



HAL
open science

Intégration d'outils de simulation et d'optimisation pour le pilotage d'une ligne d'assemblage multiproduit à transfert asynchrone

Franck Fontanili

► **To cite this version:**

Franck Fontanili. Intégration d'outils de simulation et d'optimisation pour le pilotage d'une ligne d'assemblage multiproduit à transfert asynchrone. Automatique / Robotique. Université Paris 13, 1999. Français. NNT: . tel-03742039

HAL Id: tel-03742039

<https://hal.science/tel-03742039>

Submitted on 2 Aug 2022

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.

Université PARIS 13
Groupe de Recherche en Production Intégrée (G.R.P.I)

Thèse

Présentée par

Franck FONTANILI

Pour obtenir le grade de

Docteur de l'Université Paris 13

Spécialité : **Génie Mécanique**

Sujet :

**Intégration d'outils de simulation et d'optimisation
pour le pilotage d'une ligne d'assemblage
multiproduit à transfert asynchrone**

Soutenue le 26 octobre 1999 devant les membres du Jury :

M. Jean Paul KIEFFER, Professeur, Directeur de l'Ecole Nationale Supérieure des Arts et Métiers d'Aix-en-Provence, rapporteur

M. Pierre DEJAX, Professeur, Directeur du Laboratoire d'Automatique et Productique de l'Ecole des Mines de Nantes, rapporteur

M. Raymond PONSONNET, Professeur, Directeur du Groupe de Recherche en Production Intégrée de l'IUT de St-Denis, Directeur de Thèse

M. Gérard PLATEAU, Professeur, L.I.P.N. de Paris 13

M. Patrick TRUCHOT, Professeur, ENSGSI Nancy

M. Hugues MOLET, Professeur, Ecole des Mines de Paris

M. Thierry SORIANO, Maître de conférence, ISMCM CESTI Toulon

M. Bernard LE MEAUX, Chef du dépt. Industrialisation moteurs E/K, RENAULT

M. Michel MOINET, Consultant

Problématique de la thèse et présentation générale	3
<i>Partie 1</i>	
Introduction sur la production industrielle	13
<i>Partie 2</i>	
Caractéristiques des processus d'assemblage	47
<i>Partie 3</i>	
Présentation du système expérimental	71
<i>Partie 4</i>	
Modélisation et simulation de flux	87
<i>Partie 5</i>	
Elaboration du plan expérimental sur le modèle de simulation	139
<i>Partie 6</i>	
Analyse des résultats des simulations	171
<i>Partie 7</i>	
Optimisation de la production	225
<i>Partie 8</i>	
Choix et intégration d'un outil d'optimisation	265
Bilan et perspectives de la thèse	335
Table des matières	343

***Problématique de la thèse
et
présentation générale***

I- Problématique de la thèse

Mon travail d'enseignant au département **O**rganisation et **G**estion de la **P**roduction de l'IUT de Cergy-Pontoise est en grande partie à l'origine de cette thèse. En effet, les nombreux contacts avec des industriels d'une part et la réalisation d'une ligne de production expérimentale d'autre part m'ont très rapidement permis de mettre en évidence des problèmes liés au pilotage optimal d'un système de production.

A la fin des années 80, j'ai d'abord découvert le courant novateur du "**Juste à Temps**" (JAT) et du "**Computer Integrated Manufacturing**" (CIM) qui faisaient tous deux figures de solution à la crise que rencontraient les industries occidentales. L'objectif principal de ces concepts était (et est toujours) de disposer d'un système de production capable de s'adapter très facilement et très rapidement à une demande de plus en plus variée et exigeante sur les délais et la qualité, tout en réduisant au maximum les coûts de production, et ceci dans un contexte de forte concurrence.

J'ai visité à cette époque des entreprises telles que *CITROEN*, qui disposait d'un atelier flexible d'usinage à Meudon, *IBM Orléans* qui avait une ligne d'assemblage flexible de rubans encreurs pour imprimantes à aiguilles, *THOMSON SDC* qui intégrait à Sartrouville un outil informatique de gestion des programmes pour les machines à commande numérique destiné à réduire le temps de changement de fabrication, ...

Depuis, ces concepts ont fortement évolué, et ces sites de production aussi, allant même jusqu'à disparaître pour certains d'entre eux

En moins de vingt ans, nous sommes progressivement passés d'une notion de flexibilité qui était synonyme de "*tout automatique*", à une notion de flexibilité où l'homme avait une place prépondérante au sein de l'outil de production. En effet, les premiers ateliers flexibles automatisés étaient très spectaculaires... quand ils fonctionnaient.

On s'est rapidement rendu compte de leur manque d'efficacité en terme de productivité, dû en grande partie à leur taux de fiabilité globale, produit du taux de fiabilité de chaque équipement utilisé. Plus le nombre d'équipements automatisés augmente, plus la fiabilité globale diminue. Même avec des équipements qui possèdent un taux de fiabilité élevé, par exemple 0.95, il est facile de démontrer que la fiabilité d'une ligne comportant dix de ces équipements n'est que de 0.59 (soit 0.95^{10}) !

Les solutions technologiques totalement automatisées n'étaient pas la panacée et cette prise de conscience tardive a permis aux industriels de se pencher sur l'aspect organisationnel de la production. Comme le souligne l'un des précurseurs de cette approche en France en analysant les méthodes japonaises, "*Automatiser des problèmes, ce n'est qu'amplifier ces problèmes. La première étape est donc de commencer par simplifier, organiser et éliminer les problèmes*" (P. Béranger dans "Les nouvelles règles de la production", Dunod 1987).

Depuis quelques années, on assiste donc à un retour en force de l'homme, en tant qu'opérateur, sur des lignes de production qui, quelques années plus tôt, étaient

entièrement automatisées. L'analyse que je fais de ce constat est que la tendance semble être d'automatiser des tâches faciles mais très répétitives ou pénibles et de confier à l'homme des tâches variées et délicates où son intelligence et sa dextérité seront bien supérieures à n'importe quel robot ou machine automatique. Par ailleurs, ce qui se prête bien à certains processus particuliers, comme l'usinage par exemple, n'est pas forcément très représentatif de tous les autres processus. En effet, les centres d'usinage ont la particularité de pouvoir enchaîner des opérations sur une pièce unique avec un minimum de manipulations. Une pièce usinée peut même être réalisée entièrement sans démontage. Les machines outils à commande numérique ou les machines transfert permettent donc de pousser très loin le niveau d'automatisation.

Si l'on regarde les objets qui nous entourent, combien sont issus d'un procédé d'usinage, combien sont composés d'une seule pièce ? Très peu.

Si l'on analyse un peu plus ces objets, on en déduit finalement que la plupart d'entre eux sont issus d'un processus d'assemblage : un stylo, un cahier, un ordinateur, une lampe de bureau, une fenêtre, une voiture, un pantalon, etc...

C'est dans ce type de processus d'assemblage que les difficultés sont les plus grandes, indépendamment du procédé d'assemblage choisi (collage, rivetage, sertissage, frettage, encliquetage, vissage, etc...). Tout d'abord, pour assembler un produit, il faut disposer de tous ses composants, au bon endroit et au bon moment, d'où des problèmes de synchronisation et d'approvisionnement. Ensuite, pour des raisons de productivité ou de technicité, les différentes opérations d'assemblage se font souvent sur plusieurs postes de travail, ce qui entraîne des problèmes d'équilibrage de charge entre les postes, de transfert et d'encours de composants et sous-ensembles.

Les processus d'assemblage sont donc à l'origine et au centre de ma réflexion car ils me semblent être particulièrement riches en pistes de recherche.

Les premiers articles que j'ai trouvés sur ce thème m'ont révélé que de nombreux chercheurs s'intéressaient en particulier à l'optimisation combinatoire des gammes d'assemblage en faisant souvent abstraction du système de production lui-même. ***Il m'a semblé qu'une piste était à explorer concernant les lignes d'assemblage.***

Parallèlement, j'ai été amené en 1994 à concevoir et à réaliser une plate-forme pédagogique pour des étudiants du DUT **O**rganisation et **G**énie de la **P**roduction, destinée à mettre en pratique les enseignements théoriques dispensés dans ce cursus. La principale contrainte du cahier des charges était de reproduire en milieu universitaire les problèmes d'organisation, de gestion, de pilotage et d'optimisation d'un vrai système de production. Pour orienter mes recherches, je suis parti du constat que le gestionnaire de production gère plus des flux que des procédés. Les procédés utilisés à chaque étape de fabrication de composants, de sous-ensembles ou d'un produit se limitent pour lui à un temps opératoire. Ce temps opératoire correspond exclusivement à la transformation, autrement dit à l'apport de valeur. Que cette valeur corresponde à une transformation physique de la matière par usinage, emboutissage, moulage, formage, assemblage ou autre, ne présente que peu d'intérêt pour le gestionnaire de production. Ce qui lui importe essentiellement est donc la durée de chaque opération et l'ordre des opérations. Il n'a que très peu de maîtrise

sur les opérations elles-mêmes : son rôle consiste essentiellement à gérer l'ensemble du processus d'un point de vue "flux", c'est-à-dire à regarder les composants et sous-ensembles comme des entités qui circulent ou attendent entre différentes opérations où ils sont transformés, divisés ou assemblés pour obtenir un produit à livrer au client dans un délai imparti. Le seul aspect des procédés qui peut le concerner est la qualité. Il faut que le procédé choisi permette d'obtenir un résultat correspondant aux exigences du client en terme de qualité. Toutes ces réflexions m'ont donc progressivement amené à m'intéresser non seulement aux processus d'assemblage, à cause de leur représentativité, mais aussi aux systèmes de production permettant de réaliser ces processus d'assemblage.

Comme je l'ai évoqué précédemment, le marché a évolué depuis une vingtaine d'années et s'oriente de plus en plus vers des produits à faible durée de vie et à fortes variantes. Il est donc nécessaire pour les industriels de disposer de systèmes de production, d'assemblage en particulier, permettant de répondre à cette demande par une grande flexibilité, ce que les Anglo-saxons appellent le "*High-Mix Low-Volume*" (voir l'ouvrage portant ce nom de R. Michael Mahoney, Prentice Hall 1997)

Une enquête rapide auprès d'industriels et différentes lectures m'ont permis de faire ressortir un type particulier de système de production destiné à l'assemblage en petites à moyennes séries, répétitives ou pas, de produits à variantes ou options. Ce type de système est composé d'un dispositif de transfert asynchrone comportant des convoyeurs automatisés qui entraînent des palettes entre les différents postes de travail manuels ou automatisés. Leur flexibilité est obtenue à la fois par leur structure et leur modularité : les postes de travail peuvent être implantés suivant différentes topologies (en U, en L, en boucle ouverte ou fermée, en série, en dérivation, etc...) et peuvent être facilement reconfigurés soit directement sur la ligne, soit en dehors de la ligne, puis "connectés" suivant le principe de la brique Lego. Ces systèmes de production permettent donc de réduire ou même d'éliminer l'un des principaux soucis du gestionnaire de production qui est de déterminer la taille des lots de transfert entre 2 postes de travail en fonction de critères économiques prenant en compte la durée du transfert par rapport à la durée des opérations.

De plus, cette solution de transfert automatisé est une réponse aux problèmes fréquents que l'on rencontre dans des installations où le transfert entre les opérations est réalisé manuellement, avec ou sans l'assistance de moyens mécanisés (chariots, transpalettes, etc...). En effet, l'implantation des différents postes de travail, même si elle est optimale, ne permet pas toujours un écoulement continu des articles traités. On est donc amené à regrouper les articles par lots de transfert, dans des conteneurs par exemple, ce qui occasionne donc des temps d'attente importants : dans un lot de transfert de 100 articles, chaque article passe 99% de son temps de défilement en attente/stockage, et ceci en supposant le temps de transfert négligeable ! Il est donc facile de démontrer qu'un écoulement pièce à pièce entre les postes de travail entraîne une fluidité optimale et une réduction importante des délais, sans changer ni les étapes du processus, ni la nature ou la durée des procédés utilisés à chaque étape.

Ce type de ligne illustre aussi la tendance évoquée plus haut qui est d'automatiser des actions simples et répétitives, le transfert faisant partie de ces actions. Il est toutefois utile

d'émettre quelques réserves sur une généralisation abusive de l'automatisation du transfert. En effet, le transfert n'apportant aucune valeur ajoutée au produit, le plus efficace serait donc de l'éliminer, purement et simplement, ou de le réduire au minimum comme cela se fait dans les processus continus. L'automatisation du transfert engendre donc des coûts qui n'apporte pas directement de valeur aux produits fabriqués. C'est donc une solution à retenir dans des cas bien spécifiques, comme celui qui nous intéresse dans ce travail, après avoir notamment vérifié que l'implantation des postes de travail ne peut être améliorée. Les lignes flexibles qui traitent plusieurs produits avec des gammes de fabrication différentes ont des postes de travail qui ne sont pas forcément implantés de façon optimale pour chaque produit. L'automatisation du transfert présente deux intérêts : non seulement minimiser les temps de défilement de chaque article, mais aussi affranchir les opérateurs et le gestionnaire de la difficulté à traiter et organiser une multitude de routages entre les postes de travail. Le gestionnaire de production se concentre donc exclusivement sur le choix des paramètres de pilotage nécessaires à chaque lancement. Les opérateurs n'ont à s'occuper que des opérations sans se soucier du transfert. L'automatisation du transfert me semble donc être justifiée dans de tels cas.

En m'inspirant d'exemples réels pris dans des entreprises du secteur informatique telles que *IBM Orléans* ou *APPLE Computer Ltd. of Ireland*, du secteur automobile (*VALEO, Compagnie Européenne de Sièges pour l'Automobile*), ou de secteurs logistiques (*SODIS* pour la distribution de livres, *SERVAIR* pour la distribution de plateaux repas dans les avions), j'ai donc réalisé une ligne flexible expérimentale basée sur une structure de transfert automatisé par convoyeurs à bandes et palettes entre des postes de travail implantés en dérivation. L'une des particularités de ce système est qu'il n'a pas été conçu pour fabriquer un produit particulier, ce qui peut sembler en totale opposition à la démarche habituelle pour ce type de système qui est de concevoir un produit puis de concevoir le système qui va fabriquer ce produit. Sa conception est le fruit de l'observation de multiples systèmes de production, automatisés ou pas, et du constat que parmi tous ces systèmes, les plus flexibles étaient justement ceux qui n'étaient pas dédiés à un produit. On peut par exemple citer les ateliers de mécanique générale de type *Job Shop* où l'on trouve des tours, des fraiseuses, des perceuses, des postes d'assemblage, etc.. et où l'on détermine la gamme de fabrication des produits en fonction des équipements existants. On va parfois même jusqu'à concevoir ou reconcevoir le produit en fonction du système de production.

C'est donc ce principe qui a été retenu, privilégiant le processus au produit. Pour fabriquer une grande variété de produits, la conception a donc consisté essentiellement à rechercher une architecture de ligne permettant d'obtenir la flexibilité nécessaire. La principale contrainte du système qui peut empêcher la fabrication d'un produit est la capacité dimensionnelle des palettes. Abstraction faite des procédés de fabrication, ce système est donc capable de traiter l'enchaînement des opérations de n'importe quel produit, à condition que celui-ci ait des dimensions inférieures ou égales à celles de la palette.

Ce système expérimental offre donc une grande flexibilité quant aux produits fabriqués. En contrepartie, la flexibilité oblige à un surdimensionnement ou une surcapacité de certains éléments, par exemple le nombre de palettes en réserve ou la longueur de certains tronçons de convoyeurs. La conception de ce système m'a obligé à me poser un certain

nombre de questions, mais je dois avouer que c'est surtout son exploitation qui a déclenché l'identification de nombreux problèmes.

La flexibilité et la surcapacité évoquées plus haut, qui ne me semblaient pas être des contraintes à priori, bien au contraire, s'avèrent en réalité pénalisantes dans certains cas de production et suivant certains critères. A titre d'exemple, la flexibilité oblige le gestionnaire de production à déterminer à chaque nouveau lancement la valeur de certains paramètres de pilotage comme le Temps Inter Lancement (TiLa) entre 2 palettes consécutives ou encore la capacité maximale d'accueil en palettes des tronçons en amont des postes de travail (StAm), pour n'en citer que deux.

On peut dire que la spécificité d'un système flexible, quel qu'il soit, est qu'il comporte de nombreux paramètres variables qu'il va falloir déterminer à chaque nouveau lancement pour obtenir le rendement optimum, alors qu'un système dédié est supposé être réglé une fois pour toutes.

C'est principalement cet aspect qui sert de fil conducteur à ce travail de recherche. Il s'agit donc de proposer une méthodologie et des outils permettant aussi bien au concepteur qu'à l'exploitant d'un système flexible d'assemblage de déterminer les valeurs optimales de certains paramètres du système. Chaque fois que cela sera possible, on mettra en évidence la possibilité de transférer cette méthodologie ou ces outils à des systèmes de production plus généraux.

II- Résumé de la thèse

Cette thèse est structurée en huit parties qui permettent un cadrage progressif de la problématique.

Les parties 1, 2, 3, 4 et 7 présentent un état de l'art des thèmes abordés à partir duquel une contribution a été apportée, soit en tant que synthèse, soit comme complément.

Les parties 5, 6 et 8 sont développées suivant une démarche expérimentale visant à élaborer une méthodologie et des outils d'optimisation.

Dans la **partie 1 : "Introduction sur la production industrielle"**, le contexte général de la recherche est présenté, à la fois pour l'aspect organisationnel de la production et aussi pour l'aspect technique de l'implantation des lignes. Un certain nombre de définitions et de typologies sont définies afin de lever toute ambiguïté sur des propos qui seront tenus ensuite. En effet, dans le domaine de la gestion de production, malgré quelques tentatives de normalisation, on peut s'accorder pour dire que l'une des principales difficultés est d'uniformiser le langage. Ce n'est pas une nouvelle tentative de terminologie qui est

proposée, mais uniquement un préliminaire permettant de préciser le contexte. Le dernier chapitre fait apparaître les caractéristiques principales des lignes de production en ce qui concerne leur implantation et le mode de transfert entre les opérations.

Un recentrage du contexte sur les processus d'assemblage est présenté dans la **partie 2 : "Caractéristiques des processus d'assemblage"**. Après une présentation de la typologie des principaux processus d'assemblage, une analyse fonctionnelle d'un système de production à transfert automatisé destiné à l'assemblage est proposée suivant le formalisme SADT. Cette proposition est le résultat d'une synthèse bibliographique et d'observations diverses. Quelques solutions technologiques sont présentées pour illustrer chaque fonction, à la fois pour fixer les idées, mais aussi pour permettre une meilleure compréhension de la description du système expérimental qui est faite dans la partie 3.

