Quality (HOQ)) représente la première phase du QFD. Cet outil est fondamental et stratégique car il permet de traduire les besoins des clients en mesures techniques appropriées pour répondre à ces besoins. La plupart des études sur le QFD se concentrent principalement sur cette phase et ignore les autres. Les chercheurs et les utilisateurs ont utilisé différents modèles de la HOQ afin de le rendre adaptable aux besoins d'un projet particulier. Une HOQ typique peut être présentée comme l'indique la figure 2.

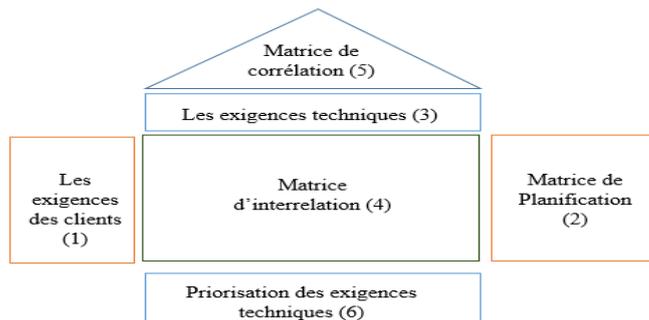


Figure 2 : La maison de qualité

Où :

- Zone 1 : présente les différents besoins clients et leur importance ;
- Zone 2 : analyse concurrentiel pour évaluer la performance du produit de la société productrice et des produits similaires de ses principaux concurrents ;
- Zone 3 : les exigences techniques sont les attributs du produit ou service qui peut être mesurés afin de satisfaire les besoins clients ;
- Zone 4 : la relation entre les exigences techniques (TRs) et les besoins des clients (CNs) pour définir l'impact de chaque (TR) sur chaque CN ;
- Zone 5 : cette matrice permet de déterminer le type de corrélations entre les TRs, soit positif, négatif ou neutre ;
- Zone 6 : priorisation des exigences techniques, c'est le résultat principal du processus HOQ.

2.6 Les avantages et les inconvénients du QFD

Plusieurs recherches telles que [Bouchereau et Rowlands, 2000 ; Costa et al., 2000 ; Chan and wu, 2002 ; Pullman et al., 2002 ; Carnevalli et Miguel, 2008 ; Korayem et Iravani, 2008 ;

Mehrjerdi, 2010 ; Abdolshah et Moradi, 2013 ; Gündoğdu et al., 2020 ; Li et al., 2020 ; Ocampo et al., 2020 ; Pandey, 2020] ont discuté les avantages et les inconvénients du QFD. Ces dernières peuvent être résumées comme suit, voir tableau 1.

Tableau 1: Les avantages et les inconvénients du QFD

Avantages	Inconvénients
Augmente la satisfaction des clients	Se caractérise par la subjectivité Présente les opinions des clients et des experts en des valeurs numériques précis
Meilleures opportunités de communication entre les équipes	Sa mise en œuvre est complexe et prend du temps (la grande taille de matrice)
Réduction du temps de conception	QFD traditionnel ignore la corrélation entre les besoins clients. La corrélation entre les exigences de conception n'est pas prise en compte dans la hiérarchisation de ces derniers.
Raccourcir le délai de mise sur le marché	L'application du QFD pour la conception de produits alimentaires durables est limitée

3 METHODOLOGIE DE RECHERCHE

Dans notre travail, nous présentons une revue et une analyse de la littérature sur le QFD étalée de l'année 2005 jusqu'à 2020. De plus, une recherche bibliographique approfondit a été menée via des bases de données telles que Science direct et Google scholar pour chercher des articles étudiant la thématique du QFD dans leurs titres, leurs résumés, leurs mots clés et leurs méthodologies de recherche. Plus de deux cent quatre-vingts articles sont collectés à partir des journaux scientifiques populaires.

L'objectif principal de notre étude est de mener une analyse catégorique des méthodes et des techniques hybridés ou utilisées conjointement avec le QFD.

4 RESULTATS

Premièrement, nous avons analysé les articles selon les nombres des approches hybridés avec QFD. La figure 3 donne des informations précieuses sur la méthodologie adoptée par les chercheurs lors d'utilisation du QFD.

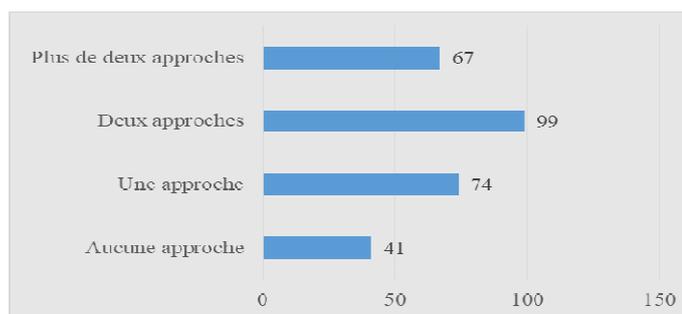


Figure 3. Nombre d'approches hybridées avec QFD

D'après la littérature analysée, 41 articles (15%) utilisent seulement le QFD et ne comptent pas sur son hybridation avec d'autres méthodes sur un total de deux cent quatre-vingt-un publications (281), tels que les travaux de [Marsot, 2005 ;

Chakraborty et Dey, 2007 ; Sørensen et al., 2010 ; De Pelsmaeker et al., 2015 ; Lam et Bai, 2016 ; Riesener et al., 2019 ; Duru et al., 2020] etc. 74 recherches (26%) combinent une seule approche avec le QFD, par exemple la méthode AHP, Goal programming, DEA, théorie des ensembles flous, ANP, Kano, TRIZ, etc. [Iranmanesh et Thomson, 2008 ; Ramanathan et Yunfeng, 2009 ; Lam et Lai, 2015 ; Lima-Junior et Carpinetti, 2016 ; Mao et al., 2019 ; Palominos et al., 2019 ; Li et al., 2020]. L'hybridation du QFD avec deux autres méthodes est la plus utilisée avec 35%, on peut citer, par exemple la combinaison AHP et ANP avec le QFD pour résoudre le problème de la sélection des emplacements des installations [Patrovi, 2006]. A propos l'utilisation de plus de deux approches avec le QFD, nous avons trouvé 67 articles, telle que la combinaison d'un ensemble des méthodes MCDM et la logique floue avec le QFD pour la conception et le développement de produits alimentaires durables. Par la suite, nous avons fait une analyse plus détaillée sur les différentes méthodes utilisées, afin de dégager l'objectif de l'hybridation de chacune avec le QFD. La figure 4 présente les différents outils et techniques, combinés avec le QFD, trouvés dans les articles analysés.

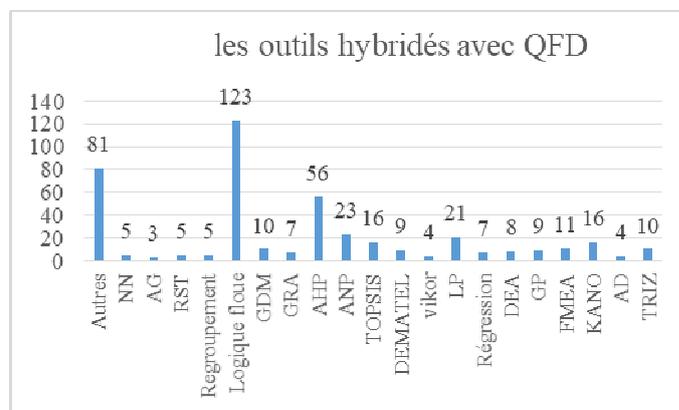


Figure 4 : Les différents outils et techniques combinés avec le QFD

Comme l'indique la figure 4, la logique floue et ses extensions est la technique la plus utilisée avec le QFD (123 articles). En deuxième rang viennent les méthodes classées comme « autres » (81 articles), par exemple chaîne de Markov [Asadabadi, 2017], les analyses fonctionnelles et comportementales [Renaud et al., 2019], Pareto [Erdil, 2019], etc. La troisième méthode est l'AHP (56 articles), puis ANP (23 articles), par la suite les programmations (linéaire ou non linéaire) avec 21 articles. Le sixième rang est partagé entre le modèle Kano et la méthode TOPSIS avec 16 articles, en suite TRIZ avec 10 articles, etc. Par la suite, nous avons groupé ces différents outils en des groupes capitaux, comme l'indique le tableau 2, en se basant sur l'étude de [Abdolshah et Moradi, 2013] pour faciliter l'analyse et l'interprétation.

Tableau 2 : Les principaux groupes des outils hybridés avec le QFD

Groupes capitaux	Nombre d'articles
Logique Floue et ses extensions	123
MCDM	102
DMAI	15
Programmations mathématiques	41
Modèles Avancés	40
Autres	81

Les six principaux groupes des méthodes hybridées avec le QFD ont été divisés en quatre sous-groupes comme le montre la figure 5.

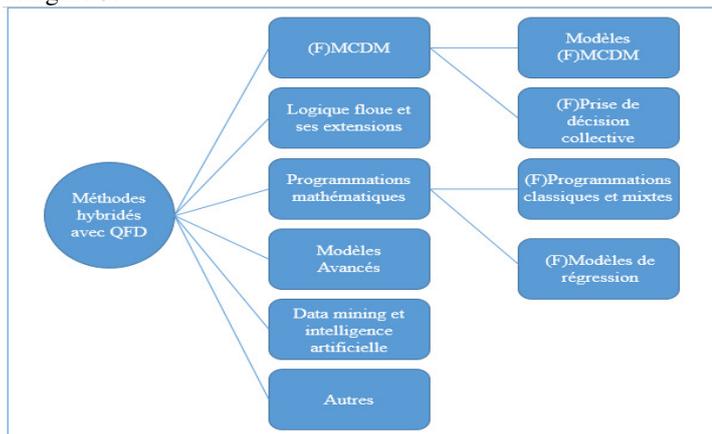


Figure 5 : Catégorisation des méthodes et techniques

5 DISCUSSION

5.1 Méthode de Prise de décision collective

Nous avons identifié 10 articles où les chercheurs ont combiné la méthode de prise de décision collective avec le QFD. D'après notre analyse bibliographique, nous pouvons dire que la raison principale de l'intégration de (F) GDM dans le QFD est afin de gérer les informations d'entrée imprécises et vagues dans cet outil (Les opinions des clients qui sont exprimés en différents formats) et faciliter le processus de la prise de décision.

5.2 Les méthodes d'Aide à la décision multicritère

5.2.1 Processus de Hiérarchie Analytique (AHP)

La méthode AHP est la méthode la plus utilisée de MCDM avec le QFD. Parmi les limites majeures du QFD traditionnel est sa faiblesse concernant la priorisation des éléments de HOQ, la priorisation des besoins clients et les caractéristiques techniques [Zhang et al., 2014 ; Bhattacharya et al., 2005 ; Wu et al., 2017 ; Efe, 2019b]. Dans les travaux de [Kubler et al., 2016], une revue sur FAHP a été réalisée et ils ont dégagé deux objectifs lors de l'hybridation (F)AHP avec QFD qui sont, soit pour déterminer l'importance des besoins clients, soit pour représenter la relation entre les besoins clients et les caractéristiques de conception.

[Chowdhury et al., 2020] ont développé un modèle FAHP-QFD multi-phase axé sur la durabilité pour déterminer le portefeuille de fournisseurs optimal avec les quantités de commande optimales correspondantes en examinant simultanément les changements dans les exigences de durabilité des parties prenantes. Ils ont construit trois HOQ afin d'étudier l'impact de la nature dynamique des attentes de durabilité des parties prenantes sur la détermination d'un portefeuille optimal de fournisseurs et les quantités de commandes optimales correspondantes. Dans ce cas, le FAHP a été utilisé pour hiérarchiser les exigences de durabilité des parties prenantes. [Chin et al., 2019] ont présenté un ensemble des techniques pour une meilleure compréhension des besoins des passagers dans le but d'améliorer la qualité de service de TGV et d'augmenter la satisfaction des passagers. Ils ont combiné l'AHP à valeurs d'intervalle avec le QFD pour refléter l'incertitude et le flou de l'évaluation subjective et pour hiérarchiser les besoins des passagers.

D'après les articles analysés, le but d'utiliser l'AHP dans le processus QFD est, généralement pour prioriser les exigences des clients et déterminer leurs importances, voir figure 6.

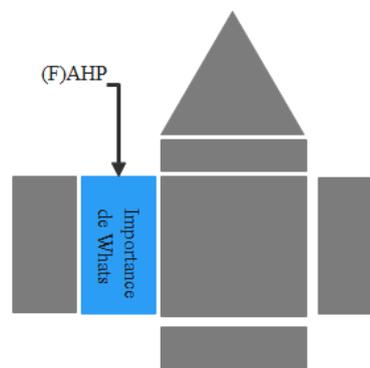


Figure 6 : le rôle de l'AHP dans le QFD

5.2.2 Processus de réseau analytique (ANP)

Le QFD traditionnel définit les degrés d'importance des besoins des clients sans prendre en compte leurs interdépendances, à cause de la difficulté de cette matrice. Aussi, la corrélation entre les exigences de conception n'est pas prise en compte dans leur hiérarchisation [Chan and Wu, 2005 ; Efe, 2019a ; Gündoğdu et Kahraman , 2020). L'ANP, voir figure 7, peut résoudre ce problème, grâce à sa capacité d'incorporer des relations interdépendantes entre des éléments de différents niveaux ou à l'intérieur de niveaux.

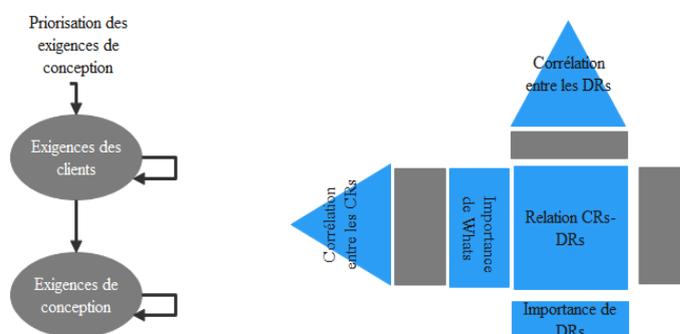


Figure 7 : Hybridation de l'ANP avec le QFD

Dans ce cadre, [Wey et Chiu, 2013] ont affirmé que le processus d'analyse QFD peut être formulé comme un modèle ANP. Ils ont proposé une approche qui combine le QFD et l'ANP pour comprendre les besoins des piétons et déterminer les exigences techniques alternatives (ATR) pour le développement axé sur les besoins des citoyens urbains afin d'améliorer la conception de l'environnement de la marche. L'approche ANP a été appliquée pour déterminer les relations d'interdépendance entre les besoins des citoyens urbains et les ATR, ainsi que la dépendance intérieure entre eux.

5.2.3 Les méthodes de compromis

Nous avons identifié 16 articles sur l'intégration de la Technique pour l'Ordre de Préférence par similarité de Solution Idéale (TOPSIS) avec le QFD et seulement 4 articles sur la combinaison QFD-VIKOR (Solution d'optimisation et de compromis multicritères). Par exemple, [Gündoğdu et Kahraman, 2020] se sont concentrés sur la conception d'une technologie de robots du point de vue de l'utilisateur dans un environnement flou sphérique. Ils ont développé un modèle qui intègre le QFD avec une nouvelle extension de la logique floue, appelée la logique floue sphérique, pour surmonter l'incertitude des opinions linguistiques et TOPSIS pour obtenir un classement final des alternatives. Afin d'analyser et évaluer les risques de sécurité opérationnelle dans la construction navale, [Efe, 2019a] a utilisé le QFD et VIKOR dans un environnement flou, où la réduction des accidents de travail sont considérés comme exigences clients, et les modes de

défaillance présentent les exigences de conception. La méthode VIKOR basée sur les nombres flous intuitionnistes qui intègre les trois classements des paramètres de criticité et obtient un classement unique. Elle classe les modes de défaillance liés aux accidents de travail dans la construction navale et détermine le plus important.

D'après la littérature, l'objectif de l'hybridation de TOPSIS avec le QFD est le classement des alternatives et la détermination de la meilleure. Pour la méthode VIKOR, elle est généralement utilisée avec le QFD afin d'améliorer son résultat et déterminer le classement final des exigences de conception (DRs).

5.2.4 Autres méthodes MCDM

a. Laboratoire d'essai et d'évaluation de la prise de décision (DEMATEL)

Dans cette revue, nous avons compté 9 articles en rapport avec la combinaison du QFD avec DEMATEL. En général, afin d'améliorer la précision de résultat du QFD, l'intégration de l'outil DEMATEL permet de traiter la faiblesse lors de la détermination de relation de corrélation entre les attributs du client (CA) et les exigences techniques (TR). Cette technique est capable de visualiser les interrelations complexes entre les facteurs interdépendants et d'établir une relation de cause à effet entre eux.

Dans l'étude de [Ocampo et al., 2020], DEMATEL flou est utilisé pour comprendre et déterminer les interrelations et les poids des paramètres de décision dans chaque phase. Toutefois, [Wu et al., 2017] ont proposé une approche hybride pour traiter les lacunes et les limites de l'approche QFD traditionnel. La méthodologie proposée pour atteindre cet objectif combine DEMATEL et VIKOR pour le QFD dans un environnement flou hésitant et teste l'efficacité de cette approche dans l'industrie Automobile. La théorie d'ensemble flou hésitant (HFSs) est utilisée pour traiter le flou et l'imprécision des jugements des experts, capter l'hésitation des experts. DEMATEL Hésitant a été utilisé pour analyser les relations d'interdépendance entre besoins et déterminer leur poids. Hésitant VIKOR pour déterminer les relations entre les éléments de HOQ et classer les caractéristiques d'ingénierie.

b. La théorie grise (Grey theory) :

Deng a introduit la théorie grise en 1982. En général, un système avec des informations complètes est appelé « blanc », au contraire, un système où ses informations ne sont pas disponibles, le système est appelé « noir ». Lorsqu'un système se présente avec des informations incomplètes ou incertaines, le système est appelé « gris ».

L'objectif principal derrière l'application de la théorie des ensembles gris au QFD est de résoudre les problèmes des informations incomplètes (peu de données) et incertaines [Liu et Cheng, 2016 ; Chen et Chou, 2011] ont utilisé l'analyse relationnelle grise (GRA) au sein du QFD pour identifier objectivement la force de la relation entre les techniques d'amélioration et les besoins des lecteurs.

5.3 La logique floue est ses extensions

Le QFD est basé sur les opinions et les jugements des clients et des experts. Ces divers données linguistiques constituées les inputs de processus QFD et sont supposées être précises et sont traitées comme des données numériques précises. La manière où le QFD traditionnel traite ces données est caractérisé par la subjectivité, l'incertitude et l'imprécision, ce qui influence la prise de décision. Afin de traiter cet inconvénient et réduire l'incertitude dans le processus QFD, la théorie de la logique floue peut être utilisée.

La logique floue est capable d'évaluer les décisions subjectives, par la transformation des variables linguistiques en des valeurs floues. [Bottani et Rizzi, 2006] ont proposé une approche basée sur le QFD et la logique floue pour aider les entreprises à déterminer les actions stratégiques importantes afin d'améliorer ses performances logistiques.

Cette théorie permet de transformer les données linguistiques exprimées par les clients en nombres flous, qui sont supposées être précises et sont traitées comme des données numériques dans le QFD traditionnel. [Khademi-Zare et al., 2010] ont développé deux modèles de priorisation basés sur le QFD flou pour aider la Communication Mobile de l'Iran (MCI) de comprendre les attentes de ses clients et déterminer les plus importants, puis hiérarchiser les actions stratégiques pour répondre à leurs attentes.

Un aperçu des applications particulières de logique floue dans le QFD est donné dans les recherches de [Chan et Wu, 2002 ; Abdolshah et Moradi, 2013 ; Mehrjerdi, 2010]. Malgré la capacité de la logique floue à modéliser l'incertitude, mais ce dernier présente plusieurs inconvénients, plusieurs chercheurs ont essayé d'améliorer cette approche en proposant d'autres extensions, comme montre la figure 8.

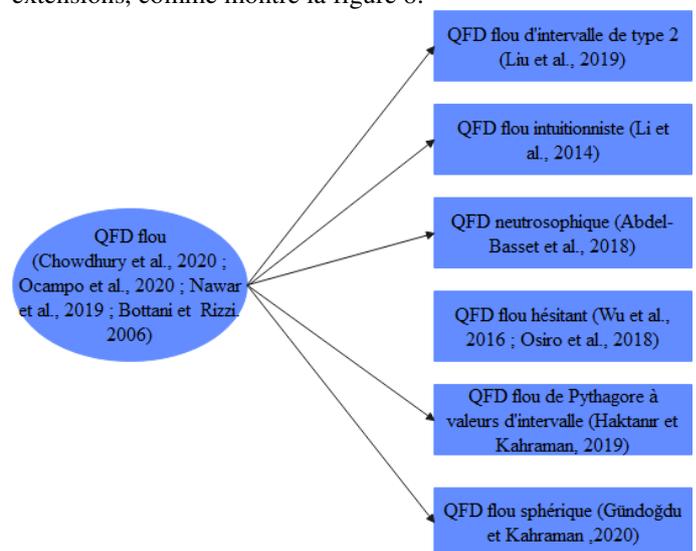


Figure 8 : QFD flou et ses extensions

5.4 Les programmations mathématiques

Nous pouvons diviser ce groupe en deux sous-groupe, telles que les techniques des programmations classiques et les modèles de régression.

5.4.1 Les programmations classiques

Les modèles de programmation linéaire, non linéaire ou mixte sont les plus utilisés et fusionnés avec le QFD afin de l'optimiser les caractéristiques d'ingénierie en respectant un ensemble des contraintes. Ils traitent les inconvénients du QFD pour élargir ses objectifs et non seulement la maximisation de satisfaction clients [Chen et Ko, 2009 ; Zhaoling et al., 2008 ; Chen et Ko, 2010 ; Delice et Güngör, 2009].

5.4.2 Les modèles de régression (linéaire ou non-linéaire)

Ces modèles sont généralement fusionnés avec le QFD pour identifier les relations fonctionnelles vagues et inexacts entre les exigences du client et les caractéristiques d'ingénierie en raison de ses avantages d'objectivité et de réalité [Fung et al., 2006 ; Sener et Karsak, 2011].

5.5 Les modèles avancés

Parmi les méthodes avancées les plus utilisés avec le QFD, nous pouvons mentionner TRIZ, FMEA, le modèle Kano et

La conception axiomatique.

5.5.1 TRIZ

Il est indiqué dans notre étude que la plupart des recherches ont tendance à ignorer la matrice de corrélation entre les exigences techniques, qui est située au-dessus du HOQ, en raison de problèmes de coût, de temps et d'évaluation, ce qui peut beaucoup affecter le résultat [Wu et al., 2017 ; Efe, 2019a ; Gündoğdu et al., 2020]. La raison principale d'utilisation de TRIZ avec le QFD est pour résoudre le problème de contradictions des caractéristiques techniques [Haktanir et Kahraman, 2019 ; Li et al., 2020] ont proposé une approche nommée PQT dans le but de déterminer les techniques requises pour la conception de système d'alarme de navire du point de vue des besoins des utilisateurs. TRIZ a été intégré avec le QFD pour traduire les exigences de la conception afin de développer des solutions innovantes.

5.5.2 FMEA

Le but ultime de la combinaison QFD-FMEA dans le développement du nouveau produit ou l'amélioration des produits existants ou des processus est la puissance de cet outil à évaluer les causes, les effets et les éléments critiques de système. Cette combinaison peut apporter des avantages tels que la prévention des défauts, l'amélioration de la sécurité et l'augmentation de la satisfaction des clients [Korayem et Iravani, 2008 ; Almannai et al., 2008 ; Liu, 2009 ; Hassan et al., 2010 ; Chen et Ko, 2010].

5.5.3 Modèle Kano

La combinaison du QFD avec Kano vise à exploiter la force de ce modèle dans l'identification et la compréhension des besoins des clients en les considérant comme attributs et en les regroupant dans des catégories afin de les prioriser.