La **partie 3 : "Présentation du système expérimental"** décrit la ligne de production qui sert de support expérimental à cette recherche. Une description détaillée du sous-système physique et du sous-système informationnel est donnée afin de pouvoir analyser le plus précisément possible le comportement dynamique de cette ligne.

Ce travail fait en grande partie appel à une démarche expérimentale visant à observer des phénomènes, à identifier des problèmes, à proposer des solutions, à mesurer leurs effets, et enfin à tenter d'expliquer certains résultats. L'intérêt de cette démarche expérimentale est qu'elle offre un champ d'investigations très large, puisque chaque expérimentation peut faire apparaître de nouveaux phénomènes que l'on tentera de comprendre et d'élucider.

Bien que pour cette recherche, on se soit servi d'un système physique destiné à des expérimentations, il est difficile d'imaginer que la démarche suivie puisse être reproductible en environnement industriel sans poser bon nombre de problèmes. Tout d'abord, le système risque d'être immobilisé pour la réalisation des expérimentations, ce qui peut être préjudiciable au niveau économique. Ensuite les expérimentations peuvent entraîner des dysfonctionnements. Enfin, on peut difficilement envisager de faire des expérimentations nécessitant une modification de la structure physique ou informationnelle du système. Pour toutes ces raisons, la méthodologie proposée fait appel d'une part à des expérimentations sur le système réel afin de disposer d'un certain nombre de données, et d'autre part à des expérimentations sur un modèle de simulation de flux à événements discrets, validé par les informations relevées sur le système réel. La collecte des données sur le système réel met aussi en évidence le problème du temps de réponse : quand une expérimentation est lancée, le délai peut atteindre plusieurs dizaines de minutes ou plus pour obtenir certaines données. Un modèle de simulation validé fournit les mêmes données que le système réel avec des valeurs statistiquement acceptables et en un temps de réponse environ 200 fois plus faible que le système réel (sur un micro-ordinateur de type PENTIUM 200 Mhz). Contrairement à un modèle purement mathématique, le modèle de simulation à événements discrets permet de reproduire très précisément le comportement du système réel, sans nécessairement émettre des hypothèses simplificatrices qui peuvent entraîner

des effets pervers sur les résultats et ainsi fausser complètement la nature du problème. L'une des principales difficultés dans l'utilisation de la simulation de flux est l'étape de modélisation permettant de passer du système réel au modèle virtuel du système à simuler. Dans la **partie 4 de cette thèse "Modélisation et Simulation de flux de production"**, après une présentation de la simulation de flux permettant de mettre en évidence l'intérêt de cet outil, on propose donc le développement d'une étape intermédiaire entre le système réel et le modèle virtuel. Le principal objectif de cette étape intermédiaire est de formaliser le système réel suivant une représentation graphique faisant apparaître à la fois les flux et les données, non seulement pour faciliter la réalisation du modèle, mais aussi la communication entre les différents acteurs concernés par une étude de simulation.

La plupart des expérimentations seront donc réalisées à partir du modèle de simulation qui peut fournir des résultats en un temps de réponse bien plus faible que le système réel. La démarche expérimentale nécessite de définir les conditions des expérimentations. Il faut choisir les paramètres à faire varier et leur étendue de variation, en fonction de plusieurs scénarios ou d'un plan d'expérimentations. Cet aspect est présenté dans la **partie 5 : "Elaboration du plan expérimental sur le modèle de simulation"**.

L'analyse des résultats expérimentaux réalisés en simulation et une résolution analytique des problèmes identifiés est proposée dans la **partie 6 : "Analyse des résultats des simulations"**. Cette résolution analytique a deux objectifs : soit de vérifier expérimentalement des relations afin de les valider, soit de déterminer des relations ou des modèles pouvant ensuite être généralisées à d'autres scénarios ou situations. C'est le cas en particulier du principe du CONWIP ("CONstant Work In Process, voir bibliographie de la partie 6) qui est particulièrement bien illustré sur une ligne fermée avec convoyeurs et palettes. Ce principe est généralisable à d'autres types d'ateliers de production et l'on utilisera la ligne expérimentale pour vérifier et pour apporter une contribution à ce modèle.

Cette partie met aussi progressivement en évidence la nécessité de disposer d'un outil permettant d'optimiser automatiquement certains paramètres du système. En effet, pour ne pas multiplier le nombre d'expérimentations, on contraint les paramètres à ne prendre que certaines valeurs. Mais si l'on relâche ces contraintes pour autoriser les paramètres à prendre n'importe quelle valeur sur leur domaine de variation afin d'intégrer des phénomènes proches de la réalité, le problème devient complexe et donc difficile à optimiser par l'utilisation d'une ou de plusieurs règles logiques. Ces règles peuvent toutefois servir de point de départ à une optimisation plus précise. On introduit aussi dans cette partie la subjectivité de l'optimisation en fonction des critères retenus. On compare les résultats des expérimentations en fonction de différents critères d'optimisation.

Dans la **partie 7 " Optimisation de la production"**, on tente donc d'identifier les méthodes algorithmiques les plus appropriées permettant d'optimiser un système de production. Parmi toutes les méthodes utilisées en optimisation combinatoire, les algorithmes approchés ou heuristiques semblent être les plus adaptés à la résolution de ce type de

problème. En particulier, les algorithmes génétiques nous paraissent les plus efficaces pour résoudre des problèmes fortement combinatoires.

La **partie 8 " Choix et intégration d'un outil d'optimisation "** propose de coupler un algorithme de type heuristique à un modèle de simulation pour optimiser le système modélisé. Dans un premier temps, on réalise des expérimentations avec quatre algorithmes différents : algorithme d'exploration aléatoire, algorithme de descente stochastique, algorithme du recuit simulé et enfin algorithme génétique. Ce dernier algorithme est présenté en insistant sur la contribution apportée en ce qui concerne le codage des paramètres, le mécanisme de sélection de type élitiste, la prise en compte de la consanguinité entre les parents sélectionnés. Les résultats obtenus par cet algorithme sont comparés avec ceux obtenus, dans les mêmes conditions, avec les autres algorithmes testés. L'objectif n'est pas de comparer l'efficacité de ces algorithmes, car trop de paramètres intrinsèques peuvent influencer considérablement les résultats, mais plutôt de justifier l'utilisation d'un algorithme plutôt qu'un autre.

Plusieurs expérimentations sont réalisées, avec les mêmes hypothèses que celles de la partie 6, en vue d'utiliser le couplage simulation / algorithme génétique pour optimiser le délai de fabrication d'une campagne. Les résultats, comparés à ceux de la partie 6, montrent l'intérêt de ces deux outils.

Pour se rapprocher des problèmes industriels réels, la simulation et l'algorithme génétique sont ensuite utilisés pour faire de l'optimisation multicritère : on cherche à minimiser simultanément le délai de production, l'encours moyen et le nombre de réglages sur les postes de travail. Dans un tel cas, la principale difficulté réside dans la méthode de sélection des solutions qui satisfont tous les critères. On expérimente plusieurs méthodes et on compare les résultats obtenus.

A partir de la méthodologie suivie dans ce travail et des résultats obtenus, ce travail de recherche se termine par une proposition d'intégration des outils informatiques à disposition du gestionnaire de production. Chaque outil est d'abord décrit en fonction des données qu'il traite en entrée et des résultats qu'il fournit en sortie. Un modèle mettant en évidence les relations entre chaque outil est proposé afin d'analyser l'intégration de ces outils et d'en proposer une utilisation structurée.

Enfin, un **bilan et des perspectives de recherches** issues de cette thèse sont proposés en conclusion. Différentes pistes sont envisagées, allant de la conception de lignes de production à la validation d'algorithmes d'optimisation multicritère, en passant par l'utilisation de la simulation pour tester des modes de pilotage d'atelier tels que le ConWip.

- PARTIE 1 -
***Introduction sur la
production industrielle***

Sommaire

Introduction	15
I- Définitions, techniques, règles et principes généraux d'organisation et de gestion de la production industrielle.....	17
<input type="checkbox"/> I.1- Contexte général	17
<input type="checkbox"/> I.2- Définitions de base.....	17
<input type="checkbox"/> I.3- Définition de la production.....	20
<input type="checkbox"/> I.4- Définition et objectif de la gestion de production.....	20
II- Evolution de la production et nouvelles règles.....	22
<input type="checkbox"/> II.1- Les principales périodes	22
<input type="checkbox"/> II.2- Les nouvelles règles de la production moderne.....	23
III- Typologie de la production.....	24
<input type="checkbox"/> III.1- Structure du produit.....	24
<input type="checkbox"/> III.2- Modes de production.....	25
<input type="checkbox"/> III.3- Circulation des produits dans l'atelier	26
<input type="checkbox"/> III.4- Relation avec le client	28
IV- Synthèse des différents modes de pilotage des flux de production	32
<input type="checkbox"/> IV.1- Pilotage centralisé	32
<input type="checkbox"/> IV.2- Pilotage par l'aval et décentralisé.....	33
<input type="checkbox"/> IV.3- Pilotage par l'amont et décentralisé.....	34
<input type="checkbox"/> IV.4- Pilotage synchronisé entre flux physique et flux informationnel.....	34
<input type="checkbox"/> IV.5- Pilotage par les contraintes ou méthode OPT	34
V- Typologie des flux physiques et implantation des lignes de production.....	37
<input type="checkbox"/> V.1- Flux physique avec opérations et manutentions synchronisées (ou liées).....	37
<input type="checkbox"/> V.2- Flux physique avec opérations et manutentions asynchrones (ou libres).....	38
<input type="checkbox"/> V.2.a) Manutention non automatisée	39
<input type="checkbox"/> V.2.b) Manutention (ou transfert) automatisée	41
<input type="checkbox"/> V.3- Autre typologie.....	41
Conclusion	43
Bibliographie	44

Introduction

Cette partie a pour objectif de situer le contexte général de la thèse, c'est à dire le domaine de la gestion de production et des systèmes de production. S'agissant d'un domaine peu exploré dans le monde de la recherche parce qu'encore à ses balbutiements (en France, la gestion de la production n'est toujours pas considérée comme une science à part entière), il m'a semblé utile et même indispensable de commencer par donner quelques définitions. L'une des particularités de la gestion de production est qu'elle repose souvent sur des règles de logique et de bon sens plus que sur des théories mathématiques abstraites. Un minimum de rigueur est donc nécessaire pour la promouvoir au rang des sciences nobles, et ceci passe en partie par l'utilisation d'une terminologie sans ambiguïté.

C'est ce qui explique la place donnée au chapitre 1, et plus particulièrement aux définitions, qui dans un autre cadre, auraient pu être données en annexe.

Le chapitre 2 fait une présentation de l'évolution de la production depuis le début de l'ère industrielle. Il met en évidence les caractéristiques très contraignantes auxquelles la production moderne doit s'adapter. Le concept du "Juste à Temps" est évoqué ainsi que les nouvelles règles imposées par le marché.

Dans le chapitre 3, la production est classifiée suivant 4 critères :

- structure du produit,
- modes de production,
- circulation des produits dans l'atelier,
- relation avec le client.

Chacun de ces critères comporte plusieurs classes, et la combinaison de chacune de ces classes permet d'établir jusqu'à 64 configurations différentes ($4 \times 4 \times 2 \times 2 = 64$), bien que l'on se soit contenté de prendre uniquement les principaux critères et les classes les plus significatives. On entrevoit déjà la diversité qui existe au sein des systèmes de production.

Pour cette recherche, on se focalisera sur la production de produits à structure convergente fabriqués en série et sur commande.

A la fin de ce chapitre, un modèle de représentation des flux est proposé permettant de situer les principales étapes et données d'un processus de production de produits assemblés.

Le chapitre 4 présente les cinq principaux modes de pilotage de flux de production que l'on rencontre dans les ateliers de production : pilotage centralisé (poussé), par l'aval (tiré), par l'amont, synchronisé, par les contraintes.

Les flux physiques et les principales implantations de lignes de production sont classifiés dans le chapitre 5. On met surtout en relief les modes de transfert possibles entre les postes de travail.

I- Définitions, techniques, règles et principes généraux d'organisation et de gestion de la production industrielle

□ I.1- Contexte général

La production industrielle comporte trois éléments en interaction :

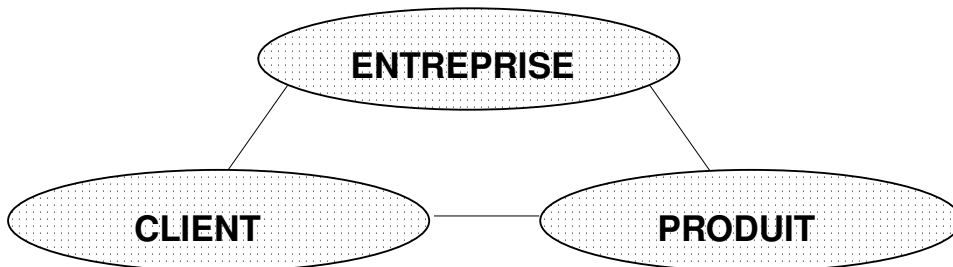


Figure 1 : Les 3 éléments de la production

Le **CLIENT** éprouve un **BESOIN** immédiat ou différé pour un **PRODUIT** réalisé par l'**ENTREPRISE**.

□ I.2- Définitions de base

- **ENTREPRISE** : La norme [AFNOR 90] définit l'entreprise comme un "Système dirigé et organisé en services dont la finalité est de générer de la valeur ajoutée"

- **CLIENT** : Dans la même source, le client est : "La personne ou l'entité pour qui le produit a été conçu".

- **PRODUIT** : L'Analyse de la Valeur présente le produit comme "ce qui est (ou sera) fourni à un client pour répondre à son besoin" [AFNOR 90]. Cette définition correspond donc au produit final commercialisé par l'entreprise.

De façon plus concrète, dans le domaine industriel, ce produit final peut être un **ENSEMBLE** (ou produit composé ou appareil) ou un **COMPOSANT indivisible** (ou pièce) correspondant à un bien d'équipement, à un bien de consommation durable ou à un composant destiné à être intégré dans un bien d'équipement ou de consommation.

La définition du produit nous amène à expliquer de nouveaux termes :

- **ENSEMBLE** : [BLONDEL 97] définit un ensemble comme "la réunion et l'assemblage de plusieurs pièces élémentaires, constituant un élément à nouveau stockable dans cet état".

Cette définition s'applique aussi à la notion de **SOUS-ENSEMBLE**, dont la principale différence avec l'**ENSEMBLE** est qu'il ne correspond pas à un produit fini.

- **PIECE détachée ou COMPOSANT indivisible** : On appelle pièce tout constituant simple du produit composé de matière et de **VALEUR AJOUTEE**.

- **VALEUR AJOUTEE** : C'est la différence entre la valeur de la production et les consommations de biens et de services fournis par des tiers pour réaliser cette production [BLONDEL 97]. Elle peut donc s'exprimer simplement par l'opération suivante :

$$\begin{array}{r} \text{VENTES (produits, pièces détachées)} \\ - \text{ACHATS (machines, matière première, salaires)} \\ \hline = \text{VALEUR AJOUTEE} \end{array}$$

A ces définitions de base, on peut rajouter quelques autres termes qui seront rencontrés dans ce mémoire :

- **ORDRE de FABRICATION (O.F)** : C'est la commande à l'atelier d'un produit donné, pour une quantité donnée et généralement dans un délai à respecter.

- **STOCK et ENCOURS** : [AFNOR 91] et [AFGI 91] distinguent précisément les STOCKS des ENCOURS :

- Le **STOCK** correspond , à un moment donné et à un endroit donné , à la quantité d'un article non encore utilisé soit pour l'encours, soit pour la consommation, mesurée dans une unité de stockage appropriée. Le stock se situe donc à l'interface des flux externes de l'entreprise.

- L'**ENCOURS** a pour effet ou pour fonction de constituer un "tampon" de régulation des flux internes. De façon plus générale, un encours correspond à l'accumulation d'une différence de flux entre deux ressources. Au point de vue comptable, tous les articles (pièces, sous-ensembles et ensembles) en cours de fabrication sont considérés comme étant sortis du stock fournisseurs ou matières premières et non encore rentrés dans le stock clients ou produits finis. Donc la totalité des articles présents dans le flux interne du processus, qu'ils soient en attente devant une ressource, en transfert d'une ressource vers une autre, ou en transformation sur une ressource, est considéré comme de l'ENCOURS.

- **CAPACITE** : La CAPACITE d'une ressource correspond à la quantité maximale d'unités d'œuvre pouvant être raisonnablement (ou théoriquement) atteinte dans une période donnée et dans le cadre de certaines hypothèses de travail.

Exemple :

- *Chaîne d'embouteillage ayant une capacité de 6000 bouteilles/heure.*
- *Tour à commande numérique ayant une capacité de 36 heures/semaine.*

- **CHARGE** : La CHARGE d'une ressource correspond à une demande planifiée sur une période donnée et exprimée avec la même unité que la capacité. Pour une ressource donnée, la charge doit bien sûr être inférieure ou égale à la capacité.

La définition du **PRODUIT** présentait celui-ci comme correspondant au **BESOIN** d'un **CLIENT**. Pour l'entreprise, deux situations peuvent se présenter, suivant que le produit correspond à :

- un **BESOIN immédiat**, ce qui signifie que le délai de livraison acceptable par le client est nul;
- un **BESOIN différé**, ce qui signifie que le client accepte un délai de livraison non nul. Ce délai doit toutefois être le plus faible possible, car l'entreprise évolue dans un marché où la concurrence est sévère et le délai peut être un critère de décision du client par rapport à des produits similaires d'entreprises concurrentes.

La première situation implique que le produit soit stocké en quantité suffisante sur le lieu de vente. L'entreprise doit donc produire sur des prévisions de vente et constituer un stock de produits capable de répondre à la demande.

La seconde situation implique que l'entreprise produise avec un délai inférieur ou égal au délai acceptable par le client.

Dans [BITEAU 98], cette situation est caractérisée par un ratio d'incertitude **Ri** dont la valeur peut être inférieure à 1 si le lancement de la production se fait à partir d'une commande ferme, et supérieure à 1 si l'entreprise est obligée de commencer sa production avant de savoir exactement quelle sera la demande du client.