Vue que la conception esthétique est très importante pour le succès d'un produit sur le marché, [Yadav et al., 2013] ont proposé une approche qui combine le QFD et le modèle flou Kano pour prioriser les attributs esthétiques du profil de la voiture pour obtenir une plus grande satisfaction du client.

5.6 Data mining et l'intelligence artificielle

Les Méthodes Datamining et l'Intelligence Artificielle (DMAI) peuvent être utilisées aux différentes étapes du processus QFD, grâce à leurs capacités à résoudre des problèmes complexes [Abdolshah et Moradi, 2013].

[Bouchereau et Rowlands, 2000] ont discuté dans leur article la capacité des réseaux de neurones (NN) à surmonter la faiblesse du QFD en termes de traitement d'une grande quantité de données d'entrée. [Siraj et al., 2008] ont développé un modèle de prévision QFD à l'aide de NN et des techniques statistiques pour la sélection des spécifications des machines. Ils ont utilisé la technique NN pour surmonter la faiblesse du QFD dans les jugements subjectifs des relations entre les paramètres. [Li et Kuo, 2007] ont développé une nouvelle méthode QFD améliorée (EQFD) pour concevoir les éléments virtuels de jeux de rôle en ligne massivement multi-joueurs (MMORPG). Cette étude intègre EQFD et un réseau neuronal chaotique génétique afin d'exploiter les données dans la base de données client du MMORPG et déterminer les besoins des utilisateurs et l'importance de chaque exigence.

Aussi, parmi les méthodes de data mining utilisées avec le QFD, nous trouvons le concept des ensembles approximatifs. Le but principal d'introduire cette approche avec le QFD est de gérer l'imprécision et l'incertitude qui est exprimé par la région limite non vide, au contraire de la théorie des ensembles flous, cette approche évite la sélection subjective des fonctions d'appartenance [Yang et al., 2011].

5.7 Autres

[Moubachir et Bouami, 2015] ont proposé une approche qui facilite la transition entre les phases QFD. Cette approche est basée sur les analyses morphologiques pour traiter les corrélations entre les Spécification de Conception (DSs). Pour la résolution des difficultés de cette méthode, les auteurs ont utilisé la théorie des graphes et ils ont représenté la relation entre les DS en tant que problème de graphes.. Cette démarche a été simulée en utilisant le langage de programmation Python et Tkinter comme interface utilisateur graphique. [Asadabadi, 2017] a appliqué la méthode de chaîne de Markov avec le QFD pour résoudre le problème de sélection des fournisseurs. La chaîne de Markov génère un modèle de priorités des besoins clients (CNs), suit l'évolution des priorités des besoins des clients et cherche un modèle pour eux. Ce modèle est l'input de l'ANP-QFD dans le but de trouver les priorités des exigences du produit (PRs). Ensuite, ces PRs sont utilisés pour classer les fournisseurs.

5.7.1 La décomposition en cascade inverse du QFD

Ce nouveau type du QFD est proposé par [Duan et al., 2020] pour construire l'arbre de pertinence des accidents de qualité, afin d'identifier et analyser les causes de qualité possibles et profondes aux étapes de conception et de production. Cette nouvelle extension du QFD implique trois HOQs, où la première phase lie les caractéristiques d'accident de qualité aux caractéristiques physiques. La deuxième permet de traduire ces caractéristiques en des paramètres de conception. La dernière phase transforme ces paramètres en exigences fonctionnelles. La nouveauté de ce type est que le « Qui » se place en haut de la HOQ et « Comment » sera à gauche.

5.7.2 QFD Inverse (Reverse QFD ; R-QFD)

La figure 9 présente un modèle de formulation d'une stratégie de fabrication.

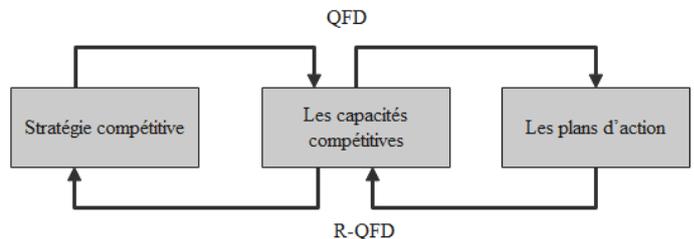


Figure 9 : Un modèle de formulation de stratégie de fabrication

De gauche à droite, le QFD est adopté pour formuler les stratégies manufacturières en transformant les exigences de la stratégie concurrentielle en importance des capacités concurrentielles, puis en importance des plans d'action. Pour réaliser le processus de rétroaction, une nouvelle méthode QFD inverse (R-QFD) est élaborée afin d'analyser comment les plans d'action contribue à l'amélioration des capacités [Wang et al., 2020].

5.7.3 Blitz QFD

Blitz QFD a été introduit pour la première fois par Zultner au milieu des années 1990 [Zultner, 1995]. L'objectif de Blitz QFD est de traiter quelques limites du QFD traditionnel. Il offre un moyen plus simple et plus rapide, et par la suite permet de raccourcir le temps de cycle du développement de produits et d'identifier les exigences des clients les plus importantes. La nouveauté de ce modèle est que la voix des

affaires est la première étape au lieu de voix des clients [Nawar et al., 2019].

Afin de faire face aux problèmes environnementaux et répondre aux demandes émergentes de développement du produit/service écologique, les chercheurs ont essayé d'introduire des nouvelles extensions du QFD classique afin d'incorporer les exigences de la durabilité. Parmi ces extensions, nous pouvons mentionner : QFDE, ECO-QFD, Green QFD, Green QFD II, QFD multi-phases [Chowdhury et al., 2020 ; Popoff et Millet, 2017 ; Chowdhury et Quaddus, 2016 ; Wood et al., 2016 ; Younesi et Roghanian, 2015 ; Bereketli et Genevois, 2013 ; Kuo et al., 2009 ; Utne, 2009].

6 CONCLUSION

Dans cet article, nous fournissons une analyse de la littérature et une classification des méthodes et techniques d'aide à la décision combinés avec le QFD. De plus, nous expliquons les buts et les finalités des hybridations. Six principaux groupes de méthodes et techniques pour soutenir le QFD ont été présentés. En terme des résultats, nous avons constaté que : (i) Les méthodes MCDM sont les plus utilisées avec le QFD, en particulier la méthode FAHP pour bien hiérarchiser les besoins des clients et les caractéristiques techniques objectivement. Cependant, la méthode FANP est implémentée pour résoudre et faciliter le problème de contradiction entre les exigences clients et les exigences de conception, (ii) La majorité des articles ont tenté de combiner la logique floue et ses extensions avec le QFD afin de fournir un cadre quantitatif pour évaluer le processus de prise de décision subjective, c-à-d. surmonter et faire face à l'imprécision et l'incertitude des opinions des experts et des clients. (iii) La plupart des recherches se contentent sur la construction de la première phase du QFD, appelée la planification de produit. (iv) Le nombre des articles des revues sur le QFD qui ont analysé son hybridation avec les méthodes d'aide à la décision est très réduit.

7 REFERENCES

- Abdel-Basset, M., Manogaran, G., Mohamed, M., & Chilamkurti, N. (2018). Three-way decisions based on neutrosophic sets and AHP-QFD framework for supplier selection problem. *Future Generation Computer Systems*, 89, 19-30.
- Abdolshah, M., & Moradi, M. (2013). Fuzzy quality function deployment: an analytical literature review. *Journal of Industrial Engineering*, 2013. <http://dx.doi.org/10.1155/2013/682532>
- Akao, Y. (1997). QFD: Past, present, and future. *International symposium on QFD*, 97(2), 1-12
- Almannai, B., Greenough, R., & Kay, J. (2008). A decision support tool based on QFD and FMEA for the selection of manufacturing automation technologies. *Robotics and Computer-Integrated Manufacturing*, 24(4), 501-507.
- Asadabadi, M. R. (2017). A customer based supplier selection process that combines quality function deployment, the analytic network process and a Markov chain. *European Journal of Operational Research*, 263(3), 1049-1062.
- Bereketli, I., & Genevois, M. E. (2013). An integrated QFDE approach for identifying improvement strategies in sustainable product development. *Journal of Cleaner Production*, 54, 188-198.
- Bhattacharya, A., Sarkar*, B., & Mukherjee, S. K. (2005). Integrating AHP with QFD for robot selection under requirement perspective. *International journal of production research*, 43(17), 3671-3685.
- Bottani, E., & Rizzi, A. (2006). Strategic management of logistics service: A fuzzy QFD approach. *International journal of production economics*, 103(2), 585-599.
- Bouchereau, V., & Rowlands, H. (2000). Methods and techniques to help quality function deployment (QFD). *Benchmarking: An International Journal*, 7(1), 8-20
- Carnevali, J. A., & Miguel, P. C. (2008). Review, analysis and classification of the literature on QFD—Types of research, difficulties and benefits. *International Journal of Production Economics*, 114(2), 737-754.
- Chai, J., & Ngai, E. W. (2020). Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead. *Expert Systems with Applications*, 140, 112903.
- Chakraborty, S., & Dey, S. (2007). QFD-based expert system for non-traditional machining processes selection. *Expert Systems with Applications*, 32(4), 1208-1217.
- Chan, L. K., & Wu, M. L. (2005). A systematic approach to quality function deployment with a full illustrative example. *Omega*, 33(2), 119-139.
- Chan, L. K., & Wu, M. L. (2002). Quality function deployment: A literature review. *European journal of operational research*, 143(3), 463-497.
- Chen, L. H., & Ko, W. C. (2009). Fuzzy linear programming models for new product design using QFD with FMEA. *Applied Mathematical Modelling*, 33(2), 633-647.
- Chen, L. H., & Ko, W. C. (2010). Fuzzy linear programming models for NPD using a four-phase QFD activity process based on the means-end chain concept. *European Journal of Operational Research*, 201(2), 619-632.
- Chen, Y. T., & Chou, T. Y. (2011). Applying GRA and QFD to improve library service quality. *Journal of Academic Librarianship*, 37(3), 237-245.
- Chen, Z. S., Martínez, L., Chang, J. P., Wang, X. J., Xiong, S. H., & Chin, K. S. (2019). Sustainable building material selection: A QFD-and ELECTRE III-embedded hybrid MCGDM approach with consensus building. *Engineering Applications of Artificial Intelligence*, 85, 783-807.
- Chin, K. S., Yang, Q., Chan, C. Y., Tsui, K. L., & Li, Y. L. (2019). Identifying passengers' needs in cabin interiors of high-speed rails in China using quality function deployment for improving passenger satisfaction. *Transportation Research Part A: Policy and Practice*, 119, 326-342.
- Chowdhury, M. M. H., & Quaddus, M. A. (2016). A multi-phased QFD based optimization approach to sustainable service design. *International Journal of Production Economics*, 171, 165-178.
- Chowdhury, M. M. H., Paul, S. K., Sianaki, O. A., & Quaddus, M. A. (2020). Dynamic sustainability requirements of stakeholders and the supply portfolio. *Journal of Cleaner Production*, 255, 120148.
- Costa, A. I. A., Dekker, M., & Jongen, 4. 7. 4. W. (2000). Quality function deployment in the food industry: a review. *Trends in food science & Technology*, 11(9-10), 306-314.
- De Pelsmaeker, S., Gellynck, X., Delbaere, C., Declercq, N., & Dewettinck, K. (2015). Consumer-driven product development and improvement combined with sensory analysis: A case-study for European filled chocolates. *Food quality and preference*, 41, 20-29.
- Delice, E. K., & Güngör, Z. (2009). A new mixed integer linear programming model for product development using quality function deployment. *Computers & Industrial Engineering*, 57(3), 906-912.
- Duan, P., He, Z., He, Y., Liu, F., Zhang, A., & Zhou, D. (2020). Root cause analysis approach based on reverse cascading decomposition in QFD and fuzzy weight ARM for quality accidents. *Computers & Industrial Engineering*, 147, 106643.
- Duru, O., Galvao, C. B., Mileski, J., Robles, L. T., & Gharehgozli, A. (2020). Developing a comprehensive approach to port performance assessment. *Asian Journal of Shipping and Logistics*. In press <https://doi.org/10.1016/j.ajsl.2020.03.001>
- Efe, B. (2019a). Analysis of operational safety risks in shipbuilding using failure mode and effect analysis approach. *Ocean Engineering*, 187, 106214.
- Efe, B. (2019b). Fuzzy cognitive map based quality function deployment approach for dishwasher machine selection. *Applied Soft Computing*, 83, 105660.
- Erdil, A. (2019). An Evaluation on Lifecycle of Products in Textile Industry of Turkey through Quality Function Deployment and Pareto Analysis. *Procedia Computer Science*, 158, 735-744.
- Garcia Martinez, M., de Carlos, P., de Felipe, I., Briz, J., Morais, F., & Navarro, M. (2007). Quality Function Deployment: Can Improve Innovation Efficiency in the Food Industry? *1st International European Forum on Innovation and System Dynamics in Food Networks*. Innsbruck-Igls, Austria 15-17 February 2007, pp. 23-33.
- Gündoğdu, F. K., & Kahraman, C. (2020). A novel spherical fuzzy QFD method and its application to the linear delta robot technology development. *Engineering Applications of Artificial Intelligence*, 87, 103348.
- Haktanir, E., & Kahraman, C. (2019). A novel interval-valued Pythagorean fuzzy QFD method and its application to solar photovoltaic technology development. *Computers & Industrial Engineering*, 132, 361-372.
- Hassan, A., Siadat, A., Dantan, J. Y., & Martin, P. (2010). Conceptual process planning—an improvement approach using QFD, FMEA, and ABC methods. *Robotics and Computer-Integrated Manufacturing*, 26(4), 392-401.

- Iranmanesh, H., & Thomson, V. (2008). Competitive advantage by adjusting design characteristics to satisfy cost targets. *International Journal of Production Economics*, 115(1), 64-71.
- Khademi-Zare, H., Zarei, M., Sadeghieh, A., & Owlia, M. S. (2010). Ranking the strategic actions of Iran mobile cellular telecommunication using two models of fuzzy QFD. *Telecommunications Policy*, 34(11), 747-759.
- Korayem, M. H., & Iravani, A. (2008). Improvement of 3P and 6R mechanical robots reliability and quality applying FMEA and QFD approaches. *Robotics and Computer-Integrated Manufacturing*, 24(3), 472-487.
- Kubler S, Robert J, Derigent W, Voisin A, & Le Traon Y. (2016). A state-of-the-art survey & testbed of fuzzy AHP (FAHP) applications. *Expert Systems with Applications*; 65: 398- 422.
- Kuo, T. C., Wu, H. H., & Shieh, J. I. (2009). Integration of environmental considerations in quality function deployment by using fuzzy logic. *Expert systems with applications*, 36(3), 7148-7156.
- Lam, J. S. L., & Bai, X. (2016). A quality function deployment approach to improve maritime supply chain resilience. *Transportation Research Part E: Logistics and Transportation Review*, 92, 16-27.
- Lam, J. S. L., & Lai, K. H. (2015). Developing environmental sustainability by ANP-QFD approach: the case of shipping operations. *Journal of Cleaner Production*, 105, 275-284.
- Li, F., Chen, C. H., Lee, C. H., & Khoo, L. P. (2020). A user requirement-driven approach incorporating TRIZ and QFD for designing a smart vessel alarm system to reduce alarm fatigue. *Journal of Navigation*, 73(1), 212-232.
- Li, M., Jin, L., & Wang, J. (2014). A new MCDM method combining QFD with TOPSIS for knowledge management system selection from the user's perspective in intuitionistic fuzzy environment. *Applied soft computing*, 21, 28-37.
- Li, S. G., & Kuo, X. (2007). The enhanced quality function deployment for developing virtual items in massive multiplayer online role playing games. *Computers & Industrial Engineering*, 53(4), 628-641.
- Lima-Junior, F. R., & Carpinetti, L. C. R. (2016). A multicriteria approach based on fuzzy QFD for choosing criteria for supplier selection. *Computers & Industrial Engineering*, 101, 269-285.
- Liu, A., Xiao, Y., Lu, H., Tsai, S. B., & Song, W. (2019). A fuzzy three-stage multi-attribute decision-making approach based on customer needs for sustainable supplier selection. *Journal of Cleaner Production*, 239, 118043.
- Liu, H. T. (2009). The extension of fuzzy QFD: From product planning to part deployment. *Expert Systems with Applications*, 36(8), 11131-11144.
- Liu, H. T., & Cheng, H. S. (2016). An improved grey quality function deployment approach using the grey TRIZ technique. *Computers & Industrial Engineering*, 92, 57-71.
- Mao, Q., Li, N., & Peña-Mora, F. (2019). Quality function deployment-based framework for improving the resilience of critical infrastructure systems. *International Journal of Critical Infrastructure Protection*, 26, 100304.
- Marsot, J. (2005). QFD: a methodological tool for integration of ergonomics at the design stage. *Applied Ergonomics*, 36(2), 185-192.
- Mehrjerdi, Y. Z. (2010). Quality function deployment and its extensions. *International Journal of Quality & Reliability Management*. 27(6), 616-640
- Moubachir, Y., & Bouami, D. (2015). A new approach for the transition between QFD phases. *Procedia Cirp*, 26, 82-86.
- Nawar, E. A., Backar, S. H., & El-Dardiry, M. A. (2019). Case Study on Fuzzy Blitz Quality Function Deployment. *Archive, The Academic Research Community publication*, 2(3), 256-264.
- Ocampo, L. A., Labrador, J. J. T., Jumao-as, A. M. B., & Rama, A. M. O. (2020). Integrated multiphase sustainable product design with a hybrid quality function deployment–multi-attribute decision-making (QFD-MADM) framework. *Sustainable Production and Consumption*, 24, 62-78.
- Osiro, L., Lima-Junior, F. R., & Carpinetti, L. C. R. (2018). A group decision model based on quality function deployment and hesitant fuzzy for selecting supply chain sustainability metrics. *Journal of Cleaner Production*, 183, 964-978.
- Palominos, P., Quezada, L. E., & Gonzalez, M. A. (2019). Incorporating the voice of the client in establishing the flexibility requirement in a production system. *International Journal of Production Economics*, 211, 34-43.
- Pandey, M. M. (2020). Evaluating the strategic design parameters of airports in Thailand to meet service expectations of Low-Cost Airlines using the Fuzzy-based QFD method. *Journal of Air Transport Management*, 82, 101738.
- Partovi, F. Y. (2006). An analytic model for locating facilities strategically. *Omega*, 34(1), 41-55.
- Popoff, A., & Millet, D. (2017). Sustainable life cycle design using constraint satisfaction problems and Quality Function Deployment. *Procedia Cirp*, 61, 75-80.
- Prasad, B. (1998). Review of QFD and related deployment techniques. *Journal of manufacturing Systems*, 17(3), 221-234.
- Pullman, M. E., Moore, W. L., & Wardell, D. G. (2002). A comparison of quality function deployment and conjoint analysis in new product design. *Journal of Product Innovation Management*, 19(5), 354-364.
- Ramanathan, R., & Yunfeng, J. (2009). Incorporating cost and environmental factors in quality function deployment using data envelopment analysis. *Omega*, 37(3), 711-723.
- Renaud, J., Houssin, R., Gardoni, M., & Armaghan, N. (2019). Product manual elaboration in product design phases: Behavioral and functional analysis based on user experience. *International Journal of Industrial Ergonomics*, 71, 75-83.
- Revelle, J. B., Moran, J. W., & Cox, C. A. (1998). *The QFD handbook*. John Wiley & Sons.
- Riesener, M., Rebentisch, E., Doelle, C., Kuhn, M., & Brockmann, S. (2019). Methodology for the Design of Agile Product Development Networks. *Procedia CIRP*, 84, 1029-1034.
- Sener, Z., & Karsak, E. E. (2011). A combined fuzzy linear regression and fuzzy multiple objective programming approach for setting target levels in quality function deployment. *Expert Systems with Applications*, 38(4), 3015-3022.
- Shillito, M. L. (1994). *Advanced QFD: linking technology to market and company needs*. John Wiley & Sons.
- Siraj, F., Nordin, N., & Yusoff, N. (2008). Quality function deployment analysis based on neural network and statistical results. *Int'l J. Simulation: Systems, Science and Technology*, 9(2), 73-81.
- Sivasamy, K., Arumugam, C., Devadasan, S. R., Muruges, R., & Thilak, V. M. M. (2016). Advanced models of quality function deployment: a literature review. *Quality & quantity*, 50(3), 1399-1414.
- Sørensen, C. G., Jørgensen, R. N., Maagaard, J., Bertelsen, K. K., Dalgaard, L., & Nørremark, M. (2010). Conceptual and user-centric design guidelines for a plant nursing robot. *biosystems engineering*, 105(1), 119-129.
- Utne, I. B. (2009). Improving the environmental performance of the fishing fleet by use of Quality Function Deployment (QFD). *Journal of Cleaner Production*, 17(8), 724-731.
- Wang, J., Wu, H., & Chen, Y. (2020). Made in China 2025 and manufacturing strategy decisions with reverse QFD. *International Journal of Production Economics*, 224, 107539.
- Wey, W. M., & Chiu, Y. H. (2013). Assessing the walkability of pedestrian environment under the transit-oriented development. *Habitat International*, 38, 106-118.
- Wood, L. C., Wang, C., Abdul-Rahman, H., & Abdul-Nasir, N. S. J. (2016). Green hospital design: integrating quality function deployment and end-user demands. *Journal of Cleaner Production*, 112, 903-913.
- Wu, X., Nie, L., & Xu, M. (2017). Robust fuzzy quality function deployment based on the mean-end-chain concept: Service station evaluation problem for rail catering services. *European Journal of Operational Research*, 263(3), 974-995.
- Yadav, H. C., Jain, R., Shukla, S., Avikal, S., & Mishra, P. K. (2013). Prioritization of aesthetic attributes of car profile. *International Journal of Industrial Ergonomics*, 43(4), 296-303.
- Yang, M., Khan, F. I., Sadiq, R., & Amyotte, P. (2011). A rough set-based quality function deployment (QFD) approach for environmental performance evaluation: a case of offshore oil and gas operations. *Journal of Cleaner Production*, 19(13), 1513-1526.
- Younesi, M., & Roghanian, E. (2015). A framework for sustainable product design: a hybrid fuzzy approach based on quality function deployment for environment. *Journal of Cleaner Production*, 108, 385-394.
- Zhang, F., Yang, M., & Liu, W. (2014). Using integrated quality function deployment and theory of innovation problem solving approach for ergonomic product design. *Computers & Industrial Engineering*, 76, 60-74.
- Zhaoling, L., Qisheng, G., & Dongling, Z. (2008). Product design on the basis of fuzzy quality function deployment. *Journal of Systems Engineering and Electronics*, 19(6), 1165-1170.
- Zheng, X., & Pulli, P. (2007). Towards high quality mobile services for Senior citizens in smart living environments. In *IEEE International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*, 26-28 April 2007, Seoul, Korea, pp. 164-170.
- Zultner, R. E. (1995). *Blitz QFD: Better, faster, and cheaper forms of QFD*. American Programmer, 8, 24-24.

CIGI 2021

Quality of sawmilling output predictions according to the size of the lot - The size matters!

VINCENT MARTINEAU^{1,2}, JONATHAN GAUDREAU^{1,2}, MICHAEL MORIN^{1,3}, STEVE VALLERAND⁴

¹ FORAC Research Consortium

² Department of Computer Science and Software Engineering

³ Department of Operations and Decision Systems

Université Laval, Québec, QC, Canada

vincent.martineau.1@ulaval.ca

jonathan.gaudreault@ift.ulaval.ca

michael.morin@osd.ulaval.ca

⁴ FPInnovations, Québec, QC, Canada

steve.vallerand@fpinnovations.ca

Résumé – Lors de l'évaluation de modèles d'apprentissage automatique supervisé, on considère généralement le rendement de prédiction moyen obtenu sur les tests individuels comme mesure de choix. Toutefois, lorsque le modèle est destiné à prédire quels produits du bois seront obtenus lors du sciage de certains billots, c'est généralement la performance pour un lot complet qui importe. Dans cet article, nous montrons l'impact de cette nuance en termes d'évaluation du modèle. En fait, la qualité d'une prédiction (globale) s'améliore considérablement lorsque l'on augmente la taille des lots, ce qui offre un solide soutien à l'utilisation de ces modèles en pratique.