Ces deux cas vont être détaillés par la suite, et d'autres définitions complémentaires à celles-ci seront données chaque fois que cela sera nécessaire.

□ I.3- Définition de la production

[GIARD 88] définit la production comme étant une transformation de ressources appartenant à un système productif et conduisant à la création de biens et de services. Les ressources peuvent être de quatre types :

- des équipements (machines, ...),
- des hommes (opérateurs, ...),
- des matières (matières premières et composants),
- des informations techniques ou procédurales (gammes, nomenclatures, fiches opératoires, ...).

La production d'un bien s'effectue par une succession d'opérations consommant des ressources et transformant les caractéristiques morphologiques ou spatiales de "matières".

Comme on l'a déjà vu dans la définition des STOCKS et ENCOURS, un élément fondamental de la production est la notion de **FLUX** entre les différentes ressources. Un flux correspond à une circulation d'entités physiques ou informationnelles au travers d'un processus, caractérisé par un débit. En production, on pourra donc distinguer essentiellement deux types de flux : Les flux de matières et les flux d'informations.

Une autre distinction des flux concerne la localisation des flux dans la chaîne logistique qui comporte trois maillons : approvisionnement, production et distribution. Les flux **EXTERNES** concernent uniquement l'approvisionnement et la distribution, alors que les flux **INTERNES** correspondent à l'ensemble de la production. La Gestion de Production s'intéresse essentiellement aux flux INTERNES de l'entreprise (voir figure 2).

□ I.4- Définition et objectif de la gestion de production

[BLONDEL 97] définit la Gestion de Production (G.P) comme la fonction qui permet de réaliser les opérations de production en respectant les conditions de qualité, délai, coûts qui résultent des objectifs de l'entreprise et dont le but est d'assurer l'équilibre entre :

- le taux d'emploi des ressources,
- le niveau des encours et des stocks,
- les délais.

[MOLET 97] remarque tout à fait justement que cet équilibre est très difficile à obtenir puisque "l'on veut réduire en même temps stocks, délais, pannes tout en accroissant la flexibilité, la variabilité des produits,... autant d'objectifs complexes, multiples mais

souvent contradictoires et dont l'importance relative peut varier à chaque moment". Il en conclut d'ailleurs que "la gestion de production reste, malgré les apports des outils, la gestion des compromis".

Dans [COURTOIS 95], l'objectif principal de la Gestion de Production est de gérer les flux de matières et d'informations par rapport aux objectifs prioritaires définis par la Direction Générale de l'entreprise. Le schéma suivant montre l'ensemble des flux que gère, totalement ou partiellement, la gestion de production.

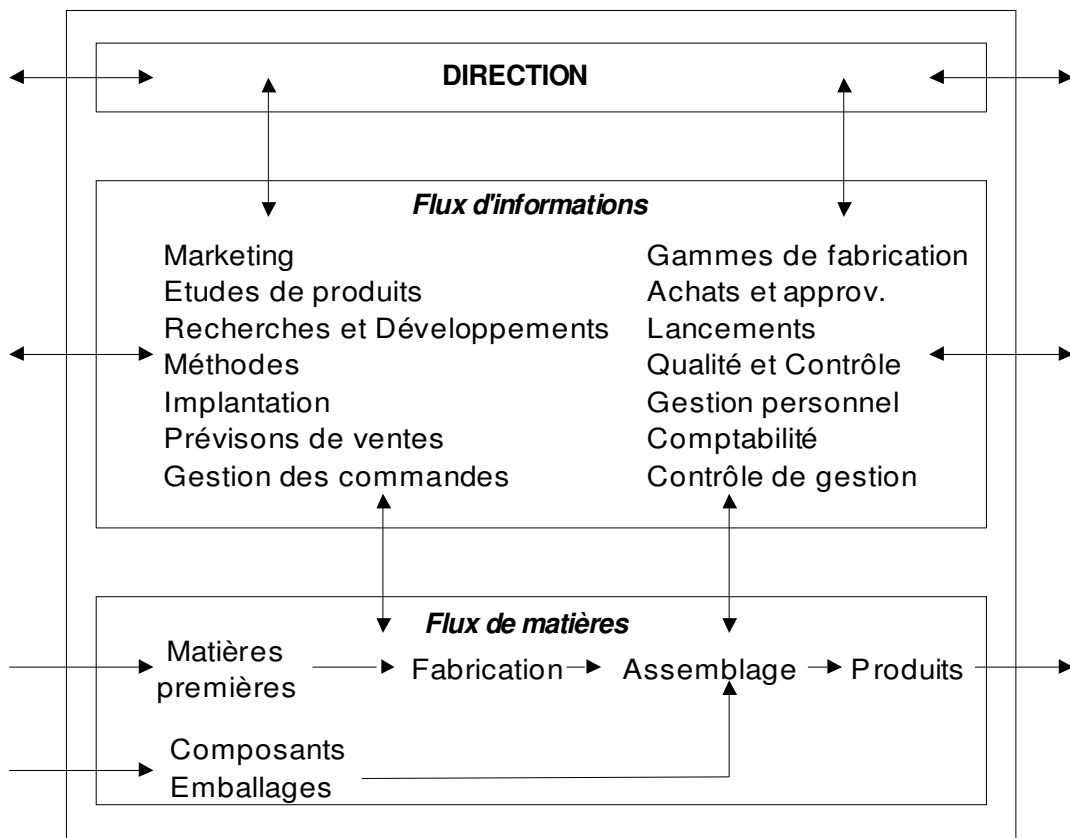


Figure 2 : Les flux informationnels et physiques

II- Evolution de la production et nouvelles règles

□ II.1- Les principales périodes

Dans [COURTOIS 95] et [BLONDEL 97], on distingue trois à quatre périodes qui marquent chacune une évolution de la production industrielle :

1^{ère} période : les débuts de l'industrie (à partir de la fin du 19^{ème} siècle), il y a à peine 100 ans... ! La production est alors proche de l'artisanat : faibles quantités, grande diversité, personnel très qualifié (compagnons).

2^{ème} période : de la première guerre mondiale à 1975, période incluant les fameuses "Trente glorieuses" de l'industrie de 1945 à 1975, pendant laquelle la demande est très importante, et même supérieure à l'offre. Les marges sont confortables et les principales caractéristiques de la production sont les suivantes : fabrication en très grandes séries, faible diversité (on connaît la phrase d'Henry Ford " Le client qui désire une Ford T peut demander n'importe quelle couleur, pourvu qu'elle soit noire"), personnel peu qualifié, travail découpé en tâches élémentaires simplifiées et rapides pour garantir un enchaînement rapide des opérations (voir le début du film de Charlie Chaplin "Les Temps Modernes", qui ne doit pas être aussi caricatural qu'on l'imagine....). Dans ce contexte, pour que l'entreprise existe, il suffit de **PRODUIRE PUIS VENDRE**.

3^{ème} période : Durant cette période transitoire, de 1975 à 1985, l'offre et la demande s'équilibrent, le client a le choix du fournisseur. C'est l'après-choc pétrolier, il est nécessaire de faire des prévisions commerciales, d'organiser les approvisionnements, de réguler les stocks. Il faut alors **PRODUIRE CE QUI SERA VENDU**.

4^{ème} période : Depuis la fin des années 70, les marchés sont fortement concurrentiels et surtout se mondialisent. L'offre est supérieure à la demande et de nouvelles contraintes apparaissent : maîtrise des coûts, qualité, délais de livraisons courts et fiables, produits personnalisables et à faible durée de vie,....Les séries sont toujours importantes, mais très diversifiées : plus de 60 000 variantes pour la Renault 18 à la fin des années 70, plus de 200 000 variantes sur la Peugeot 306 au début des années 90 (environ 250 variantes de moteurs et 1000 variantes de caisses habillées). Toujours dans l'industrie automobile, la taille moyenne annuelle d'une série de même véhicule est de 40 véhicules, le véhicule le plus vendu atteint moins de 8000 exemplaires.

Des pays comme le Japon sont les précurseurs de l'utilisation de nouvelles méthodes dont le mot clé est le "**JUSTE A TEMPS**". L'entreprise doit tendre à **PRODUIRE CE QUI EST DEJA VENDU**.

□ II.2- Les nouvelles règles de la production moderne

Dans le contexte de la 4ème période définie précédemment, l'entreprise est face à plusieurs difficultés :

- Evoluer dans un marché volatile, mal maîtrisé, ou les clients sont imprévisibles, infidèles à une marque et sensibles au délai ou à son respect, à la qualité, au service après-vente.
- Trouver des compromis entre stocks minimums, délais minimums et aléas minimums.
- Réduire les coûts de production, limiter les investissements, disposer de ressources flexibles.
- etc...

III- Typologie de la production

III.1- Structure du produit

Le produit, ainsi que son flux de production, peut avoir plusieurs structures [COURTOIS 95] :

- **Convergente** : c'est le cas des produits réalisés à partir de l'assemblage de composants. Cette structure est caractérisée par une arborescence présentant plusieurs niveaux qui correspondent à des sous-ensembles du produit final. La forme générale du schéma ci-dessous permet de mettre en évidence l'autre appellation donnée à cette structure : "structure en A" [MARRIS 96].

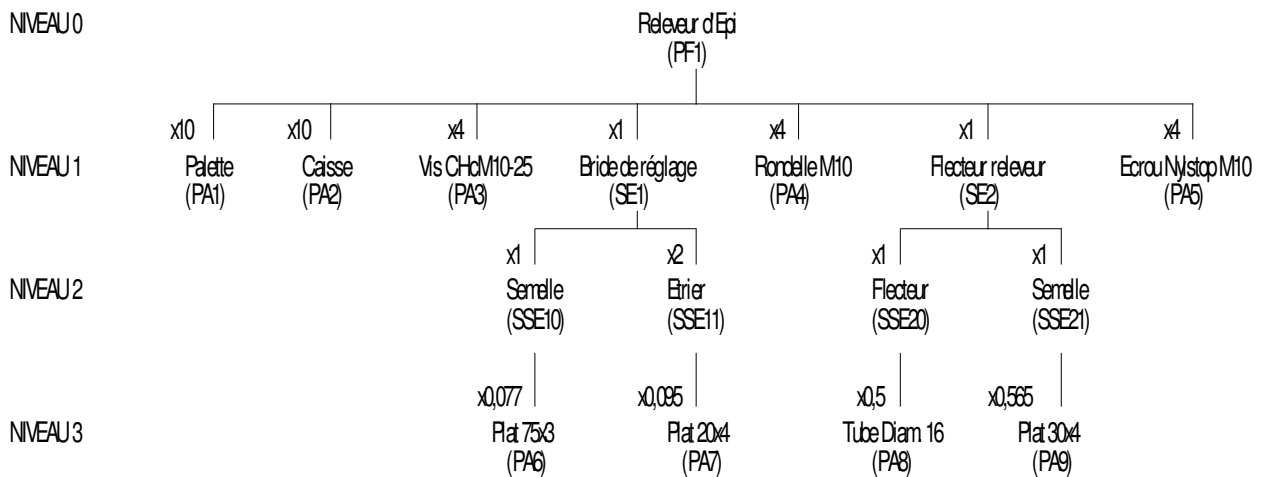


Figure 3 : Structure de produit Convergente

- **divergente ou "structure en V"** : cette structure est celle des produits réalisés à partir de la transformation d'une matière première unique : pétrole, lait, acier, etc..

- **parallèle** : les produits sont réalisés à partir de quelques matières premières faiblement transformées : industries de l'emballage, du pneumatique, ...

- **à points de regroupement ou "structure en T"** : cette structure dérivée de la structure convergente, est typique des produits réalisés à partir de l'assemblage de composants soit spécifiques (structure convergente) soit standards (points de regroupement) : cas de l'industrie automobile, où un même moteur peut être implanté dans plusieurs véhicules différents. Une structure en T cherche à concilier la production de masse et la personnalisation des produits. Il s'agit en fait d'un produit de conception modulaire destiné à permettre la réalisation de produits sur mesure en combinant de différentes façons des sous-ensembles dont certains sont standardisés. Une telle conception conduit ainsi à définir un produit complexe au moyen d'une nomenclature arborescente à plusieurs niveaux, chaque niveau correspondant à un stade de la décomposition

Les structures *divergentes* et *parallèles* correspondent à des usines fabriquant des produits peu variés, en très grandes séries, sur des périodes très longues avec un marché peu fluctuant.

Dans la suite de ce travail, on ne s'intéressera qu'aux structures *convergentes* et à *points de regroupements*, les plus fréquentes, correspondant aux industries manufacturières de type automobile, électronique, électroménager, ameublement, aéronautique, etc..

□ III.2- Modes de production

Les entreprises industrielles qui réalisent des produits implantent et organisent leurs ressources de production en fonction non seulement de la structure du produit, mais aussi du volume de production. On peut ainsi établir une classification des modes de production [WOODWARD 65] [MULKENS 93] [BARANGER 87] :

- **production unitaire ou par projet** : le produit est généralement complexe, nécessite la coordination de plusieurs ressources devant intervenir simultanément ou

séquentiellement afin de livrer au moment convenu. C'est le cas des grands projets industriels ou civils comme la réalisation d'un navire, d'un ouvrage de génie civil, de l'organisation d'une manifestation sportive internationale,

- **production continue** : elle correspond aux produits qui subissent des transformations en continu, par le biais d'opérations parfaitement synchronisées au niveau de leur temps opératoire : aciéries, cimenteries, stations d'épuration, raffineries sont des exemples de processus qui correspondent à cette typologie. Les équipements de production sont dédiés et d'un niveau d'automatisation très élevé.

- **production de masse** : caractéristique des produits à structure en A ou V, réalisés par fabrication et/ou assemblage en très grande quantité, mais avec très peu de variantes. Les ressources de production (hommes et machines) sont donc fortement spécialisées et dédiées à des tâches précises. Le niveau d'automatisation est en général élevé.

Exemple : fabrication de roulements à billes, d'ampoules d'éclairage, etc....

- **production en petites à moyennes séries, répétitives ou pas** : c'est le cas des produits à structure en A ou T, "personnalisables" en fonction des besoins du client : produits de base avec possibilité d'options ou variantes. Les ressources de production sont très polyvalentes, flexibles, capables de passer rapidement d'une production à une autre. Le niveau d'automatisation est généralement faible ou nul.

Dans la suite, on ne s'intéressera qu'aux deux derniers modes de production correspondant aux *produits à structure convergente (en A ou T) fabriqués en série.*

□ III.3- Circulation des produits dans l'atelier

Cette typologie est très liée au critère précédent. On distingue 2 à 3 grandes classes : circulation en Job Shop et en Flow Shop [WIDMER 91].

- **Circulation des produits en Job Shop** : C'est le cas d'un atelier général, destiné à fabriquer une grande variété de pièces. Les produits circulent de machines en machines suivant un routage correspondant à leur gamme de fabrication.

- **Circulation des produits en Flow Shop** : tous les articles suivent le même cheminement. C'est le cas des lignes transferts dédiées (voir chapitre V.3) où les articles "visitent" systématiquement chaque poste de travail implanté sur la ligne, et toujours dans le même ordre. Cependant, les chercheurs ont décomposé cette classe en plusieurs sous-classes parmi lesquelles on trouve : les circulations de type "*pur flow shop*", où tous les temps opératoires sont positifs, les circulations de type "*flow shop généralisé*" où certains temps opératoires pouvant être nuls (la pièce ne devant pas subir une opération sur une machine particulière), et enfin les circulations de type "*flow shop de permutation*" dans lesquelles la séquence des pièces est la même sur toutes les machines (pas de dépassement autorisé).

En complément à la description des critères énoncés précédemment, on peut remarquer les particularités suivantes :

- Certains produits de grande série ont une durée de vie de plus en plus courte, ce qui implique qu'au niveau des équipements de production, il est nécessaire de prévoir une possibilité de réadaptation afin d'étaler les investissements.

- On constate une augmentation très importante des produits de grande série "à option" ce qui correspond à un mode de production de masse avec la flexibilité des petites à moyennes séries. En effet, les lignes de production doivent être capables de réaliser en grande quantité des produits pouvant posséder de nombreuses variantes et même des produits différents appartenant toutefois à une même famille. L'exemple le plus représentatif de ce mode de production est l'industrie automobile, essentiellement au niveau du processus d'assemblage.

- Vis à vis du client et du marché, une des principales qualités de l'entreprise industrielle est son temps de réponse qui doit être bien sûr le plus faible possible. On parle aujourd'hui de "*réactivité*". Le régime transitoire entre deux productions stabilisées comporte des opérations de réglage, de changement d'outillage, de modification de cadence, qui doivent se faire le plus rapidement possible afin de réduire la période d'arrêt.

□ III.4- Relation avec le client

Cette typologie distingue 2 catégories de relation avec le client [GIARD 88] [COURTOIS 95] : Production sur stock, production sur commande et mixte.

- Production sur Stock :

Une production sur stock est déclenchée par anticipation d'une demande s'exerçant sur un produit dont les caractéristiques sont définies par le fabricant. Ce type de production s'applique dans les cas suivants :

- l'éventail des produits finis est restreint ;
- la demande des produits est prévisible;
- le délai de fabrication est supérieur au délai admissible par le client;
- la saisonnalité du produit est trop forte pour justifier le maintien de ressources en hommes et machines qui seraient excédentaires une bonne partie de l'année.

L'équation qui régit le stock de produits finis est :

Stock produits finis	=	stock de sécurité	+	quantité de produits fabriqués	-	quantité de produits vendus
----------------------------	---	-------------------------	---	--------------------------------------	---	-----------------------------------

Le stock de sécurité permet d'amortir les variations de ventes par rapport aux prévisions.

L'objectif est de minimiser ce stock, soit à partir de prévisions de ventes très précises, soit grâce à une très grande réactivité du système de production lui permettant de se réguler par rapport aux ventes effectives (si les ventes diminuent, ralentir la production; si les ventes augmentent, accélérer la production).

La production sur stock repose donc sur une prévision très fine de la demande.

L'inertie du système de production pouvant être très importante, il est primordial d'avoir une boucle de retour des ventes effectives très rapide.

On peut assimiler un tel système à une boucle de régulation (voir figure 4) [BLONDEL 97]. Le système de production est piloté par l'écart entre les ventes effectives et les prévisions :

- Si l'écart est positif, l'excédent des ventes est tiré du stock de sécurité.
- Si l'écart est négatif, il faut ralentir ou même arrêter le système afin de ne pas augmenter le stock de produits finis.

Dans les deux cas, il est nécessaire d'augmenter ou de ralentir la cadence de production afin de tendre vers un écart nul entre prévisions et ventes effectives.

Comme le montre le schéma-blocs de la figure 4, le système de production peut rapidement s'avérer instable, si ce n'est chaotique [MASSOTTE 94] [ALFARO 97] quand les variations de la demande sont importantes. Mais comme le démontre d'ailleurs P. Massotte, même avec une demande constante, le système a une forte probabilité de comportement chaotique !

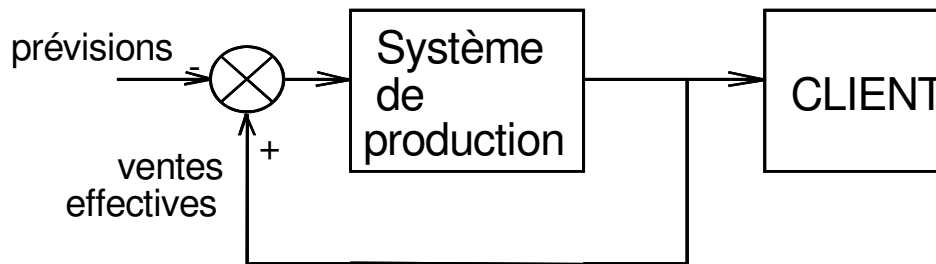


Figure 4 : Boucle de régulation du système de production

- Production sur commande :

Cette situation semble être la plus favorable à l'entreprise car elle lui permet de produire uniquement sur commande, mais à condition que le délai du cycle (achat + fabrication + assemblage + livraison) soit inférieur ou égal au délai acceptable par le client.

En théorie, si la condition précédente est remplie, aucun stock n'est nécessaire.

Dans certains cas où le délai de fabrication est trop long, il est possible d'anticiper l'achat et la fabrication des composants, et de procéder à l'assemblage dès que l'on a une commande ferme. Ceci implique aussi d'avoir de bonnes prévisions de ventes afin de ne pas constituer de stocks excessifs de composants. On parle alors de **Production MIXTE**, qui est de plus en plus rencontrée. En effet, dans ce type de production, le produit fini peut être personnalisé le plus en aval possible, tout en étant constitué de composants et sous-ensembles standards.

Le principal inconvénient de la production sur commande ou mixte est l'obligation d'avoir un système de production présentant le minimum d'aléas afin de ne pas perturber les délais. La maintenance préventive et prédictive doit donc jouer un rôle très important.

C'est ce type de fonctionnement que l'on appelle aujourd'hui le "**JUSTE A TEMPS**" (JAT) dont le principe est :

**IL FAUT ACHETER OU PRODUIRE SEULEMENT CE DONT ON A
BESOIN, QUAND ON EN A BESOIN.**

La différence avec la production sur stock réside dans la chronologie des opérations de ventes et de production :

Production sur stock : PRODUCTION puis VENTE;

Production sur commande : VENTE puis PRODUCTION.

Comme on l'a vu dans le chapitre II.1, nous sommes maintenant entrés dans une période dans laquelle les entreprises sont amenées à fonctionner en Juste à Temps, avec des productions très diversifiées, sur commande ou mixtes.

Bien que ce contexte semble plus favorable à l'entreprise vis à vis de sa trésorerie, il implique une gestion très précise, une forte réactivité de l'entreprise, une grande rationalisation des équipements de production, une grande maîtrise des flux physiques et informationnels, ...

La figure 5 de la page suivante présente un modèle de représentation des flux physiques et informationnels d'une production mixte en Juste à Temps. Il s'agit d'un macro-modèle, d'un niveau de description très général, qui permet de situer l'un des centres d'intérêt de ce travail : le processus d'assemblage des composants et sous-ensembles.

Un modèle de représentation des flux physiques et informationnels d'un processus d'assemblage en JAT

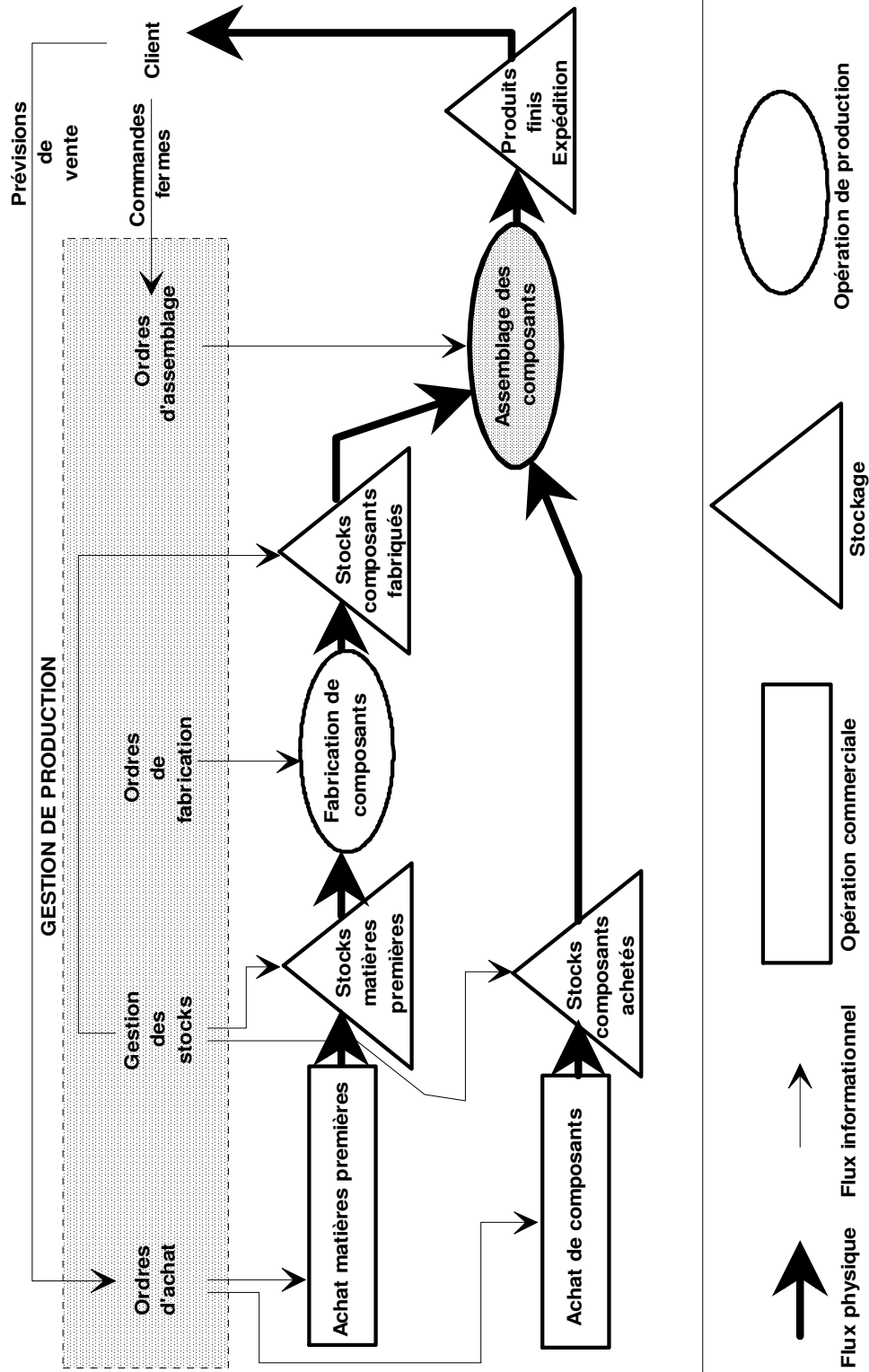


Figure 5 : Macro-modèle des flux dans un processus d'assemblage

IV- Synthèse des différents modes de pilotage des flux de production

Les principaux flux de production qui traversent l'entreprise sont :

- **les flux physiques** : composants achetés, fabriqués, pièces de rechange, sous-ensembles, produits finis, etc...
- **les flux informationnels** : commandes, ordres de fabrication, ordres d'approvisionnement, gammes, fiches opératoires, fiches de suivi, etc...

Les principaux objectifs d'amélioration des entreprises industrielles fonctionnant en Juste à Temps sont :

- **de simplifier les flux physiques** par l'amélioration de l'implantation des équipements; Plusieurs méthodes ont été développées pour répondre à cet objectif : Technologie de Groupe (TG) [MUTEL 92] , méthode des chaînons, algorithme de Kusiak, etc....

- **d'accélérer les flux physiques** en évitant les pannes, en diminuant les temps de changement de série, en améliorant la qualité, en développant la polyvalence des opérateurs, etc... La maintenance préventive et prédictive, le SMED (Single Minute Exchange of Die) [SHINGO 90], la TPM (Total Productive Maintenance) [NAKAJIMA 89], le SPC (Statistical Process Control)[VINCENT 93], sont des méthodes de plus en plus mises en œuvre dans les entreprises.

- **de rendre les flux informationnels plus cohérents et de faciliter la communication** et l'échange de données. C'est ce dernier objectif qui présente le plus de difficultés pour les entreprises, compte tenu de la quantité, de la diversité et de la difficulté à fiabiliser des informations, mais aussi à cause de l'éclatement des données dans l'ensemble des services. La Gestion des Données Techniques (GDT), l'Echange de Données Informatisées (EDI), la norme STEP, etc... répondent, tant bien que mal, aux interrogations des industriels dans ce domaine encore peu débroussaillé.

On distingue différents modes de pilotage des flux de production qui sont conditionnés par les flux d'informations qui contrôlent la production.

□ IV.1- Pilotage centralisé

Ce mode de pilotage s'applique aussi bien à la production de masse qu'à la petite série, chaque fois que des produits définis et planifiables existent. Le pilotage est réalisé à partir d'un programme directeur de production (PDP) sur un horizon de planification de

plusieurs semaines à plusieurs mois (plus précisément de 1,5 à 2 fois le cycle de production). Ce type de pilotage s'est particulièrement développé avec la création du MRP (Material Requirements Planning ou calcul des besoins nets) puis du MRP2 (Manufacturing Resource Planning ou Management des Ressources de Production) [ORLICKY 75],[WIGHT 84]. Le flux d'informations est déconnecté du flux de produits de par sa centralisation.

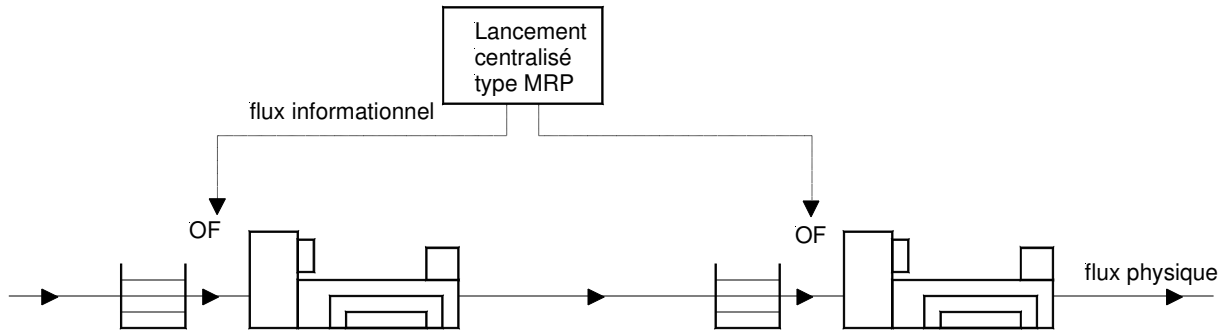


Figure 6 : Pilotage centralisé (type MRP)

□ IV.2- Pilotage par l'aval et décentralisé

C'est le mode de pilotage développé par l'industrie japonaise au début des années 60 et baptisé "Kanban", du nom des étiquettes servant d'ordres de fabrication [SHINGO 90].

Il est adapté aux productions en séries répétitives avec une demande régulière. C'est essentiellement un système de pilotage de l'atelier, ce qui fait qu'il est souvent couplé avec un système centralisé de type MRP. Le flux de produits est tiré par l'aval (le client), ce qui correspond à un flux d'informations remontant à contre-courant le flux physique.

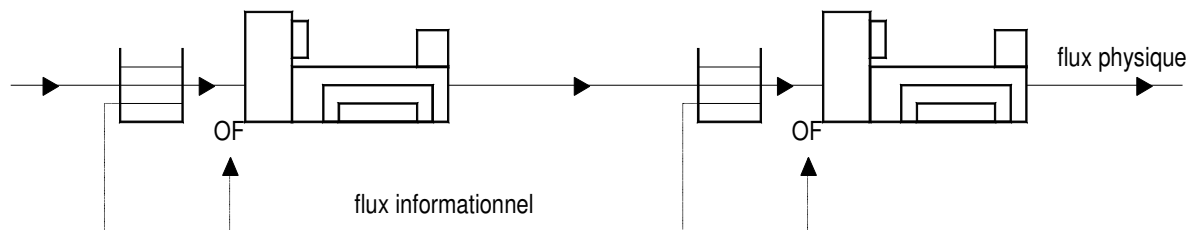


Figure 7 : Pilotage par l'aval et décentralisé (type Kanban)

□ IV.3- Pilotage par l'amont et décentralisé

Le flux de produit est dans ce cas poussé par l'amont. Ce pilotage est bien adapté à la fabrication sur commande de produits non planifiables, par petites séries non nécessairement répétitives [MULKENS 93].

□ IV.4- Pilotage synchronisé entre flux physique et flux informationnel

Ce pilotage est particulièrement adapté aux processus dont le transfert de produit est automatisé (système de convoyage, chariots filoguidés, etc...). On peut ainsi associer à chaque produit toutes les informations nécessaires à son élaboration (par exemple à partir d'étiquettes électroniques) [MULKENS 93].

□ IV.5- Pilotage par les contraintes ou méthode OPT

La méthode *OPT* (*Optimised Production Technology*) est née au Etats-Unis dans les années 80 [GOLDRATT 92]. Elle s'adresse, à priori, à des entreprises déjà partiellement informatisées qui utilisent un logiciel de type MRP. Mais les neuf règles de bon sens qui constituent son principe sont applicables par n'importe quelle entreprise, sans avoir nécessairement recours à l'informatique. La caractéristique de cette méthode est de prendre pour postulat l'existence d'une contrainte, un "goulet d'étranglement", quelque part sur la ligne de production : manque de personnel, machine inadaptée, etc.... Débit de sortie et niveau des stocks dépendent de ce goulet qui va donc déterminer le niveau d'utilisation des autres ressources, non critiques, appelées "non goulet". [MARRIS 95] a prolongé les travaux de [GOLDRATT 92] et défini cette méthode comme la *Théorie des Contraintes* ou *Management Par les Contraintes* (TOC ou MPC).

Les neuf règles de cette théorie :

1- Equilibrer les flux et non les capacités.

Bien que cela semble être judicieux, il est pratiquement impossible d'avoir un équilibrage parfait des capacités des différentes ressources d'une ligne de production, ne serait ce qu'à cause des aléas (pannes, défauts, rupture d'approvisionnement, etc...), mais aussi parce que la gamme de production n'est pas découpée en phases et opérations de

durées égales. Cette première règle propose donc de piloter la production à partir du flux de la ressource goulet.

2- Le niveau d'utilisation d'un non-goulet n'est pas déterminé par son propre potentiel, mais par d'autres contraintes du système.

3- Utilisation et plein emploi d'une ressource ne sont pas synonymes.

4- Une heure perdue sur le goulet est une heure perdue pour tout le système.

Cette règle permet, dans un premier temps, de limiter les actions préventives, de surveillance ou d'amélioration uniquement à la ressource goulet. En effet, il ne faut pas que la ressource goulet tombe en panne, qu'elle ne soit plus approvisionnée par son fournisseur. Il faut qu'elle ait des temps de changement de série rapide. En clair, il faut qu'elle ait une charge aussi proche que possible de sa capacité, donc un rendement opérationnel proche de 1.

5- Une heure gagnée sur un non-goulet est un leurre.

Si une amélioration doit être réalisée sur une ressource non-goulet, elle ne produira aucun effet sur l'ensemble du flux. Il est par exemple inutile de réduire le temps de changement de série ou de moderniser une ressource non goulet.

6- Le goulet détermine à la fois le débit de sortie et le niveau des en-cours.

Cette règle peut s'interpréter de la façon suivante : si on désire réduire le délai de production et la taille des en-cours, il ne suffit pas d'augmenter la taille et la fréquence des approvisionnements.

7- Souvent le lot de transfert ne doit pas être égal au lot de production.

C'est le cas du chevauchement d'opérations [AFNOR 91], où l'on transfère du poste amont au poste aval une quantité inférieure à celle qui est lancée.

8- Les lots de fabrication doivent être variables et non fixés.

Il s'agit de fractionner les lots en cours de cycle de production pour obtenir encore plus de souplesse. Cette règle s'oppose complètement à la notion de série économique de Wilson... Si les temps de changement de série sont courts, si les machines sont bien implantées, si l'on pilote la production par le flux de la ressource goulet, il n'y a plus besoin de créer des séries économiques !

9- Etablir la planification en prenant en compte simultanément toutes les contraintes de capacité et de priorité.

On ordonnance (en capacité finie) en priorité la production des ressources goulets. Les autres ressources peuvent être ordonnancées par des règles simples, puisqu'elles sont surcapacitaires.

La devise de la Théorie des Contraintes :

La somme des optimums locaux n'est pas l'optimum global du système .

L'identification des ressources goulets :

- Une ressource dont le niveau d'en-cours amont est important par rapport aux autres ressources est peut être un goulet.
- Les produits finis livrés en retard possèdent des composants réalisés sur la ressource goulet.
- La ressource goulet est celle qui a le taux d'occupation = charge/capacité le plus élevé.

V- Typologie des flux physiques et implantation des lignes de production

- V.1- Flux physique avec opérations et manutentions synchronisées (ou liées) "Paced flow lines"

Ce type de flux correspond à une production de masse (voir chapitre I.1) réalisée sur des équipements de production automatisés : transferts rotatifs ou linéaires.

Les produits présents à chaque poste de travail sont traités de façon simultanée, avant d'être transférés simultanément aussi vers le poste suivant [ANDERSON 97] .

L'implantation des postes de travail peut être circulaire, disposés autour d'un plateau tournant qui comporte en général de 4 à 24 positions. A un instant donné, à chaque position correspond une phase de la gamme de fabrication ou d'assemblage du produit.