Abstract – When comparing supervised learning models, one generally considers the average prediction performance obtained over individual test samples. However, when using machine learning to predict which lumber products will be obtained when sawing logs, it is usually the performance over the entire lot that matters. In this paper, we show the impact of this by evaluating a model performance for various batch sizes. The quality of a (global) prediction improves tremendously when batch size increases, which offers a strong support for the use of such models in practice.

Mots clés - Simulation de débitage, évaluation de modèles d'apprentissage supervisé, application d'apprentissage automatique, industrie des produits forestiers.

Keywords – Sawing Simulation, Supervised Learning Models Evaluation, Machine Learning Application, Forest products industry.

1 INTRODUCTION

For many reasons, the forest product industry needs forecasting what its inventory would allow producing. There exist needs for performance forecasting, configuration decisions and wood allocation between several plants. It has been shown that accurately forecasting sawmills output can improve profitability due to better decision-making [Morin et al. 2020]. Tools for such evaluations include sawmill simulators, such as the Optitek [FPInnovations, 2014]. These tools are designed to simulate the transformation process of a log at a given plant [Lin et al. 2011]. However, simulation is computationally expensive with large volume of woods, for companies owning multiple plants or when testing multiple sawmill configurations. As a result, supervised learning has been recently proposed to improve or replace slow simulators by a fast relation from the space of the logs characteristics to the space of the sawmill outputs, i.e. lumbers [Morin et al. 2015].

Morin et al. (2020) showed that, although supervised learning models produced for that particular problem are only

approximate, they remain more than useful for decision-making, e.g. to make decision regarding wood allocation between sawmills. In this paper, we explain why seemingly low performance on individual logs does not hinder a models' usefulness for decision-making. As demonstrated by our experiments, the errors in the prediction for some given logs are partially compensated by the errors on other logs as the batch size increases. In addition, sophisticated models, i.e. models trained using a machine learning procedure, proved useful and more accurate than a simple average for all batch sizes and especially for smaller batches. This explains the good performance of such models for decision-making, especially on the wood allocation problem [Morin et al. 2020]. It makes these models especially promising in the specific use cases where a simulator tends to be too slow, i.e. on large amount of data.

The paper is organized as follows. We first review the related literature and concepts in Section 2. In Section 3, we present the experimental process and data supporting our contributions, interpret the precision, the recall and the F1-score performance

measure in the context of our machine learning problem, and present the results. We conclude in Section 4.

2 RELATED CONCEPTS AND LITERATURE

In the following subsections, we introduce general notions regarding sawing simulation as well as supervised learning for sawing simulation.

2.1 Sawing Simulators

Sawing simulators make it possible to anticipate which products a log will generate when processed at a given sawmill. These tools generally take three inputs, namely a sawmill model (described in a suitable formal language), a feasible product list, and a virtual description of the log.

The *sawmill model* details the equipment, the links between the machines as well as their capacities. The *feasible product list* describes the lumbers that can be produced by the factory along with their value. Log sawing by-products, such as sawdust and shavings, are also associated with values. A product value can either be based on its true market value, or it can reflect its importance w.r.t. the current needs of the company. Finally, a *virtual log* is a representation of a log that needs to be cut, often a three-dimensional scan. It is generally modeled as a point cloud approximating the surface of the log [Thomas 2013].

Simpler representations, such as *parametric descriptions*, might be used in some studies. Parametric descriptions are vectors of characteristics such as the length of the log, its strong diameter, its weak diameter, its curvature, its tapering, and its volume. Other characteristics, such as the log species, could be introduced in the parametric description.

Given a sawmill model, a feasible product list and a virtual log, the sawing simulator performs a simulation of all the feasible cuts to choose the one leading to an optimal yield for this log, i.e. the simulator finds a set of lumbers maximizing the total value of the transformed log [Todoroki 1990]. An example of a sawing simulation result can be seen in Figure 1.

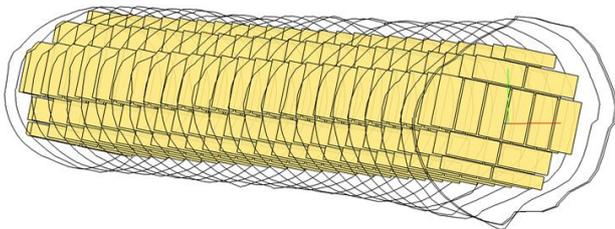


Figure 1. A virtual log and the optimal basket of products obtained by the sawmill simulation.

2.2 Supervised Learning for Sawing Simulation

Although there exist some machine learning model using the three-dimensional log scans as input (e.g. [Selma 2018]), we focus in this paper on models using the parametric description of a log. Both types of models provide as output a vector of counts for each feasible lumber products. This vector of counts describes the *basket of products* one should obtain if processing the log at the modeled sawmill.

Supervised learning algorithms aims to build a mapping from a feature vector space to a label space using a training set of examples consisting of pairs of inputs and outputs. In our context, the features are the logs parametric descriptions, and

the labels are the vectors of counts representing the baskets of products.

In supervised learning, there are two main types of output labels, namely discrete (and finite), and continuous. The former leads to what we call a classification problem whereas the latter leads to a regression problem. In what follows, we describe how basket of products, which are vectors of counts, can be encoded as labels in both the classification and the regression cases.

2.2.1 Classification for Sawing Simulation

A class is a discrete label. For classification purposes in the sawing simulation context, it is possible to consider each individual basket of products encountered in the training set, i.e. combination of lumber products, as a specific class.

The drawback of this approach is that each specific combination needs to be encountered in the data. Therefore, it is possible that some feasible combinations of products, i.e. some feasible baskets, cannot be predicted by the model. One example would be a class that is present in the test dataset and not in the training dataset. Conceptually, one of the main benefits of the approach is that all the predictions are guaranteed to be feasible for at least one existing log (as they have been seen at least once in the data).

2.2.2 Regression for Sawing Simulation

When performing a regression in our context, a regression model individually predicts the number of products of each feasible lumber from the feasible product list (i.e. each position in the vector count).

Contrary to classification it is now possible to predict any basket of products even if they were not part of the train set. For each of the feasible product, the model provides a real number. The drawback of the approach is that there are no such things as decimal lumber counts in practice. This might be a problem when using the model for applications that need predictions for individual logs. However, in many applications, there is few need to obtain precise counts and we aim at evaluating a global prediction, i.e. the prediction on a batch of logs. As a result, we are using the real values provided by the model directly to compute the performance on a batch of logs.

This approach comes with some challenges. One way to address the problem would be to create one model per product output. The problem with this approach is that it assumes independent outputs. In the case of our application, individual entries in the count vector are related to one another. That is, each predicted product might take the space needed for another. For that reason, we used a single model trained to predict the entire vector of counts. The specific procedure depends on the chosen learning algorithm.

2.2.3 Learning Algorithms for Sawing Simulation

In the literature, multiple learning algorithms have been evaluated in the context of sawing simulation. In this paper, we focus on three machine learning algorithms that were proved efficient for that problem by Morin et al. (2015), namely k-nearest-neighbors (KNN) [Fix and Hodges Jr. 1951], decision tree [Breiman 1996], and random forest [Breiman 2001]. We experiment with both the regression and the classification version of these algorithms.

2.2.3.1 k-Nearest-Neighbors

The KNN algorithm is based on the idea that similar logs should have similar product baskets. During training, the algorithm will

use a data structure that makes it easy to compare the distance between two logs. In this study, we use the Euclidian distance between the parametric description of two logs. The algorithm chose the k nearest logs.

In classification mode the algorithm performs a majority vote to determine the class whereas in regression mode it averages the basket of the nearest k logs. Although we tested multiple values of k , a value of one allowed us to achieve good results in training while maintaining a good generalization for examples that were not seen in the training phase. Recent applications of KNN has shown that it performs well on unbalanced data, e.g. [Cai 2020]. Since our dataset is highly unbalanced, it made the algorithm a promising choice in the context of our application.

2.2.3.2 Decision Tree

The decision tree learning algorithm uses the training examples to generate a tree-based model where the nodes are simple decision rules based on the input features. For prediction, an unseen example is passed through the tree and the model output is determined when it reaches a leaf.

In classification mode, each leaf represents the class to be predicted which is determined by the examples the training set that reached that leaf during training (by a majority vote). In the case of regression, we use a variation of the algorithm to predict multiple output values. Although, it would be possible to create a set of independent trees, product counts in the basket are dependent. We therefore used a single tree that can predict the entire basket at once such as described in [Borchani 2015]. Recent applications of the decision tree algorithm assessed it as a simple, efficient, and easy to interpret model [Lan 2020].

2.2.3.3 Random Forest

The random forest algorithm builds multiple decision trees. It does so by sampling the training set N times. Each sample is used to create a single tree [Schonlau 2020]. For prediction, the data are passed through each tree and the predictions are combined. In the case of classification, a majority vote is performed whereas in regression the trees outputs are averaged.

3 EXPERIMENTS

Our experiments are intended to show the impact of the size of the lot on the quality of the prediction for the prediction models using the KNN, the decision tree and the random forest learning algorithms.

In this section, we present the industrial data used for this study (3.1), the details of the model-building phase (3.2), and the evaluation metrics we use to compare the quality of the models (3.3). Finally, we analyze and compare the performance of our models on three aspects. First, by exploring the average prediction quality on individual logs (3.4). Second, by comparing the performance on batches of increasing size (3.5). Third, by highlighting the most common source of error in predictions (3.6).

3.1 Data

Data was provided by FPInnovations. Our dataset contains a total of 2235 logs. All the logs from this dataset are known to produce a non-empty basket of products. To train the supervised learning models, we used the following log characteristics: length, strong diameter, weak diameter, curvature, taper, and volume.

We used the Optitek sawing simulator [FPInnovations, 2014] to virtually transform each of these 2235 logs and obtain the actual basket of products. A total of 85 different sawing products are present in those vectors leading with 85 entries.

We obtained 1188 different baskets of products which leads to 1188 classes when considering the classification version of the problem. Considering the high number of classes relative to the number of examples in our dataset, it is expected that several classes will have a low representation. For example, there are roughly 30% of classes that are only present once in the dataset while one of the classes is present 58 times. This particularity of the dataset implies that certain classes in the training set will not be present in the test set (and vice-versa).

This class distribution is particularly complex for classification models because it is expected that these models do not try predicting classes that have not been seen in training. However, models should predict similar classes that are present at training. In this situation even if the classes are different, it is possible that these two classes have several common elements.

It is important to mention that the logs were not chosen for the purpose of having a balanced input or output. This means that some lengths or diameters are more common than others and some only appear once. The same thing is true for the basket of products, hence the classes.

Figure 2 shows a histogram of the length of the logs in the complete dataset. The figure allows us to see that some lengths are much more frequent and that some lengths are very rare and have only 1 example (5.43 meters). Furthermore, we see that

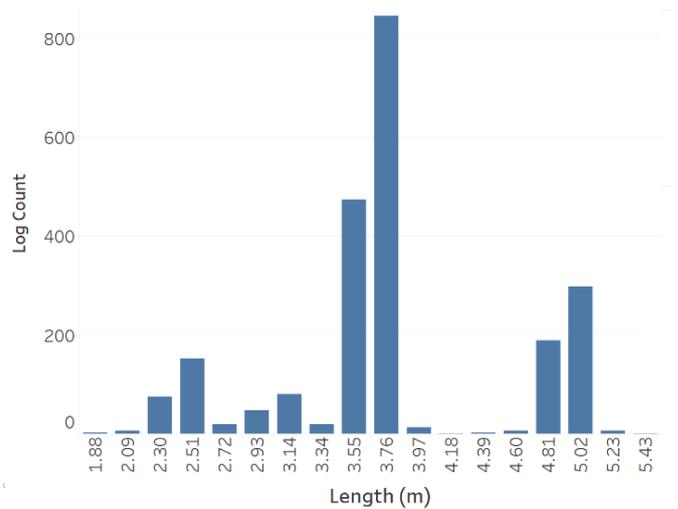


Figure 2. Distribution of logs by length (m) in the dataset

around 800 logs have a length of 3.76 meters. This distribution is expected to complicate the task of fitting a model on the data.