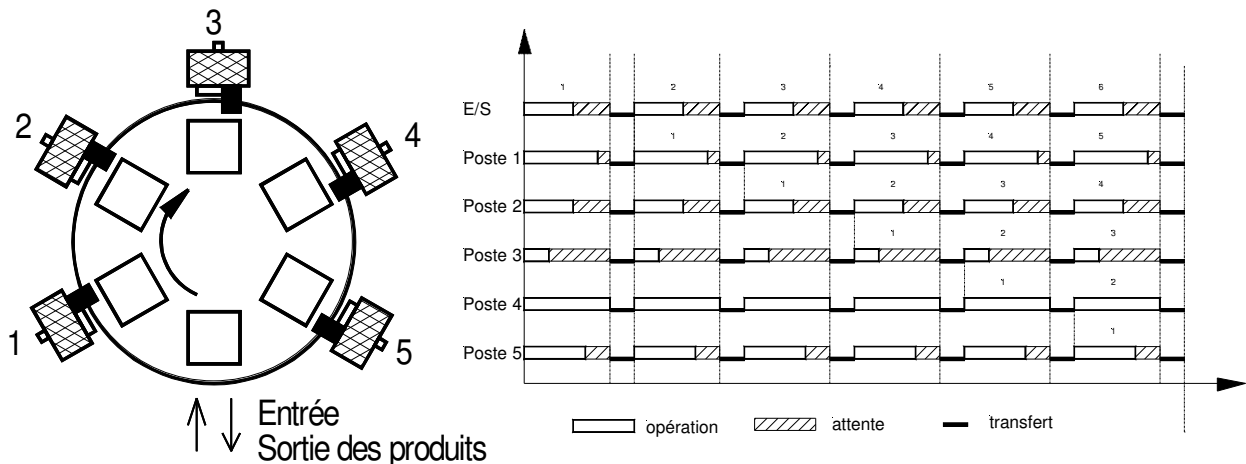


Figure 8 : Transfert et opérations synchronisées

L'implantation peut être aussi linéaire, ce qui permet une accessibilité accrue (côtés et dessus du produit) ainsi qu'un nombre d'opérations plus élevé par rapport à l'implantation circulaire.

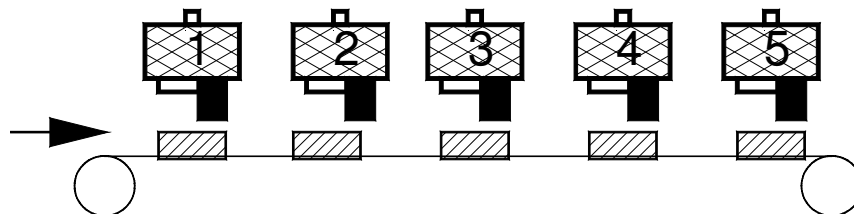


Figure 9 : Implantation Linéaire d'un Transfert Synchronisé

La cadence de production est en général très élevée (600 à 20000 produits/heure), bien que conditionnée par la cadence de l'opération la plus lente. Il est donc préférable d'équilibrer les temps opératoires de chaque poste afin de réduire les temps d'attente et de tendre vers un flux continu.

Le principal inconvénient de cette implantation réside dans les pannes et arrêts que peuvent subir les postes de travail. En effet, toute la production est stoppée en cas de problème survenant à l'un des postes de travail.

D'autre part, les équipements de production sont peu évolutifs et peu flexibles. Chaque produit passe par les mêmes postes de charge, dans le même ordre. On peut simplement prévoir que certains produits identifiables ne subissent pas certaines opérations, mais l'ordre est respecté. *Les équipements de production sont dédiés à un produit.*

Dans son article, [LITTLE 94] appelle ce type de ligne : "Dedicaced flow line"

□ V.2- Flux physique avec opérations et manutentions asynchrones (ou libres) "Unpaced flow lines"

Ce type de flux correspond à une production de masse ou en petites à moyennes séries avec une circulation des produits de type *Job Shop* ou *Open Shop*. Les équipements de production et de transfert peuvent être automatisés ou pas et être utilisés pour traiter différents produits.

Le principe général de ce flux est que le transfert des produits d'un poste de charge à un autre est indépendant des opérations qui s'y déroulent : les opérations de transfert ne sont pas synchronisées avec la fin de toutes les opérations de valeur ajoutée. C'est le type de flux qui correspond à la réalisation de produits dont les différentes opérations s'effectuent à des cadences très différentes [ANDERSON 97].

En ce qui concerne l'implantation des équipements, il faut distinguer au moins deux cas en fonction du niveau d'automatisation du système de transfert (ou de manutention) :

- **manutention non automatisé;**

- **manutention automatisé.**

V.2.a) Manutention non automatisée

Ce type de manutention se rencontre essentiellement dans les processus comportant des opérations d'usinage ou de transformation de la matière.

Les équipements sont implantés :

- **de façon fonctionnelle**, par ateliers spécialisés (voir figure 10). On regroupe les machines réalisant le même procédé : atelier soudage, atelier peinture, atelier fraisage, atelier montage , etc.... Cette implantation qui correspond à une organisation taylorienne, a dominé depuis le début du siècle et continue aujourd'hui à prévaloir dans les ateliers de sous-traitance réalisant des produits à la commande de façon non répétitive. En effet, il est possible de fabriquer tous les types de produits utilisant les moyens de l'atelier sans perturber davantage le flux.

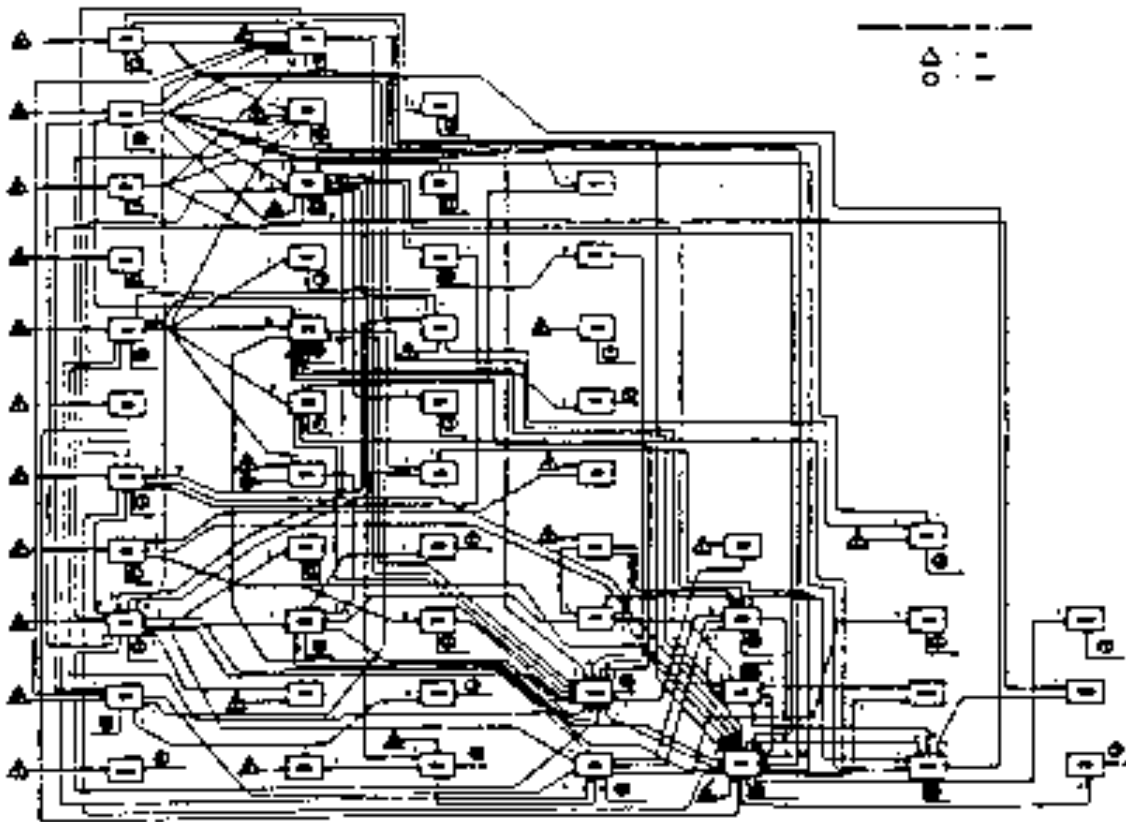


Figure 10 : Implantation fonctionnelle

- **en cellules ou îlots de fabrication.** Cette implantation est constituée de petits ateliers de production composés de machines différentes afin de réaliser entièrement ou partiellement une famille de pièces présentant des similitudes morphodimensionnelles ou de processus. Les machines sont généralement implantées dans l'ordre de la gamme dominante (gamme "mère") de la famille, de telle sorte que le flux physique soit le plus continu possible, avec le minimum de transfert ou d'attente entre deux postes (figure 11).

Ce type d'implantation est en plein développement dans les ateliers d'usinage qui travaillent par petites séries répétitives de produits variés.

La recherche est très active dans ce domaine où des méthodes d'optimisation de l'implantation sont développées.

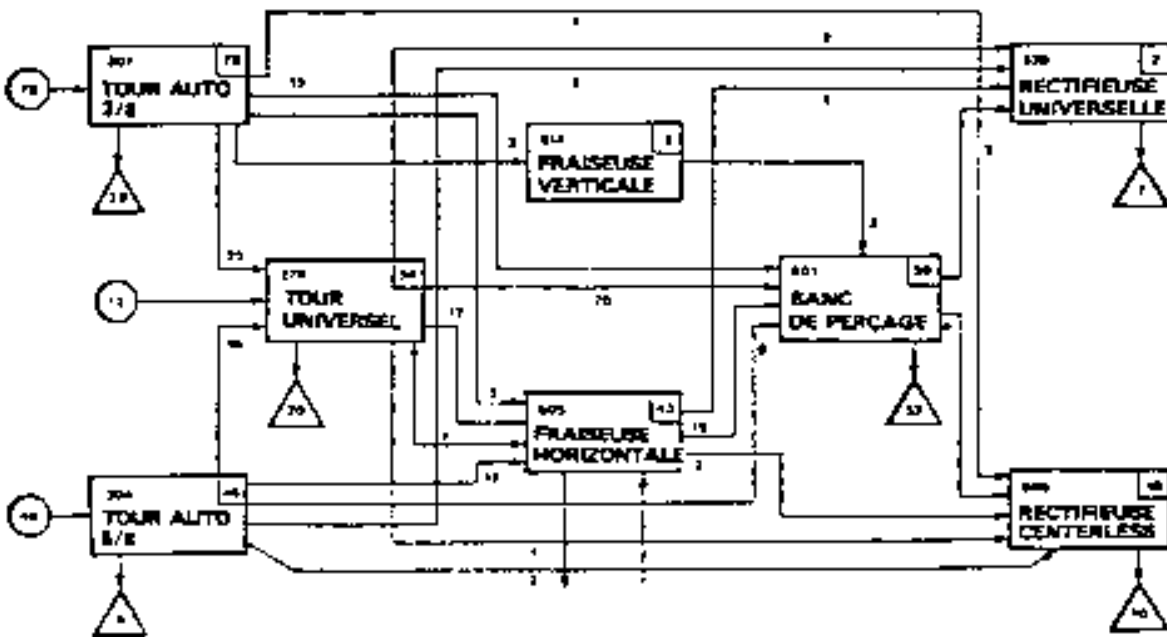


Figure 11: Implantation en îlot

V.2-b) Manutention (ou transfert) automatisée

Le transfert automatisé constitue l'ossature d'un processus d'assemblage de composants. En effet, bien que l'on rencontre quelques cas dans des ateliers flexibles d'usinage, le transfert asynchrone automatisé (Power-and-Free Transfer) est devenu l'élément essentiel des lignes d'assemblage correspondant aux critères de production suivants :

- flexibilité importante : assemblage d'une famille de produits en petites séries;
- cadence moyenne : 20 à 60 opérations par minute;
- différence de temps opératoires sur les différents postes de travail;
- faible durée de vie du produit assemblé;
- fréquents changements de produits;
- qualité optimale;
- automatisation progressive du processus, les postes de travail pouvant être manuels ou automatisés;

Ces différents critères correspondent à la grande majorité des produits assemblés de nos jours.

Ce type de système de manutention est appelé "Mixed Model Flow Line" et " Flexible Flow Line" dans la littérature [LITTLE 94], et sera donc analysé plus en détail dans la partie 2 (Caractéristiques des processus d'assemblage).

□ V.3- Autre typologie

[LITTLE 94] propose une autre typologie des systèmes d'assemblage, basée non pas sur le flux physique et l'implantation, mais sur la structure et la technologie du transfert. Cette typologie comporte deux grandes familles :

V.3-a) Les lignes transfert (Flow Lines) où le produit se déplace le long de la ligne et les composants sont ajoutés par étapes successives sur plusieurs postes de travail.

Cette famille est composée de trois catégories de lignes :

- **les lignes transfert dédiées (dedicated flow lines)** : elles sont dédiées, comme leur nom l'indique, à l'assemblage d'un seul et même produit. Chaque poste de

travail ne réalise qu'une seule opération et reçoit les produits toujours dans le même ordre. Sur ce type de ligne, la seule flexibilité consiste à ne pas réaliser certaines opérations, le produit passant quand même sur le poste de travail. Les postes de travail sont généralement reliés par un dispositif de transfert synchrone pas à pas, programmé pour libérer un produit seulement si le poste aval est prêt à le recevoir. De la même façon, les produits ne peuvent entrer sur la ligne que si le premier poste est libre.

- **Les lignes transfert multiproduits (mixed model flow lines)**, dont le principe est identique aux lignes transfert dédiées, à la différence que les postes de travail sont prévus pour réaliser l'assemblage de plusieurs types de produits (mais appartenant à la même famille). La principale différence réside donc dans les dispositifs de distribution des composants aux postes de travail qui sont adaptés pour servir plusieurs types de composants pour plusieurs types de produits.

- **Les lignes transfert flexibles (flexible flow lines)** : Les postes de travail sont dans ce cas reliés par un convoyeur asynchrone (non-synchronous material handling system) qui permet d'accumuler des produits entre deux postes de travail pour maintenir un niveau d'utilisation satisfaisant. Chaque poste de travail est constitué de plusieurs machines identiques en parallèle toutes capables de réaliser plusieurs opérations sur plusieurs produits. Un poste de travail ne peut libérer un produit que si le convoyeur d'accumulation aval (buffer) a une position de libre. Si ce n'est pas le cas, le produit est bloqué (blocking). A l'inverse, un poste de travail aval peut être en attente (starvation) par manque de produit provenant du poste amont.

V.3-b) Les cellules flexibles (Flexible Assembly Cells) où les mouvements du produits sont limités voire inexistantes. L'assemblage est donc réalisé sur un seul poste de travail. La possibilité de réaliser simultanément différents sous-ensembles de produits donne à ces systèmes leur flexibilité. Ces sous-ensemble sont ensuite amenés simultanément vers une étape d'assemblage final. Cette famille de systèmes peut elle aussi être divisée en trois catégories :

- **Cellules flexibles d'assemblage avec plusieurs postes de travail (multi-station cells)**, disposés en dérivation sur un transfert asynchrone, et pouvant réaliser tout ou partie des opérations d'assemblage.

- **Cellules flexibles d'assemblage avec postes de travail à opérations multiples (multi-stage assembly cells)**

- **Cellules flexibles à station de travail unique (single-station assembly cell)** : Dans ce cas, le poste de travail est constitué d'un robot d'assemblage qui dispose d'un système de changement d'outil/effecteur. Cela permet au robot d'assembler plusieurs produits par simple changement de pince et de programme.

Conclusion

Cette première partie de thèse a été consacrée à une description générale des principales caractéristiques de la production.

L'état de l'art fait ressortir un grand nombre de configurations possibles au niveau des produits, des processus, des modes de pilotage des flux, des systèmes de transfert et de l'implantation des lignes.

Il était important dans cette partie de montrer l'étendue et la variété des environnements de production pour mettre plus en évidence les configurations qui seront étudiées en détail et qui correspondent aux travaux de recherche.

On a insisté en particulier sur la description des processus d'assemblage car ils sont fréquemment rencontrés et leur étude est donc représentative.

Bibliographie de la partie 1 : "Introduction sur la production industrielle"

[COURTOIS 95] : "Gestion de Production"

Editions d'organisation, 2ème édition, 1995

A. Courtois, M. Pillet, C. Martin

[BARANGER 87] : "Gestion de la production"

Editions Vuibert entreprise, 1987

P. Baranger

[WOODWARD 65] : "Industrialisation organisation : practice and theory"

Editions Oxford University Press, 1965

J. Woodward

[ORLICKY 75] : "Material Requirements Planning"

Editions Mac Graw Hill, 1975

J. Orlicky

[WIGHT 84] : "Manufacturing Ressource Planning"

Editions O.W.C, 1984

O. Wight

[SHINGO 90] : "Le système Shingo"

Editions d'organisation, 1990

S. Shingo

[MARRIS 96] : "Le management par les contraintes"

Editions d'organisation, 1996

P. Marris

[MULKENS 93] : "Les nouvelles organisations productives"

Revue française de gestion industrielle N°3 - 1993, pages 5 à 30

H. Mulkens

[AFNOR 90] : "Analyse de la valeur"
Norme française NF-X 50-150, 1990

[AFNOR 91] : "Concepts fondamentaux de la Gestion de Production - Vocabulaire"
Norme française NF X50-310, 1991

[AFGI 91] : "Dictionnaire des termes de gestion industrielle"
Editions AFGI, 1991
R. Biteau, A. Garreau, M. Gavaud

[MASSOTTE 94] : "Analyse du comportement d'un système de production complexe"
Revue française de gestion industrielle N°2 - 1994, pages 61 à 92
P. Massotte

[MUTEL 92] : "Optimisation des flux d'informations et des flux de matières"
Cours DEA Production Automatisée - ISMCM / ENS Cachan, 1992
B. Mutel

[NAKAJIMA 89] : "La Maintenance Productive Totale (TPM)"
Editions AFNOR Gestion, 1989
S. Nakajima

[BLONDEL 97] : "Gestion de la Production"
Editions Dunod, 1997
F. Blondel

[GIARD 88] : "Gestion de la Production"
Editions Economica, 1988, 2ème édition
V. Giard

[MOLET 97] : "Une nouvelle gestion industrielle"

Editions Hermes, 2^{ème} édition, 1997

H. Molet

[ALFARO 97] : "Vers la mise en évidence d'un régime chaotique d'un atelier flexible..."

2ème congrès Franco-Québécois de Génie Industriel, Albi 1997

M. Alfaro, G. Bonhomme, P. Charpentier, P. Martin

[GOLDRATT 92] : "Le But"

Editions AFNOR, 1992

E. Goldratt

[MARRIS 96] : "Le Management par les Contraintes"

Editions d'Organisation, 1996

P. Marris

[LITTLE 94] : "Automated assembly scheduling : a review"

Computer Integrated Manufacturing Systems, Vol. 7 N°1 02/1994, pages 51 à 61

D. Little, A. Hemmings

[WIDMER 91] : "Modèles mathématiques pour une gestion efficace des ateliers flexibles"

Editions Presses Polytechniques Romandes, 1991

M. Widmer

[ANDERSON 97] : "The management of manufacturing"

Editions Addison-Wesley, 1997

E.J. Anderson

[BITEAU 98] : "Maîtriser les flux industriels"

Editions d'Organisation, 1998

R. et S. Biteau

- PARTIE 2 -
***Caractéristiques des
processus d'assemblage***

Sommaire

Introduction	49
I- Typologie des processus d'assemblage.....	50
<input type="checkbox"/> I.1- Assemblage par "kits".....	50
<input type="checkbox"/> I.2- Assemblage par "Rendez-vous".....	51
<input type="checkbox"/> I.3- Assemblage par "Pièce maîtresse".....	50
II- Analyse fonctionnelle et description technologique des lignes d'assemblage multiproduits à transfert asynchrone automatisé	53
<input type="checkbox"/> II.1- Fonctions générales	53
<input type="checkbox"/> II.2- Fonction TRANSFERER	53
<input type="checkbox"/> II.3- Fonction ASSEMBLER	60
II.3-a) Description de la sous fonction STOCKER.....	61
II.3-b) Description de la sous fonction ALIMENTER.....	63
II.3-c) Description de la sous fonction MANIPULER.....	64
<input type="checkbox"/> II.4- Fonction CHARGER/DECHARGER	65
III- Aspects dynamiques du flux de palettes	66
Conclusion	69
Bibliographie.....	70

Introduction

Dans cette partie, on recentre le travail sur les processus d'assemblage et sur les systèmes physiques permettant de les mettre en œuvre et notamment les lignes d'assemblage multiproduits à transfert asynchrone automatisé, qui font l'objet de cette recherche.

Le chapitre 1 présente une typologie des processus d'assemblage basée sur le mode d'approvisionnement ou de synchronisation des composants. Trois principales classes ressortent de cette typologie, sachant que l'on peut trouver, pour certains produits, un mélange de deux ou trois classes. On porte une attention plus particulière à la classe "Assemblage par pièce maîtresse" qui est la plus représentative dans les processus comportant un transfert asynchrone.

Dans le chapitre 2, on propose de développer un modèle d'analyse fonctionnelle d'une ligne d'assemblage multiproduit à transfert asynchrone automatisé. Pour ce travail, le formalisme SADT est retenu. Parallèlement à l'analyse fonctionnelle, des exemples de solutions technologiques correspondant à ces fonctions sont présentés afin d'illustrer plus clairement la structure physique d'un tel système.

On décrit ensuite dans le chapitre 3 un premier aspect de la dynamique du flux de palettes sur le transfert asynchrone. Ceci est un complément indispensable à l'analyse SADT qui ne fait ressortir que l'aspect statique du système.

I- Typologie des processus d'assemblage

Pour ce chapitre et les suivants, on se limitera aux processus les plus représentatifs de la production actuelle :

- produits assemblés à structure convergente ou à points de regroupement (A ou T);
- production de masse ou en petites à moyennes séries répétitives ou pas;
- transfert inter-opérateur asynchrone.

L'assemblage est un processus qui se retrouve dans presque tous les produits fabriqués et qui peut représenter jusqu'à 40% du coût de revient des produits [CETIM 95].

[LITTLE 94] cite les résultats d'une enquête aux Etats-Unis qui indique que 53% du temps total de production et 22% du coût de main d'œuvre sont consacrés à l'assemblage.

Le produit fini résulte de l'assemblage successif de composants sur un ou plusieurs postes de travail implantés le long d'un transfert asynchrone.

On peut distinguer plusieurs types de processus d'assemblage suivant la façon d'alimenter les composants à assembler [DESCHAMPS 94].

□ I.1- Assemblage par "kits"

L'assemblage par kits consiste à préparer tous les composants à assembler et à les placer en vrac dans un support unique, généralement un simple container, puis d'introduire ce container sur la ligne de transfert. Le container est acheminé vers un des postes de travail - obligatoirement manuel - où l'opérateur réalise l'assemblage des composants présents dans le container. Le produit fini est soit évacué au niveau même du poste de travail, soit réintroduit sur la ligne transfert jusqu'à un poste de déchargement.

Le système de gestion permet aux opérateurs de commander la préparation des "kits" par anticipation, en fonction de leur charge de travail, afin de supprimer les attentes.

Ce type d'assemblage, le plus élémentaire en production industrielle, évite la présence de dispositifs d'alimentation sur les postes de travail, mais impose la préparation des kits. Il est surtout utilisé pour les petites séries de produits variés pouvant être assemblés sur la même ligne. On supprime ainsi les temps de réglage ou de changement de dispositifs d'alimentation, qui peuvent être importants par rapport au temps de fabrication d'une série.

La préparation des kits, presque systématiquement manuelle, est réalisée en temps masqué par rapport aux opérations d'assemblage, ce qui permet de passer très rapidement d'une série à une autre.

Le kit permet une gestion rigoureuse des composants puisqu'on ne sort que le nombre de composants strictement nécessaires à la quantité de produits à assembler.

□ 1.2- Assemblage par "Rendez-vous"

Ce type d'assemblage, surtout rencontré dans la production automobile (lignes de production de moteurs), correspond à un processus mixte (usinage + assemblage).

En effet, les composants, à cause de leur encombrement ou de leur coût unitaire, ne sont pas stockés à l'avance en grosse quantité. Ils sont donc réalisés à la demande, suivant un planning prévisionnel sur des postes d'usinage. Les différents composants nécessaires à l'assemblage d'un produit sont ensuite transférés séparément vers le poste d'assemblage. C'est de cette organisation que vient la notion de "**Rendez-vous**" au poste d'assemblage qui caractérise la nécessité de synchroniser l'arrivée des différents composants requis pour une opération.

La thèse de J.C Deschamps [DESCHAMPS 94] traite particulièrement de la gestion et de l'ordonnancement de ce type de processus qui exigent la prise en compte des temps opératoires mais aussi des temps de transfert afin d'avoir les N composants en même temps sur le poste d'assemblage.

□ 1.3- Assemblage par "Pièce maîtresse"

Ce type d'assemblage est vraisemblablement le plus représentatif de la tendance actuelle dans l'industrie.

Lors de la conception du processus d'assemblage, on identifie sur le produit à assembler la "pièce maîtresse", appelée aussi "composant primaire". Sa fonction, au niveau de la gamme d'assemblage, est de recevoir successivement les différents composants et sous-ensembles nécessaires à l'obtention du produit fini. Pour des raisons d'organisation et de productivité, toutes les opérations du processus d'assemblage ne sont pas réalisées au même endroit. La pièce maîtresse doit donc être transférée d'un poste à un autre.

Les composants sont en général stockés à l'avance sur les différents postes de travail en lots (ordonnés ou en vrac) de taille suffisante pour permettre un fonctionnement sans attente.

Il faut donc être en mesure d'extraire un composant de ce stock, de le mettre en position dans la pièce maîtresse et éventuellement de l'immobiliser (par vissage, serrage, collage, encliquetage, ...).

La tâche d'assemblage, en fonction de sa complexité, est soit manuelle, soit automatisée, soit mixte. Pour des produits qui correspondent aux critères énoncés dans l'introduction du chapitre I, l'automatisation de l'assemblage relève essentiellement de l'utilisation d'un robot (4 à 6 axes) et d'équipements de périrobotique tels que convoyeurs, trémies, etc.... Ces équipements peuvent être rendus particulièrement flexibles par le biais de systèmes de vision artificielle et de reconnaissance de forme, notamment pour l'extraction et l'orientation des composants.

La tâche d'assemblage et le poste d'assemblage seront étudiés plus en détail dans le chapitre II.2 .

Ce type d'assemblage par "pièce maîtresse" et plus précisément :

- ***les produits à structure convergente ou à points de regroupement;***
- ***les productions par petites ou moyennes séries, répétitives ou pas, éventuellement les productions de masse à cadence faible et personnalisation possible du produit;***
- ***les systèmes de transfert inter-postes de type asynchrone automatisé;***
- ***les lignes flexibles d'assemblage, automatisées ou pas, capables de réaliser simultanément plusieurs produits différents et de changer rapidement de production,***

sont le support de ce mémoire de recherche.

II- Analyse fonctionnelle et description technologique des lignes d'assemblage multiproduits à transfert asynchrone automatisé

□ II.1- Fonctions générales

Le système physique correspondant aux critères énoncés précédemment doit assurer les fonctions suivantes : **CHARGER/DECHARGER**, **TRANSFERER**, **ASSEMBLER**.

La figure 1 représente les fonctions générales du système physique suivant le formalisme SADT .

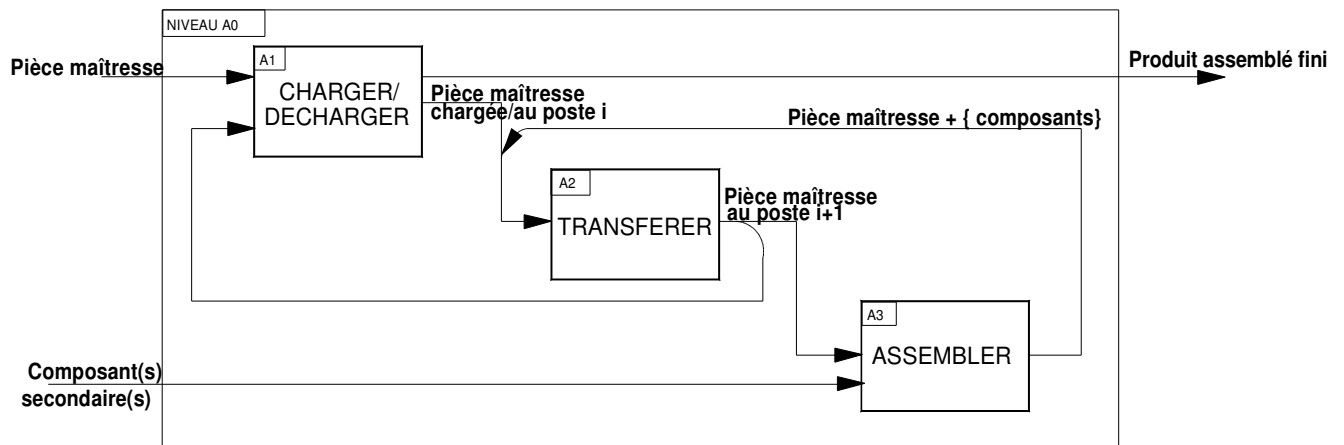


Figure 1 : Fonctions générales d'une ligne d'assemblage

□ II.2- Fonction TRANSFERER

Rappel : Ce chapitre traite uniquement du mode de transfert asynchrone.

Les mouvements des produits entre les postes d'assemblage s'effectuent sans synchronisation temporelle avec les opérations qui s'y déroulent.

La figure 2 représente la fonction **TRANSFERER** suivant le formalisme SADT :

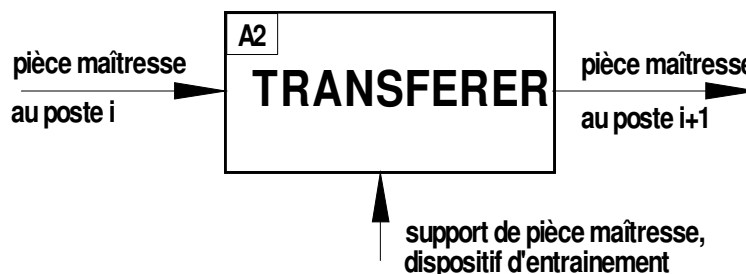


Figure 2 : Fonction TRANSFERER d'une ligne d'assemblage

Description des solutions technologiques utilisées pour cette fonction :

- **Palette** : c'est un support universel de la pièce maîtresse durant son transfert. Elle comporte souvent un posage servant d'interface, réalisé spécialement en fonction des formes et des dimensions de la pièce maîtresse et destiné à la positionner précisément pour que les opérations d'assemblage les plus délicates soient réalisées sans aléas.



Figure 3 : Exemple de palette avec un posage

On peut distinguer deux principaux types d'entraînement des palettes :

- **Les palettes motorisées** : chaque palette est pourvue de sa propre motorisation, dont l'alimentation en courant électrique est réalisée par des batteries embarquées qui se rechargent lors des opérations d'assemblage. Chaque palette dispose aussi de sa propre partie commande, ce qui autorise la programmation du routage, d'accélération, de vitesses limites, de règles de gestion de collision avec les autres palettes, On peut d'ailleurs les assimiler aux plus connus "chariots filoguidés". La différence réside justement dans le dispositif de guidage qui, pour les palettes, correspond à de simples rails comportant lignes droites, virages, aiguillage, ...

Bien que d'un intérêt évident, ce type d'entraînement ne semble pas avoir le succès qu'il mérite, vraisemblablement à cause du prix de chaque palette : environ 20 000 francs !

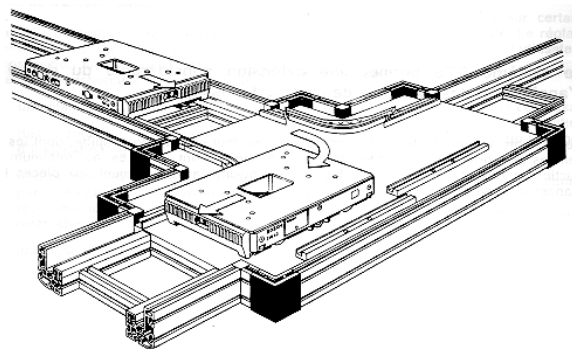


Figure 4 : Exemple d'entraînement à palette motorisée (doc. BOSH)

- **Les palettes libres** : L'entraînement des palettes est réalisé par adhérence avec un convoyeur, qui peut être à bande, à chaîne à écailles, à galets ou à courroies plates ou circulaires. Quelle que soit sa nature, le convoyeur défile en permanence et entraîne par adhérence toutes les palettes à la même vitesse. L'indépendance des palettes entre elles est obtenue par l'utilisation de butées permettant de les arrêter. Dans ce cas, le convoyeur glisse sous la palette sans l'entraîner. Ce dispositif offre la possibilité de créer des zones d'accumulation des palettes, en amont des postes de travail par exemple.

D'un coût bien plus faible (environ 1000 à 3000 francs / palette) que l'entraînement à palettes motorisées, c'est actuellement la solution la plus en vogue dans le milieu industriel.



Figure 5 : Exemple d'entraînement par palettes libres

La majorité des fournisseurs de systèmes de transfert asynchrone s'est orientée depuis peu vers une offre en rapport avec les besoins des concepteurs de lignes d'assemblage : proposer des éléments modulaires permettant non seulement la réalisation sur mesure, mais aussi l'ajout ou la suppression rapide d'éléments de convoyage.

Le concept de **T**ransfert **L**ibre **M**odulaire (TLM) correspond à cette évolution en matière de système de transfert pour les lignes d'assemblage et de nombreux fournisseurs proposent ce concept : ELCOM, BOSCH, SKF, M2A, PRODEL,

Les principales caractéristiques de ce transfert sont :

- la circulation libre, asynchrone et automatique des palettes entre les postes de travail,
- la possibilité d'arrêter les palettes sur un obstacle (butée),
- la possibilité d'accumulation sur le convoyeur en amont et en aval du poste de travail,
- la possibilité d'avoir des postes manuels et des postes automatisés sur la même ligne,
- des architectures multiples [YING-CHIN 98] :

ê **ouverte** : les produits finis sont déchargés en bout de ligne ;

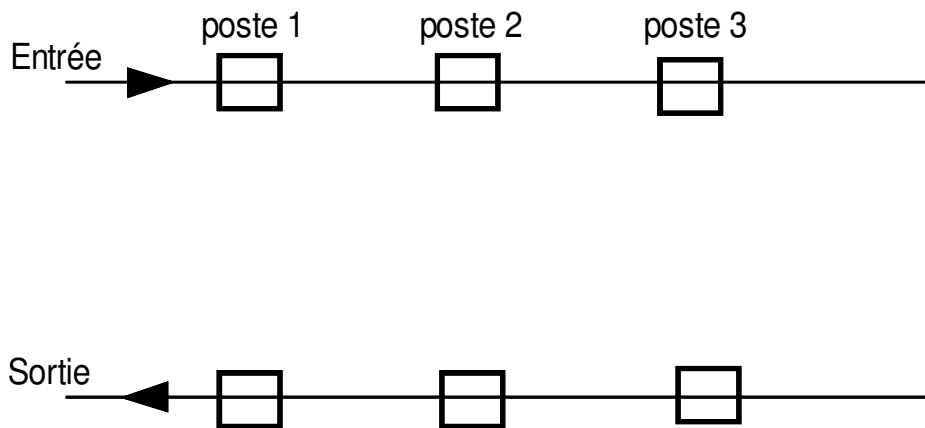


Figure 6 : Exemple d'architecture de ligne ouverte

ê **fermée** : un produit peut éventuellement effectuer plus d'un tour (par exemple s'il doit repasser plus d'une fois sur un même poste de travail ;

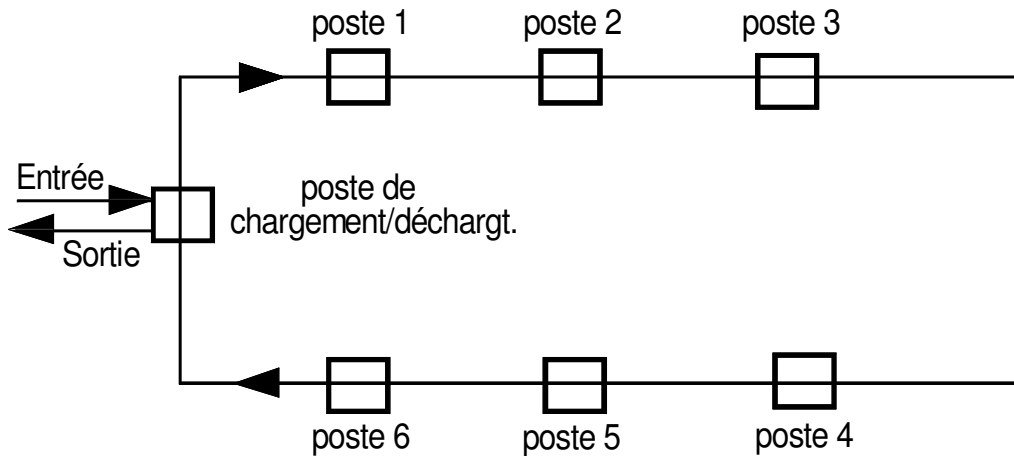


Figure 7 : Exemple d'architecture de ligne fermée

ê à **une seule piste avec postes de travail en série**. Les palettes sont bloquées et indexées au niveau du poste de travail. Cette architecture est intéressante si les temps opératoires sont proches et si toutes les palettes doivent s'arrêter à chaque poste de travail.