3.2 Model Training

To carry out the experiments, the logs in the dataset is randomly separated so that 85% of the logs are used for training, i.e. to fit a model, and 15% are used to form a test dataset to evaluate the performance of the model on unseen data. We created 10 such partitions which will allow the experiments to be repeated 10 times for each of the algorithms.

For all splits, we first trained models using the KNN, the decision tree, and the random forest algorithms on the training set. Furthermore, to establish a base case reference, we also consider a *dummy* model which systematically predicts the

average basket of products as seen in the train set. It should be noted that the dummy model prediction procedure is akin to simple forecasting heuristics such as using the historical average of the production.

3.3 Performance Measure

To evaluate the performance of the learned models, we need to assess how the basket of products they predict (vector y) compares to the actual basket of products (vector b) generated by the simulator. Vectors y and b may either correspond to the quantities obtain from the sawing of a given log or of a whole batch of logs.

In what follows, we use the well-known *precision*, *recall* and *F1-score* metrics to evaluate the models [Van Rijsbergen 1979, Powers 2020]. We recall that precision and recall are calculated using the number of true positives (TP), the number of false positives (FP), and the number of true negatives (TN) [Van Rijsbergen 1979]. In our context, we define TP , FP , and TN as follows:

- TP represents the number of lumbers that are present in both vectors.
- FP is the number of lumbers predicted but not actually produced.
- FN is the number of lumbers that were manufactured but not predicted.

Given a predicted basket y and an actual basket b , the precision is defined as

$$precision(y, b) = \frac{TP(y, b)}{TP(y, b) + FP(y, b)},$$

and the recall is defined as

$$recall(y, b) = \frac{TP(y, b)}{TP(y, b) + FN(y, b)}.$$

Finally, the F1 score is computed as the harmonic mean between precision and recall which leads to the following equation:

$$F1(y, b) = \frac{2TP(y, b)}{2TP(y, b) + FP(y, b) + FN(y, b)}.$$

It should be noted that by evaluating a set of models built using a learning algorithm and comparing their average results to the average results of models built using another learning

algorithms we are able to determine which algorithm generates the best models either for individual logs or for a batch of logs.

3.4 Performance on Individual Logs

This section is used to demonstrate the quality of our prediction for individual logs. In the case of individual log predictions, we present the performance metrics for individual predictions averaged on a batch of unseen logs and multiple replications where a replication consists of the following three steps (as described in Section 3.2). First, we partition the dataset into a training set and into a test set. Second, we use each learning algorithm to build a model based on the data of the training set. Third we evaluate each of the generated models on the test set by using it to make a prediction of the basket of each individual log it contains. The quality of the prediction for individual logs in terms of a specific metric, e.g. the F1-Score, is then averaged.

Figure 3 reports the average F1-Score for the different models we are using in this research. We are comparing the decision tree, KNN and random forest models to the dummy predictor.

All models perform much better than the dummy prediction. This result can be explained by the large variety of feasible basket of products in our dataset and their uneven distribution. This characteristic of the dataset proves to be problematic for the dummy model.

Among all the tested learning algorithms, KNN shows the worst performance. As it was shown by Selma et al. (2015), similar logs could have very different basket of products which hinders nearest neighbors approaches efficiency. Although, the difference between KNN and the dummy predictor is 28%.

Tree-based machine learning algorithms (decision tree and random forest) performed best. This is in line with the result of Morin et al. (2015). What is interesting is that the regression and classification models have similar performance. This result highlights the possibility to achieve good prediction even if the training dataset does not contain at least one example of each feasible basket of products.

Using regression also provides an advantage over classification. For our experiments, the number of classes is defined by the number of different basket of products present in our dataset. One problem with classification is that small training datasets are very likely not to contain all feasible classes. As a result, a classification model output is likely to be wrong on at least some of the products in the baskets. In regression, this is less of an issue since the model can output an unseen basket.

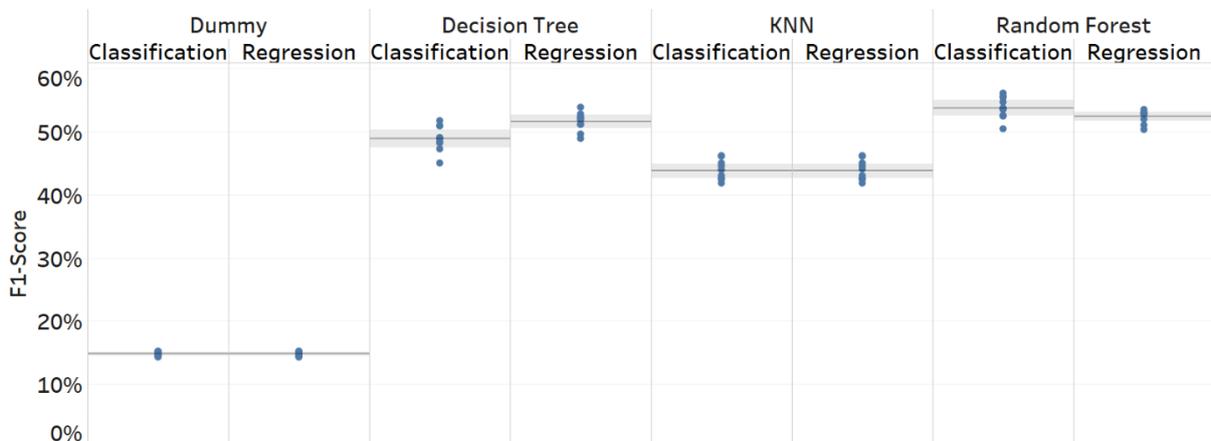


Figure 3. Average F1-score for individual predictions on the test set (95% confidence intervals; 10 replications; individual replication scores are shown in blue)

Finally, the model with the best performance is the random forest in classification. In this case the algorithm is performing 39% better than the dummy predictor.

3.5 Performance for Batches

In this section, we explore the effect of the size of the batches on the quality of the forecasts. We also validate our initial hypothesis that the errors made by a model when predicting the output of a sawmill for an unseen log is compensated for by errors made on other logs. Finally, we confirm that sophisticated models are useful for all tested batch sizes and especially on smaller batches where a simple average does not achieve a sufficient performance.

Figure 4 reports the F1-score for the prediction of the global basket of products associated with the whole test dataset. We can see a significant improvement for all models. All models improve by at least 40% in terms of F1-score. The dummy predictor has the largest improvement (approx. 76%). This result is not surprising because with a large enough sample the average should improve to be closer to a perfect prediction. This establishes a base case for the other algorithms.

It should be noted that the random forest algorithm in classification is no longer the best performer and the accuracy of its prediction is only 89%. Not only is this model no longer the best performer, but it is now our worst performer. KNN has a F1-score of about 90%, which is lower than that of the dummy predictor. The best performer is random forest in regression with a F1-score of 93%.

Figure 5 illustrates more precisely how the performance is affected by the size of the batch. We kept the same test dataset but report global results for randomly chosen subsets of different sizes.

We notice that the performance curve is logarithmic. The batch size does not need to be huge to see dramatic performance improvement. This renders the models useful for a wide range of applications involving small batches of logs such as log piles, e.g. for wood allocation purposes. Companies can therefore achieve interesting results in decision-making on such problems using such models although they appeared to be only approximate for individual predictions.

In classification all models perform better than dummy for smaller batch size; the smaller the batch is, the greater the

difference. For example, decision tree will perform 42% (in terms of F1-score) better than dummy for batch size of 1. At batch size of 50, we can see that dummy is beating the KNN algorithm and will do better than random forest at batch size of 110. With our experiment dummy never reach the level of prediction of the decision tree, but with a larger test dataset it is possible that dummy would yield better result than the decision tree classifier.

For regression, the picture is different. Again, for small batches all our algorithms perform better than the dummy predictor and again the dummy predictor performs better than KNN (starting at a batch size of 50). However, this time decision tree and random forest are doing better than dummy for batch size of 330. Not only do these models perform better, but random forest has a prediction that is 3% over the dummy in terms of F1-score.

This clearly supports the fact that the regression version of random forest dominates the other approaches and would be a safe choice whatever is the size of the lot.

Figure 6 details the F1-score, precision, and recall metrics according to the size of the batch for the best classification model (decision tree) and the best regression model (random forest). Recall and precision are relatively close to each other regardless of the size of the lot. For both models, recall and precision tends to both increase with batch size. However, we can point that the difference is much smaller for decision tree in classification mode. The distance between precision or recall and the F1 score is of about 0.2%. For random forest in regression mode, the spread is larger, and the difference is slightly over 1%. We also notice that the difference tends to be larger for smaller batch sizes than for larger batch sizes.

Figure 7 shows the 95% confidence intervals around the F1-score average for the various batch size. The spread of the 95% confidence interval diminishes significantly as batch size increases. This is especially true for small batches, e.g. when increasing the size of the batch from 1 to 10 logs. Of course, for small batches, it is possible to make a prediction that is wrong that has a large impact on the overall prediction for the batch. The fact that the interval is narrowing as the size of the batch increases confirms our hypothesis that errors tend to be canceled.

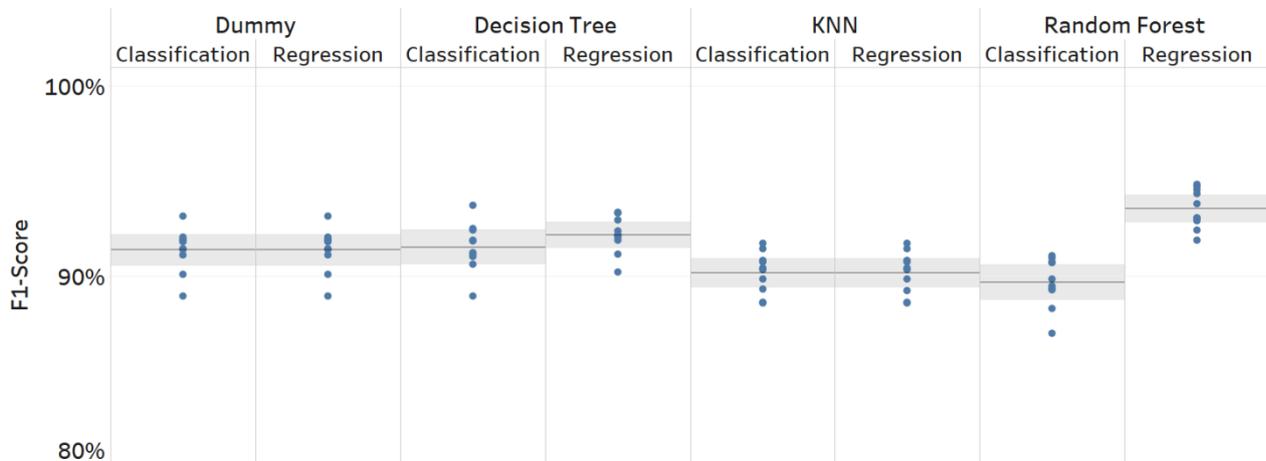


Figure 4. Average F1-score for global predictions on the test set along (95% confidence intervals; 10 replications; individual replication scores are shown in blue)