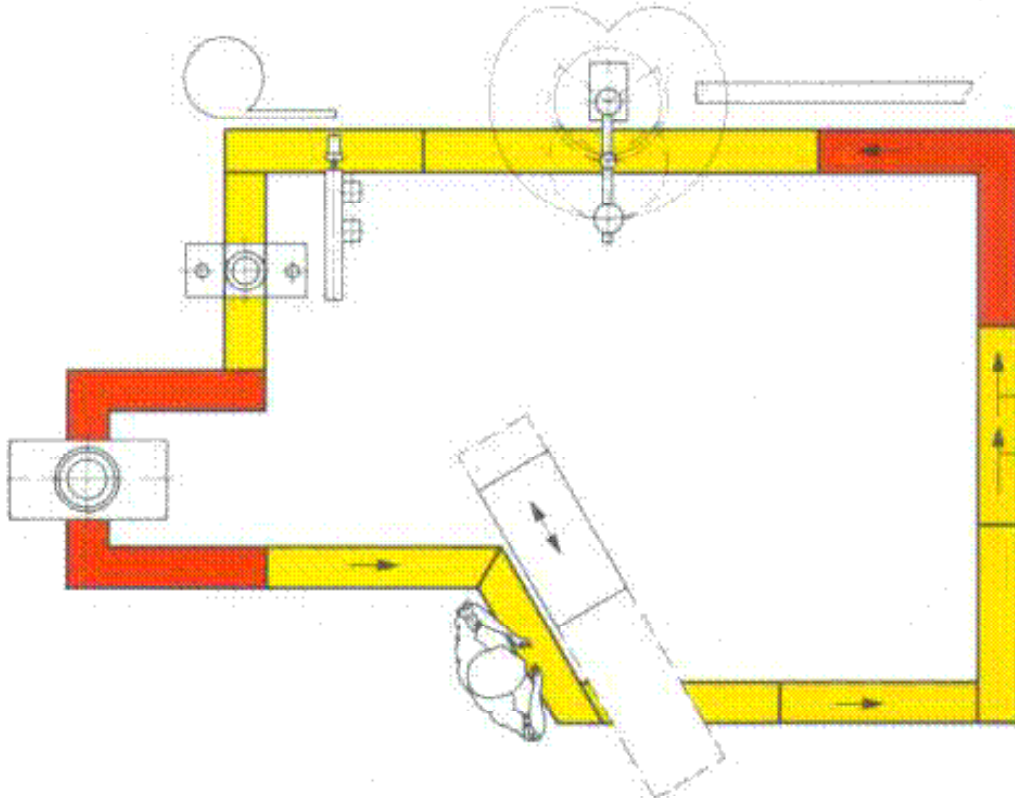


Figure 8 : Exemple d'architecture de ligne à une piste avec postes en série

ê à **une seule piste avec postes de travail en dérivation**. Cette architecture est celle qui correspond le plus au besoin de production simultanée de produits variés en quantités moyennes. En effet, les postes de travail en dérivation permettent de dériver uniquement les produits concernés par l'opération sans perturber le flux des autres produits. D'autre part, cette solution permet de créer des files d'attente locales, en amont de chaque poste, afin de réduire les temps d'attente.

ê à **deux pistes parallèles avec postes de travail en dérivation**. C'est une évolution de l'architecture précédente, permettant le passage direct d'une palette d'un poste à un autre. Cette solution, bien que plus attrayante au niveau du flux, limite les possibilités d'évolution de la ligne et augmente son coût.

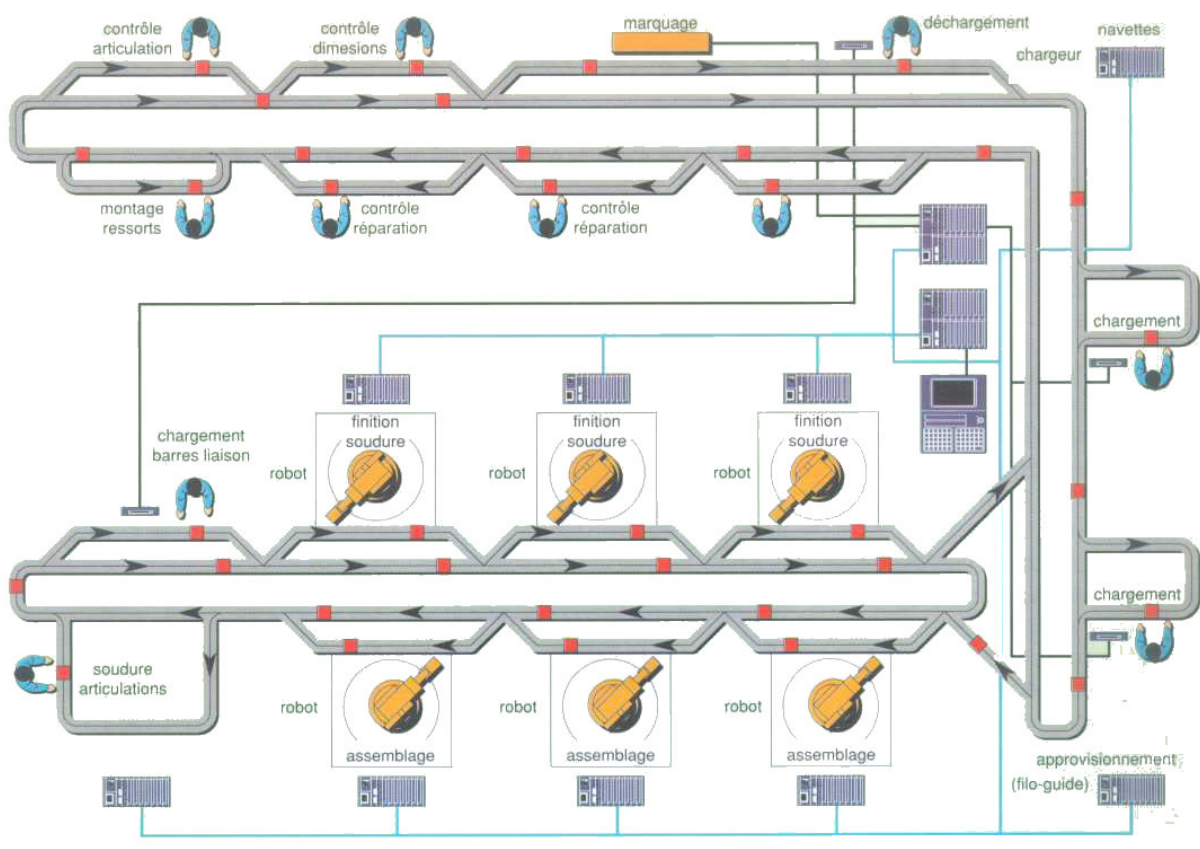


Figure 9 : Exemple d'architecture à une seule piste et postes en dérivation

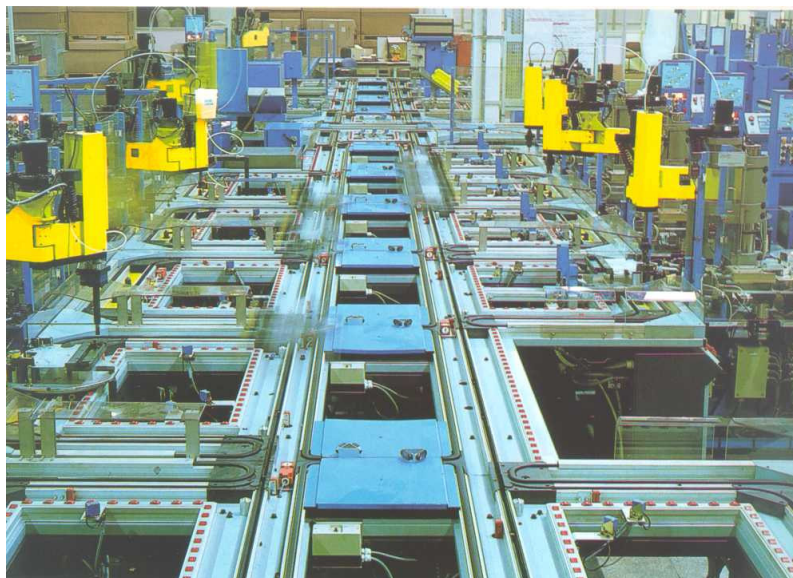


Figure 10 : Exemple d'architecture à deux pistes et postes en dérivation (doc. IBM)

En dehors des palettes et des convoyeurs, les autres éléments technologiques utilisés pour la fonction TRANSFERT sont :

- **Les butées** : elles permettent d'arrêter les palettes sur la ligne pour des opérations sans précision ou pour la gestion des priorités aux intersections.

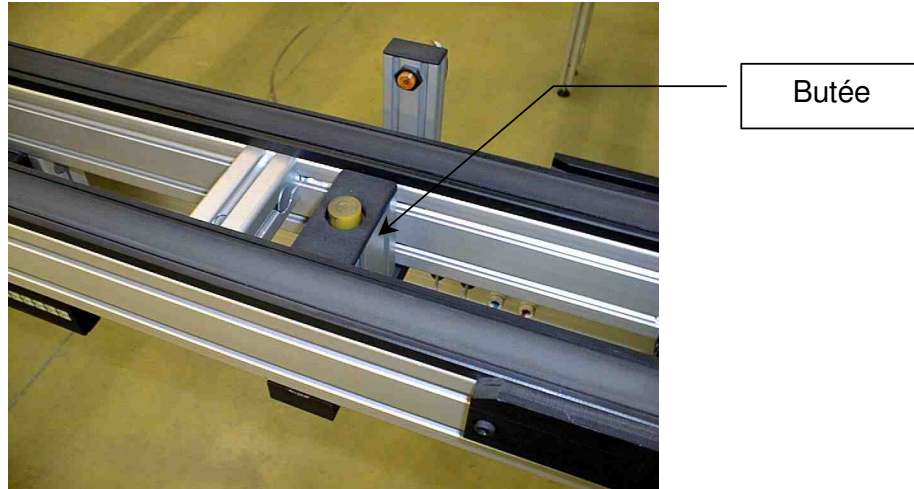


Figure 9 : Exemple de butée

- **Les indexages** : Ils sont destinés eux aussi à arrêter les palettes, mais contrairement aux butées qui laissent la palette en glissement sur le convoyeur, les indexages la soulèvent légèrement et la mettent en position précisément (environ +/- 0,03 mm) afin de pouvoir y réaliser des opérations délicates. Les indexages sont donc essentiellement situés au niveau du poste de travail, qu'il soit en dérivation ou en série.

Ces indexages sont revus dans la fonction ASSEMBLER (chap. II.3)

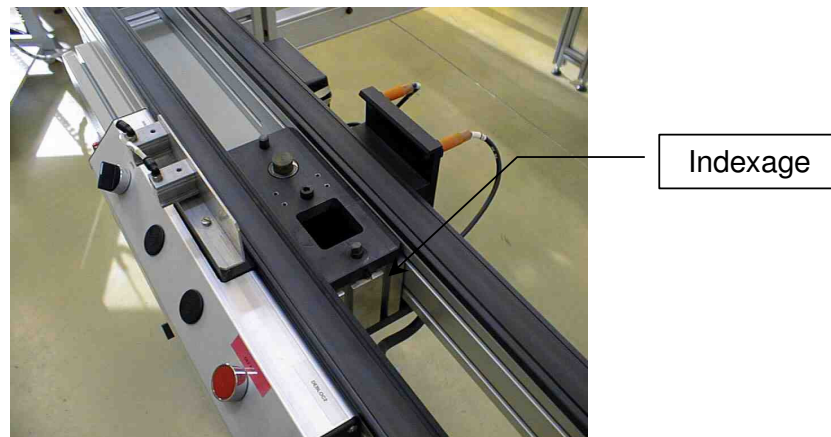


Figure 10 : Exemple d'indexage

- **Les aiguillages ou dérivations** : Ce sont les éléments technologiques indispensables si l'on désire concevoir une ligne avec des postes de travail en dérivation. Leur fonction est de dériver ou pas les palettes de la ligne principale sur une unité de transfert perpendiculaire. La dérivation se fait par lecture d'informations portées par la palette elle-même (voir partie 3 : "Caractéristiques du système expérimental").

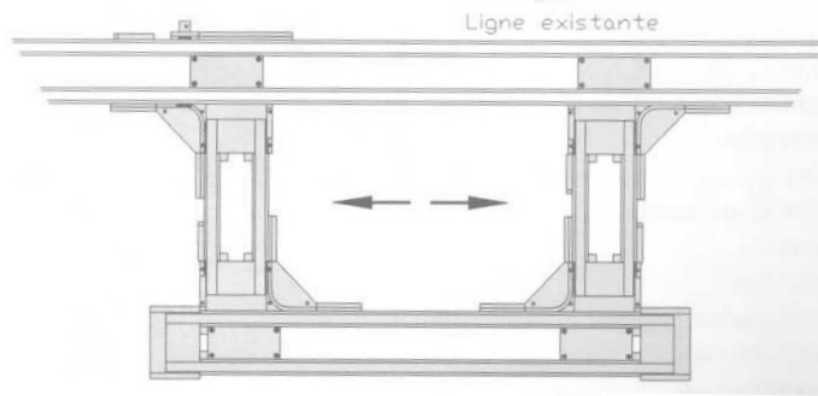


Figure 11 : Exemple de poste en dérivation (doc. ELCOM)

□ II.3- Fonction ASSEMBLER

Rappel : On se limite à l'assemblage par pièce maîtresse.

Sur un transfert libre, les postes d'assemblage peuvent être manuels ou automatisés et travailler à des cadences différentes, surtout s'ils sont placés en dérivation par rapport à la ligne.

La fonction **ASSEMBLER** se décompose en plusieurs sous fonctions dans le cas d'un assemblage par pièce maîtresse : **STOCKER, ALIMENTER, MANIPULER, INSERER.**

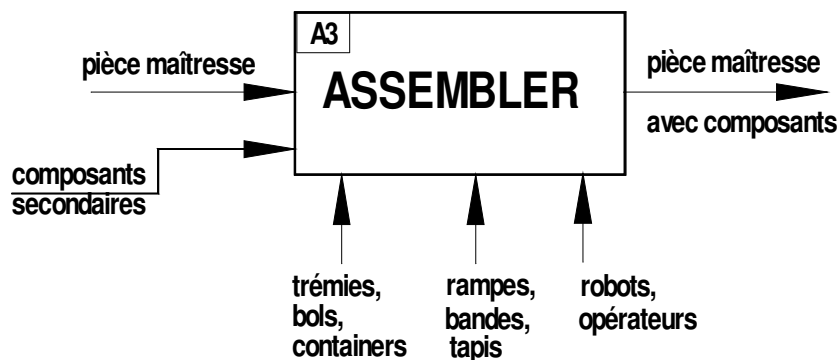


Figure 12 : Fonction ASSEMBLER

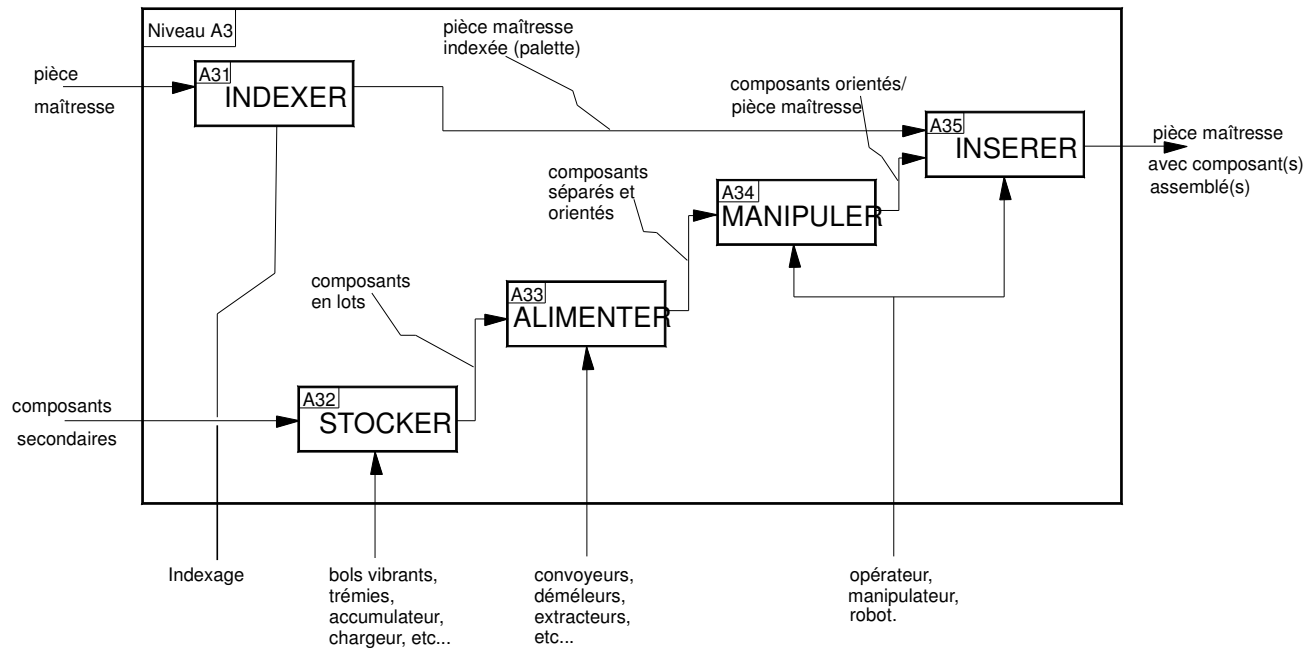


Figure 13 : Analyse détaillée de la fonction ASSEMBLER

II.3-a) Description de la sous fonction STOCKER

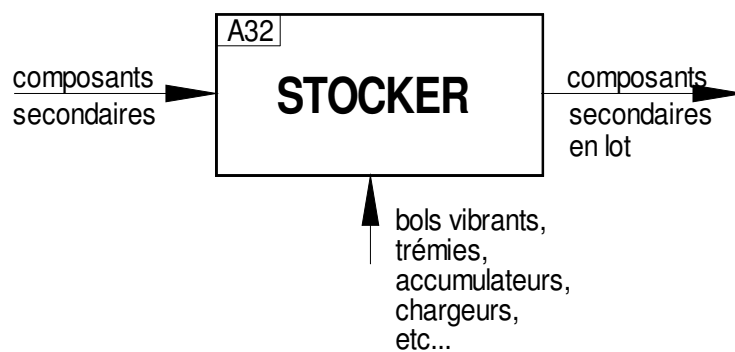


Figure 14 : Fonction STOCKER

Le stockage permet l'accumulation de composants autorisant une certaine autonomie du poste d'assemblage en le découplant du système amont qui peut être un magasin central, un atelier de fabrication ou un autre poste d'assemblage.