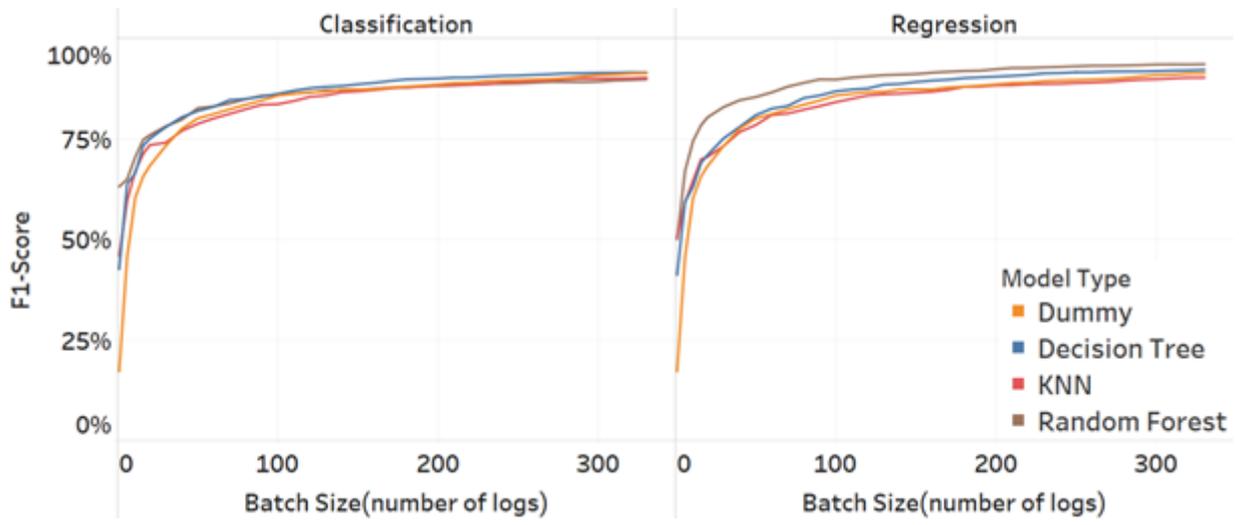


Figure 5. F1-score (in %) for KNN, Decision Tree, Random Forest, and Dummy as a function of the size of the batch (number of logs)

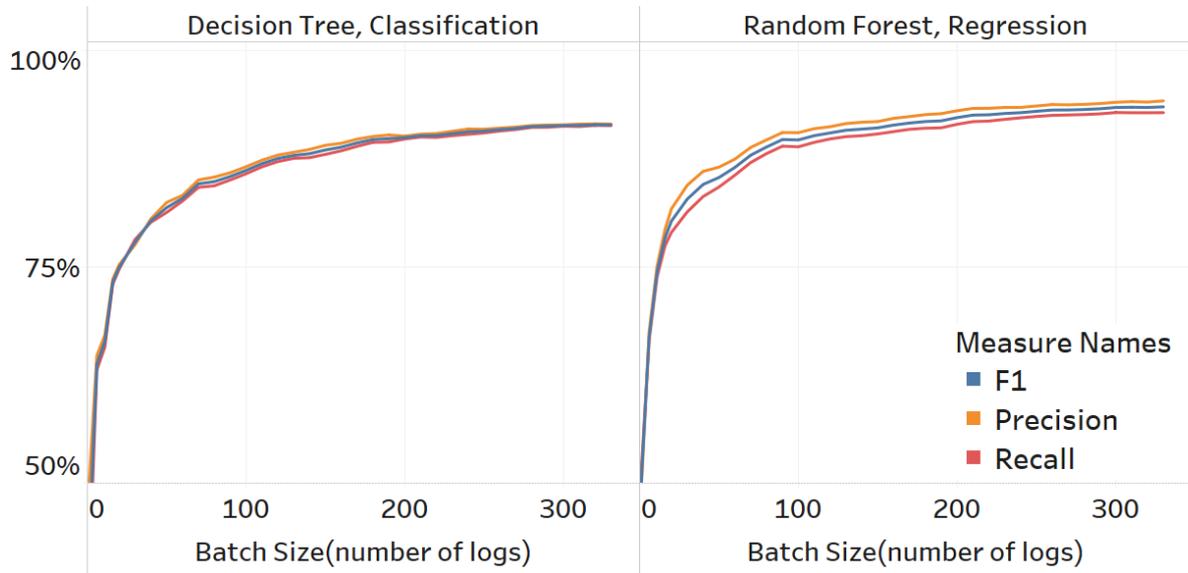


Figure 6. F1-score, precision, and recall (in %) as a function of the size of the batch (number of logs) for decision tree in classification and random forest in regression

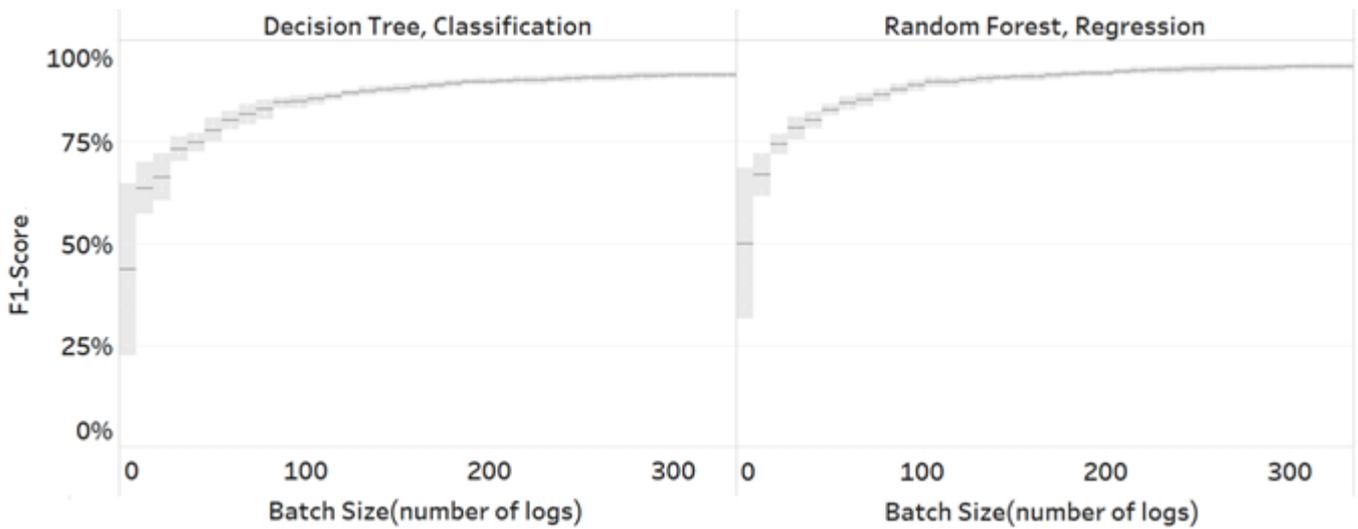


Figure 7. 95% confidence intervals for F1-score as a function of the batch size (number of logs)

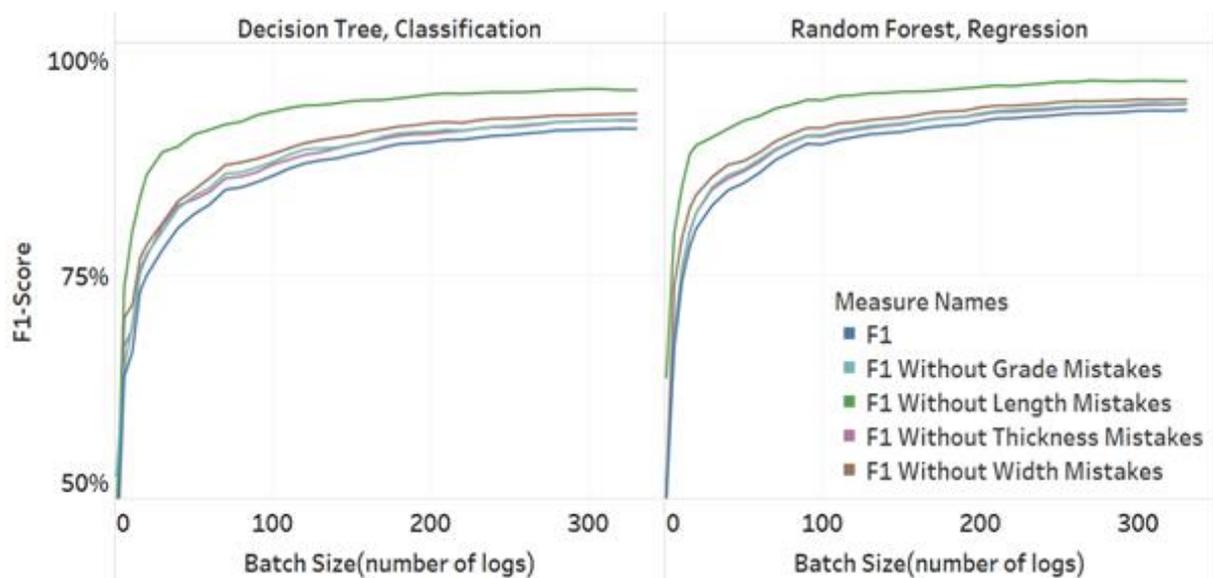


Figure 8. F1-score with and without single attribute mistakes for all types of error as a function of the size of the batch (number of logs)

3.6 Sources of error

Each given type of lumber product has a *length*, a *width*, a *thickness*, and a *grade* which, taken together, fully define the product. We call these characteristics the *lumber attributes*. By artificially ignoring errors on a specific lumber attribute, we artificially increase the score of the original model. Let us suppose we choose ignoring the length errors. Then, the magnitude of the score increment reflects the number of times the original model predicted the wrong length for a lumber without making a mistake on the other attributes.

The idea behind this analysis is to visualize “where” our models produce errors, but also to understand how to improve them in the long run. If, for instance, a model that predicts products that have the right width, length, and thickness is frequently misled on the grade, then it may be possible that the parametric data does not provide enough information to properly predict the grade. In this case it may be necessary to add information in the parametric description of the logs to improve the model.

Figure 8 reports the F1-score when we do not consider a specific type of mistakes. It shows that the length is the main source of error in our predictions. This is especially true in the case we are using a decision tree classifier as the curve without length mistakes is clearly above the others.

The curve representing the errors along the length is the curve that detaches most from the other curves. This suggests a higher error rate on length compared to other lumber attributes. Of course, other attributes also appear as sources of errors when inspecting the other curves, but the magnitude of the gain when ignoring them is less.

4 CONCLUSION

We showed how, in the context of supervised learning for sawing simulation approximation, the quality of a model can not only be quantified by its performance on individual logs, but also by its performance on batches. These results confirm that prediction errors on one log can be compensated for by errors on other logs, but also shows why more sophisticated models are preferable than a simpler model based on the historical average, especially for small batches.

Furthermore, we presented metrics that can be easily interpreted in an industrial context in the forest-product industry. Since those metrics make it easier to understand what is happening with the models in terms of over and under prediction of lumber counts, it becomes possible to explore the sources of errors in the models and interpret them in the context of the application. Those analyses pave the way to further model improvements.

5 ACKNOWLEDGEMENT

The authors would like to thank the FORAC Research Consortium and its partners. Our gratitude goes as well to the Natural Sciences and Engineering Research Council of Canada (NSERC) who provided funding for this research.

6 REFERENCES

- Andersson, K., & Luttrupp, C. (1997). Design for disassembly: Computer aid for separating surfaces and sorting borders. *Schriftenreihe WDK*, 351-354.
- Borchani, H., Varando, G., Bielza, C., & Larranaga, P. (2015). A survey on multi-output regression. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(5), 216-233.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1996). *Classification and Regression Trees*. Belmont, CA: Chapman and Hall.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45 (1): 5–32.
- Cai, L., Yu, Y., Zhang S., Song, Y., Xiong, Z. and Zhou, T., "A Sample-Rebalanced Outlier-Rejected k-Nearest Neighbor Regression Model for Short-Term Traffic Flow Forecasting," in *IEEE Access*, vol. 8, pp. 22686-22696, 2020.
- Fix, E., & Hodges Jr., J. L. (1951). *Discriminatory Analysis – Nonparametric Discrimination: Consistency Properties*. Tech. Rep., DTIC Document.
- FPIinnovations (2014). *Optitek 10. User's Manual*.
- Lan, T., Hu, H., Jiang, C., Yang, G., & Zhao, Z. (2020). A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. *Advances in Space Research*, 65(8), 2052-2061.

- Lin, W., Wang, J., & Thomas, R. E. (2011). A three-dimensional optimal sawing system for small sawmills in central Appalachia. In: Proceedings, 17th central hardwood forest conference; 2010 April 5-7; Lexington, KY; Gen. Tech. Rep. NRS-P-78. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station: 67-76. (Vol. 78, pp. 67-76).
- Morin, M., Paradis, F., Rolland, A., Wery, J., Gaudreault, J., & Laviolette, F. (2015). Machine learning-based metamodels for sawing simulation. In 2015 Winter Simulation Conference (WSC) (pp. 2160-2171). IEEE.
- Morin, M., Gaudreault, J., Brotherton, E., Paradis, F., Rolland, A., Wery, J., & Laviolette, F. (2020). Machine learning-based models of sawmills for better wood allocation planning. *International Journal of Production Economics*. 222. 107508.
- Powers, David M. W. (2020). "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation". *Journal of Machine Learning Technologies*. 2 (1): 37–63.
- Selma, C., Bril El Haouzi, H., Thomas P., Gaudreault J., & Morin M. (2018) An Iterative Closest Point Method for Measuring the Level of Similarity of 3D Log Scans in Wood Industry. In: Service Orientation in Holonic and Multi-Agent Manufacturing. *Studies in Computational Intelligence*, vol 762. Springer, Cham.
- Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, 20(1), 3-29.
- Stock, G. N., Greis, N. P., & Kasarda, J. D. (1998). Logistics, strategy and structure. *International Journal of Operations & Production Management*.
- Thomas, R.E. (2013). RAYSAW: A log sawing simulator for 3D laser-scanned hardwood logs. In: Proceedings, 18th Central Hardwood Forest Conference; 2012 March 26-28; Morgantown, WV; Gen. Tech. Rep. NRS-P-117. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station: 325-334. (Vol. 117, pp. 325-334).
- Todoroki, C. (1990). AUTOSAW system for sawing simulation. *New Zealand Journal of Forestry Science*.
- Van Rijsbergen, C. J. (1979). *Information Retrieval* (2nd ed.). Butterworth-Heinemann.

Liste des auteurs

- Addouche Sid-Ali, 530–536
Adri Ahmed, 296–303
Agard Bruno, 616–642
Akbari Haghghi Sina, 145–151
Al Masry Zena, 682–688
Alaoui Mdaghri Malika, 280–286
Alpan Gulgun, 178–185
Amegouz Driss, 280–286
Aouf Imen, 594–600
Arama Adama, 102–108
Arias Corzo Sergio Orlando, 355–362
Asgari Erfan, 498–505
Audy Jean-François, 482–488
Azzamouri Bassma, 154–161
- Bahria Nadia, 379–387, 394–401, 690–696
Bamoumen Mouna, 75–83
Barkallah Maher, 523–536
Baron Claude, 347–354
Barrios Piers, 746–753
Barth Marc, 186–193, 697–704
Barthe Delanoë Anne-Marie, 660–668
Battaïa Olga, 489
Batteux Michel, 363–370, 372–378
Belaud Jean-Pierre, 514
Belmabrouk Amal, 474–480
Bemmami Kamel-Eddin, 577–584
Ben Fradj Anes, 440–449
Ben Taleb Romain, 450–457
Ben-Ammar Oussama, 67–74
Benabbou Loubna, 705–722
Benali Maha, 228–235, 651–658
Bennouna Fatima, 280–286
Benyoucef Lyes, 152
Bernier Arnaud, 609–614
Berrado Abdelaziz, 705–714
Berrah Lamia, 466–473
Blanco Eric, 594–600
Bortolaso Christophe, 93–101
Bouchriha Hanen, 379–387, 394–401, 690–696
Boudaoud Nassim, 264–269
Boudina Ahlam, 651–658
Boukherroub Tassedat, 236–243, 440–449
Brotherton Edith, 170–177
Bélangier Olivier, 170–177
- Cabarbaye Adrien, 424–430
Cabarbaye Aurélien, 424–430
Cadoret Thomas, 730–737
Cardenas-Cristancho Diana, 304–311
Castillo Pedro, 424–430
- Chapel Claude, 466–473
Chasseray Yohann, 660–668
Chatras Clément, 134–144
Chehbi Gamoura Samia, 186–193
Chelbi Anis, 379–387, 394–401, 690–696
Cherfi-Boulangier Zohra, 264–269
Chevalier Marcel, 325–330
Chimentin Xavier, 258–263
Cholez Céline, 561–568
Clivillé Vincent, 466–473
Cloutier Caroline, 195–202
Courtin Christophe, 577–584
Cung Van-Dat, 219–227
- Da Cunha Catherine, 724–729
Dadouchi Camelia, 538–545, 634–642
Dahan Mathieu, 450–457
Dahmani Abdelhak, 152
Daille-Lefevre Bruno, 355–362
Dakkak Badr, 109–118
Damand David, 186–193, 697–704
Damien Trentesaux, 42–57
Dammak Houcine, 431–438
Danjou Christophe, 616–633, 643–658, 730–737
David Pierre, 594–600
De Jesus Garcia Edgar, 393
Derras Mustapha, 93–101
Di Mascolo Maria, 145–151
Dkhil Abdellatif, 431–438
Dolgui Alexandre, 67–74, 738–745
Dumitrescu Emil, 402–407
Dupuis Ambre, 634–642
- El Hassan El Hassan Irhirane, 109–118
Elfirdoussi Selwa, 84–91
Elomri Ameni, 754–761
Esteban Philippe, 347–354
Evrard Damien, 507–513
Eymery Angeline, 287–295
Eynard Benoît, 340–346, 746–753
- Ferguene Mounia, 538–545
Fernandez-Arias Alina, 490–497
Fontanili Franck, 93–101
Forget Pascal, 482–488

Fortin Arnaud, 675–681
Fouladirad Mitra, 319–324
Frein Yannick, 178–185, 498–505

Galmiche Théo, 553–560
Gamoura Samia, 697–704
Gardet Elodie, 577–584
Gardoni Mickael, 431–438
Gaudreault Jonathan, 762–769
Gazran Saba, 236–243
Gentil Marie-Hélène, 458–465
Ghrab Yahya, 134–144
Giampiccolo Sylvian, 682–688
Giard Vincent, 75–83
Gidel Thierry, 409–417
Godot Xavier, 355–362
Gondran Natacha, 490–497
Gosselin Louis, 515–522
Gouiaa Amal, 245–250
Grabot Bernard, 675–681
Grenier Lorenzo, 347–354
Grest Manon, 211–218
Guilhen Lucas, 643–650
Gzara Lilia, 561–568
Gülgün Alpan, 2–14

Habchi Georges, 577–584
Hachem Hassan, 319–324
Haddar Mohamed, 523–536
Haddou Benderbal Hichem, 738–745
Hafdhallaoui Emna, 251–257
Hammami Ramzi, 489–505
Hammami Roua, 245–250
Harbaoui Dridi Imen, 379–387, 394–401
Harbaoui Imen, 690–696
Havy Fabienne, 609–614
Hernández Thomas, 515–522
Hof Lucas A., 393
Hotait Hassane, 258–263
Houssin Rémy, 431–438
Hovelaque Vincent, 75–83, 154–161

Jaoua Amel, 251–257
Jarroudi Imen, 203–210
Jemai Zied, 162–169
Jouini Oualid, 162–169

Kandil Narjes, 489
Kenne Jean-Pierre, 393
Kerneis Romain, 363–370
Khebbache Selma, 372–378
Koala Denis, 178–185
Kokolo Joseph, 340–346

Koruca Ibrahim Halil, 186–193

Laaroussi Mouhcine, 715–722
Lachance Emilie, 195–202
Lamghari Amina, 482–488
Lamine Elyes, 93–101
Lamothe Jacques, 126–133
Lamouri Samir, 669–681
Lamé Guillaume, 162–169
Lauras Matthieu, 126–133, 211–218, 450–457
Le Duigou Julien, 730–737
Le Lann Jean-Marc, 660–668
Le Minh, 569–576
Lebel Luc, 170–177, 569–576
Legros Benjamin, 162–169
Lehoux Nadia, 170–177, 195–202, 515–522, 538–545, 569–576
Lehyani Fatma, 754–761
Leirtiz Jean-Michel, 271–279
Lhoste Pascal, 304–311

Mahraz Mohamed Iliasse, 705–714
Maire Jean Luc, 313–318, 546–552
Mairot Nicolas, 682–688
Maranzana Nicolas, 287–295
Maranzana Sébastien, 271–279
Marcandella Elise, 553–560
Marie Sylvain, 325–330
Martin Patrick, 355–362
Martineau Vincent, 762–769
Megoze Pomgha Patrick, 393
Mercantini Jean-Marc, 152
Merle Catherine, 458–465
Merlo Christophe, 102–108
Metaireau Agathe, 219–227
Miclo Romain, 126–133, 450–457
Millet Dominique, 418–423
Minot-Lenfant Alicia, 287–295
Mlaouhi Kaouthar, 561–568
Montarnal Aurélie, 450–457
Montecinos Julio Cesar, 715–722
Monticolo Davy, 304–311
Montreuil Benoit, 211–218
Morin Michael, 762–769
Mouatassim Marouane, 431–438
Muller Laurent, 304–311
Muller Thomas, 313–318

Nedjar Merzak, 388–392
Negny Stephane, 660–668
Nguyen Angie, 669–674
Niel Eric, 402–407
Nouira Imen, 490–497

Nozari Milad, 482–488
 Noël Frédéric, 601–608

 Oger Raphaël, 126–133
 Olivier Claude, 440–449
 Omri Nabil, 682–688
 Ouali Mohamed Salah, 251–257
 Ouhimmou Mustapha, 715–722
 Ouled Bedhief Asma, 59–66
 Oullada Oumaima, 296–303

 Pahun Laurent, 347–354
 Paquet Marc, 203–210, 236–243
 Parrenin Loic, 616–624
 Peigné Sarah, 507–513
 Pellerin Robert, 643–658, 669–674
 Philippe VIALLETELLE, 15–31
 Pialot Olivier, 418–423
 Piat Jean-Rémi, 625–633
 Pillet Davy, 313–318
 Pillet Maurice, 313–318, 546–552
 Pingaud Hervé, 93–101
 Pinquie Romain, 601–608
 Ponchet-Durupt Amelie, 264–269
 Pouydebat Olivier, 577–584
 Prioux Nancy, 514

 Rasolofondraibe Lanto, 258–263
 Raux Mathieu, 162–169
 Rifai Said, 296–303
 Rigaud Eric, 119–124
 Robert Pellerin, 32–41
 Romero Victor, 601–608
 Rose Bertrand, 271–279
 Rönnqvist Mikael, 236–243

 S A Puviyarasu, 724–729
 Sablayrolles Caroline, 514
 Saidi Hassani Alaoui Hajar, 109–118
 Sali Mustapha, 134–144
 Salvado Laura Laguna, 102–108
 Santamaria-Acevedo Gustavo, 162–169
 Santana Jussara, 746–753
 Shahin Kamrul Islam, 331–338
 Sidki Mouad, 84–91
 Siebenborn Thierry, 546–552
 Simard Vanessa, 195–202
 Simeu-Abazi Zineb, 145–151
 Simon Christophe, 331–338
 Sirgabsou Yandika, 347–354
 Slama Ilhem, 67–74
 Smati Abdelnacer, 388–392
 Soltane Amel, 530–536

 Tariq Sara, 228–235
 Tchernev Nikolay, 84–91
 Tollenaere Michel, 474–480, 754–761
 Trabelsi Imen, 523–529
 Trentesaux Damien, 466–473
 Trépanier Martin, 228–235
 Turki Sadok, 67–74

 Usuga Cadavid Juan Pablo, 675–681

 Vallerand Steve, 762–769
 Vialle Claire, 514
 Vidal Jean-Baptiste, 126–133
 Villeneuve Éric, 102–108
 Vu Hai Canh, 264–269, 319–324

 Weber Philippe, 331–338

 Xu Changyi, 402–407

 Yahouni Zakaria, 178–185
 Yelles-Chaouche Abdelkrim A., 738–745
 Younsi Karim, 388–392
 Yvars Pierre-Alain, 586–593

 Zeddini Bisma, 523–529
 Zerhouni Noureddine, 682–688
 Zhang Liwen, 93–101
 Zimmer Laurent, 586–593
 Zolghadri Marc, 523–536
 Zonghero William, 409–417
 Zouari Alaeddine, 474–480, 754–761