Le stockage des composants peut être répartie sur plusieurs niveaux liés:

ê *niveau 1* : stockage de composants pour la journée;

ê *niveau 2* : stockage de composants pour quelques minutes de fonctionnement.

*C'est par exemple le cas lorsque la fonction **STOCKER** est réalisée à partir d'une trémie (stockage niveau 1) et d'un bol vibrant (stockage niveau 2).*

La quantité de composants stockés sur un poste d'assemblage doit être calculée pour minimiser les manutentions tout en évitant les ruptures de stocks, ce qui provoquerait un arrêt de l'opération. La détermination de cette quantité peut être évaluée à partir de prévisions de commande.

En fonction de la forme géométrique des composants et de la nature du stockage, on rencontre différents types de dispositifs [OLIVIER 89] [RAMPERSAD 95] :

ê *stockage de composants isolés* : bandes ou plateaux alvéolés, utilisés surtout dans l'industrie électronique, pour le stockage de composants petits et fragiles;

ê *stockage de composants empilés, alignés ou en file (appelé aussi "stockage 1D")* : goulottes, rails, convoyeurs à bande ou vibrants avec dispositifs d'autoalignement des composants;

ê *stockage de composants en vrac planaire (2D) et volumique (3D)* : les dispositifs utilisés pour ce type de stockage sont les plus complexes et présentent des risques de défaillances qui sont à l'origine de 60% des dysfonctionnements d'un poste d'assemblage [CETIM 95]. Parmi ceux-ci, les plus fréquents sont les bols vibrants associés aux trémies, et les alimentateurs centrifuges, élévateurs ou rotatifs. La particularité de ces dispositifs est qu'ils sont, la plupart du temps, spécifiques à un composant. Ils présentent donc peu ou pas du tout de flexibilité, ce qui est donc contraire au type de production étudié.

La solution qui semble la plus flexible aujourd'hui est l'utilisation d'un robot associé à un système de vision artificielle. Il est à noter que cette solution concerne non seulement la fonction STOCKER mais aussi les autres fonctions du poste d'assemblage : ALIMENTER, MANIPULER, INSERER.

Une étude du coût d'un robot et d'un système de vision doit donc prendre en compte cet argument ainsi que sa grande flexibilité.

De nombreuses recherches ont concerné la reconnaissance de formes volumiques pour piloter un robot capable de prendre des composants directement dans un vrac. Il existe cependant encore peu d'applications industrielles.

Par contre, une variante consiste à faire passer les composants d'un vrac volumique à un vrac planaire par un système classique et universel (trémie et convoyeur à bande par exemple) puis à effectuer une reconnaissance de forme en 2D sur les composants à plat.

Le robot peut alors saisir le composant correctement disposé en orientant son préhenseur grâce au guidage de la caméra.

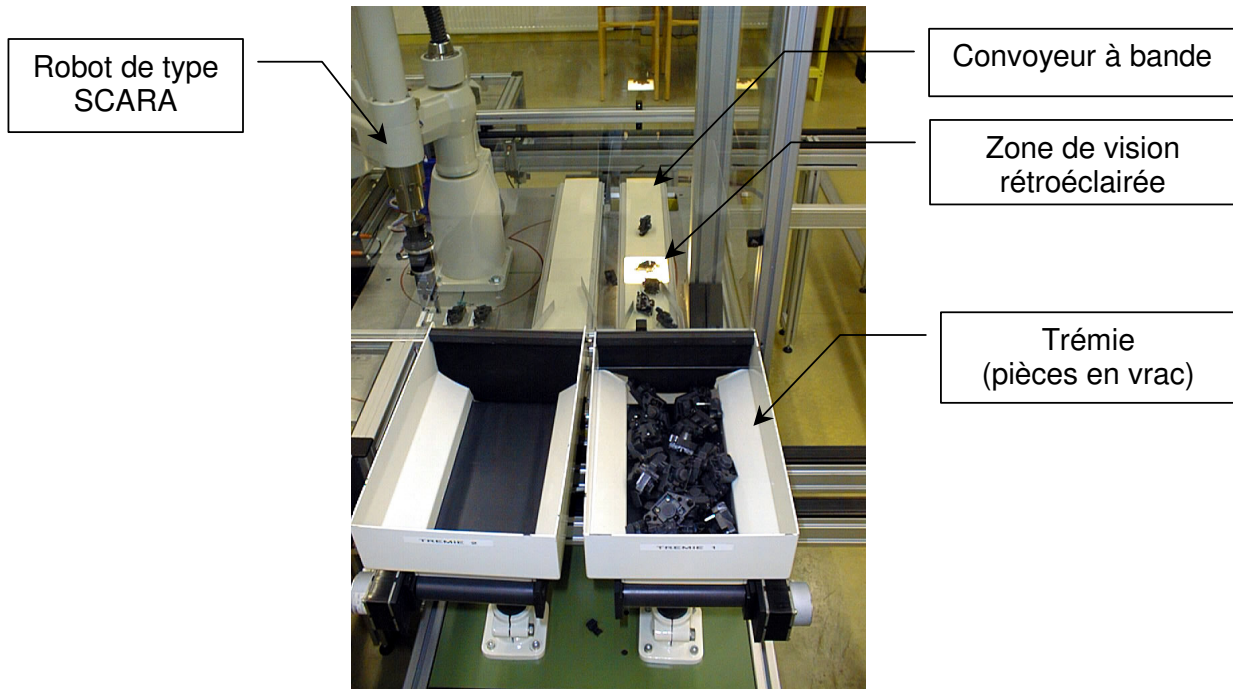


Figure 15 : Exemple de distributeur flexible assisté par vision

II.3-b) Description de la sous fonction ALIMENTER

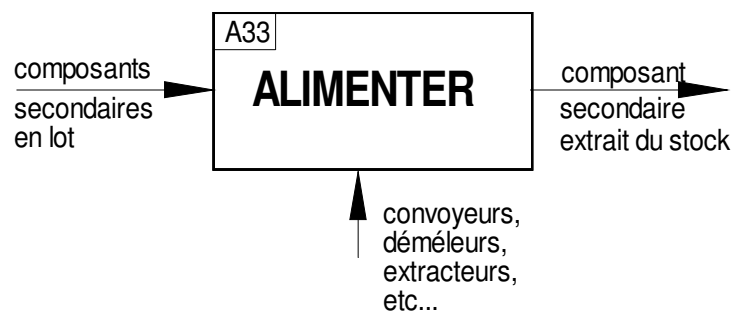


Figure 16 : Fonction ALIMENTER

La réalisation pratique de la fonction ALIMENTER est souvent associée à la fonction STOCKER. Cette fonction permet d'extraire un composant du stock et de le présenter dans une orientation permettant ensuite sa manipulation. En sortie de stock, on trouve

des dispositifs permettant de piéger les composants mal orientés et de les renvoyer dans le stocks.

Lorsque la fonction STOCKER est réalisée à partir d'un système vibrant, la sortie des composants n'est pas continue. La fonction ALIMENTER doit dans ce cas permettre aussi d'accumuler quelques composants dans la bonne orientation afin de ne pas freiner la cadence du poste d'assemblage. On utilise alors de petits convoyeurs à bande ou tout simplement des rampes gravitaires.

II.3-c) Description de la sous fonction MANIPULER

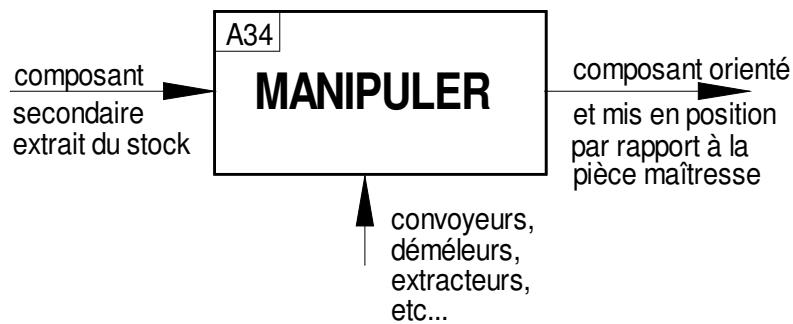


Figure 17 : Fonction MANIPULER

Deux cas peuvent se présenter pour la réalisation de cette fonction :

ê *manipulation manuelle* : c'est un opérateur qui réalise l'opération. La conjoncture économique et sociale peut avoir un effet non négligeable sur ce choix. Ceci n'est pas de mon ressort mais pourrait inspirer peut-être d'autres recherches dans le domaine des sciences sociales

Sur un poste d'assemblage, l'opérateur est utilisé soit pour réaliser des opérations à la fois simples et variées, pour lesquelles une automatisation serait trop coûteuse, soit pour des opérations d'une grande complexité où l'intelligence et la dextérité humaine sont irremplaçables.

ê *manipulation automatisée* : Les solutions technologiques pour automatiser cette fonction sont nombreuses, mais là encore, seuls les robots permettent d'obtenir la flexibilité nécessaire à une production multiproduit. Ils font l'objet de recherche permanente pour améliorer leurs performances et leur environnement. Aujourd'hui, plusieurs constructeurs annoncent des répétabilités de positionnement de +/- 5 microns. Parallèlement, les coûts ont bien diminué ces 5 dernières années.

Les architectures de type portique et Scara (3 à 4 axes), actuellement les plus répandues, sont utilisées pour des cycles où le composant manipulé reste dans une orientation parallèle au plan de prise et de dépose.

Les robots polyarticulés verticaux (5 à 6 axes) permettent des trajectoires plus complexes pour la réalisation de tâches d'assemblage selon des orientations différentes.

On peut rappeler que dans le cas où le transfert des produits est réalisé par une ligne comportant des postes de travail en dérivation, la cohabitation de postes d'assemblage manuels et automatisés ne pose aucun problème. En effet, des zones de stockage en amont des postes permettent aux postes manuels de ne pas être astreints à suivre la cadence des postes automatisés. Les temps de cycle de chacun des postes doit toutefois rester proches les uns des autres afin d'éviter une saturation en amont du poste le plus lent et une réduction du taux d'engagement du poste le plus rapide.

□ II.4- Fonction CHARGER/DECHARGER

Rappel : On se limite à l'assemblage par pièce maîtresse.

Comme on l'a vu dans les fonctions **TRANSFERER** et **ASSEMBLER**, le principe de l'assemblage par pièce maîtresse nécessite de la positionner sur la palette dès le début du processus. De même, lorsque toutes les opérations de la gamme sont terminées, le produit assemblé doit être retiré de la palette. La fonction **CHARGER/DECHARGER** est le plus souvent réalisée sur un poste de travail similaire aux postes d'assemblage, c'est à dire disposé en dérivation par rapport à la ligne principale. La seule particularité du poste de chargement/déchargement est sa zone de stockage qui doit être importante pour pouvoir accumuler plusieurs palettes en attente de lancement.

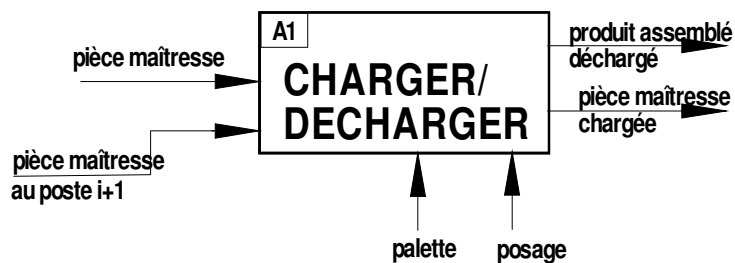


Figure 18 : Fonction CHARGER/DECHARGER

Lors des opérations d'assemblage, principalement celles qui sont automatisées, la position de la pièce maîtresse sur la palette doit être suffisamment précise et présenter une bonne répétabilité. On doit donc utiliser des posages, réalisés spécialement en

fonction de la forme et des dimensions de la pièce maîtresse, dont le rôle est d'en assurer la mise en position précise sur la palette pour que les opérations d'assemblage les plus délicates puissent s'effectuer sans aléa. Si nécessaire, des dispositifs de maintien peuvent équiper le posage pour éviter le déplacement de la pièce maîtresse ou des composants.

III- Aspects dynamiques du flux de palettes

Les transferts asynchrones permettent de déplacer les palettes d'un point à un autre tout en assurant leur accumulation momentanée. On peut donc les assimiler à des stocks, ce qui pose le problème de leur capacité d'accueil et de leurs règles de gestion.

En ce qui concerne la capacité d'accueil, celle-ci est finie, puisqu'elle correspond à la quantité de palettes pouvant être accumulées bout à bout sur la longueur totale de convoyeur. Cette situation de saturation est à éviter car elle correspondrait à un fonctionnement en mode synchrone et serait contraire à l'un des objectifs du **Juste à Temps** qui est de réduire les stocks. La charge des stocks/convoyeurs doit donc être inférieure à leur capacité. Elle doit seulement permettre de découpler les postes de travail, surtout si leurs cadences ne sont pas identiques, et compenser, sur une durée limitée, les pannes et arrêts pouvant survenir sur un des postes.

D'autre part, d'un point de vue économique, l'objectif est de minimiser le nombre de produits en-cours de processus (appelé **Work In Progress- WIP-** par les anglo-saxons) pour limiter le capital immobilisé mais aussi pour limiter l'investissement en réduisant le nombre de palettes achetées.

Il se pose donc le problème du dimensionnement des stocks/convoyeurs et du nombre total de palettes en circulation sur le transfert. **Ceci est repris dans la partie 6 : "Analyse des résultats des simulations"**.

Pour ce qui est des règles de gestion, les stocks peuvent être classés selon deux catégories [DESCHAMPS 94] :

– **Les stocks "STATIQUES"** (figure 21) sont des zones de stockage caractérisées par l'unicité du point d'entrée et de sortie, ainsi que par l'invariabilité de l'ordre des palettes dans ces stocks. Elles sont donc utilisées en FIFO par le poste desservi par le convoyeur.

Les tronçons qui composent les postes en dérivation peuvent être assimilés à des stocks statiques. Leur situation géographique par rapport au poste de travail proprement dit,

permet de les décomposer en "**stock AMONT**" et "**stock AVAL**" (voir figure 21). Leur capacité est limitée physiquement par la dimension du convoyeur, mais peut aussi être limitée logiquement par le système de pilotage. Ces données correspondent à des attributs associés au fonctionnement d'un poste en dérivation.

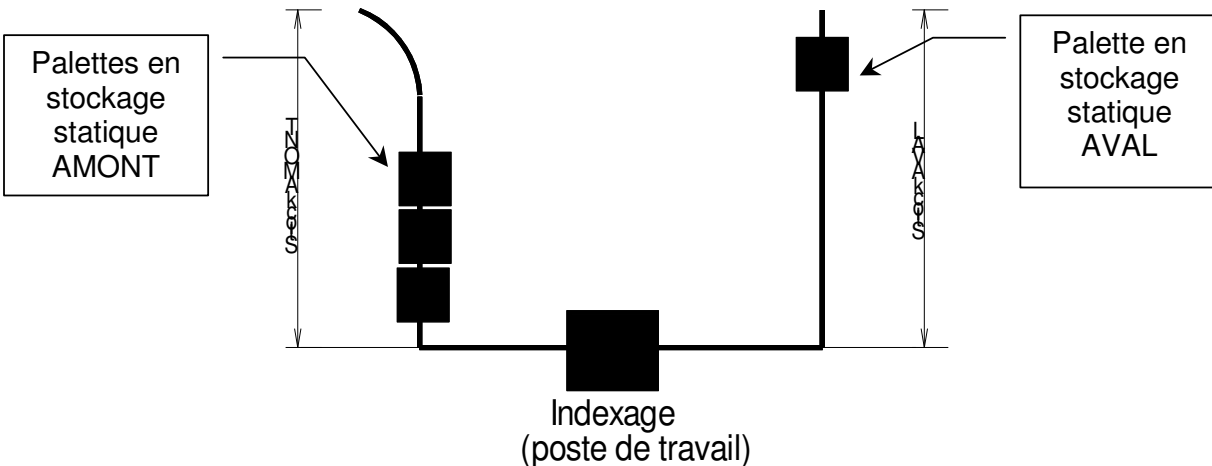


Figure 19 : Stocks/convoyeurs statiques d'un poste en dérivation

- **Les stocks "DYNAMIQUES"** (figure 22) sont des zones de stockage qui, par opposition à la notion précédente, se caractérisent par la multiplicité des points d'entrée et de sortie et la variabilité de l'ordre des palettes.

Les tronçons qui composent la ligne principale qui dessert les différents postes en dérivation sont assimilés à un stock dynamique assurant les fonctions de transfert des composants d'un poste à un autre et de stockage momentané. L'état de ce stock évolue au cours du temps puisque l'ordre absolu et la quantité de palettes sont modifiés à chaque entrée ou sortie.

La gestion de ce stock dynamique présente donc une complexité qui ne permet pas d'utiliser une simple règle de priorité .

Dans la suite, ce problème sera repris en détail afin d'essayer de dégager des principes, des méthodes, des règles ou des algorithmes permettant de maîtriser le fonctionnement du stock dynamique et son interaction avec le stock statique.

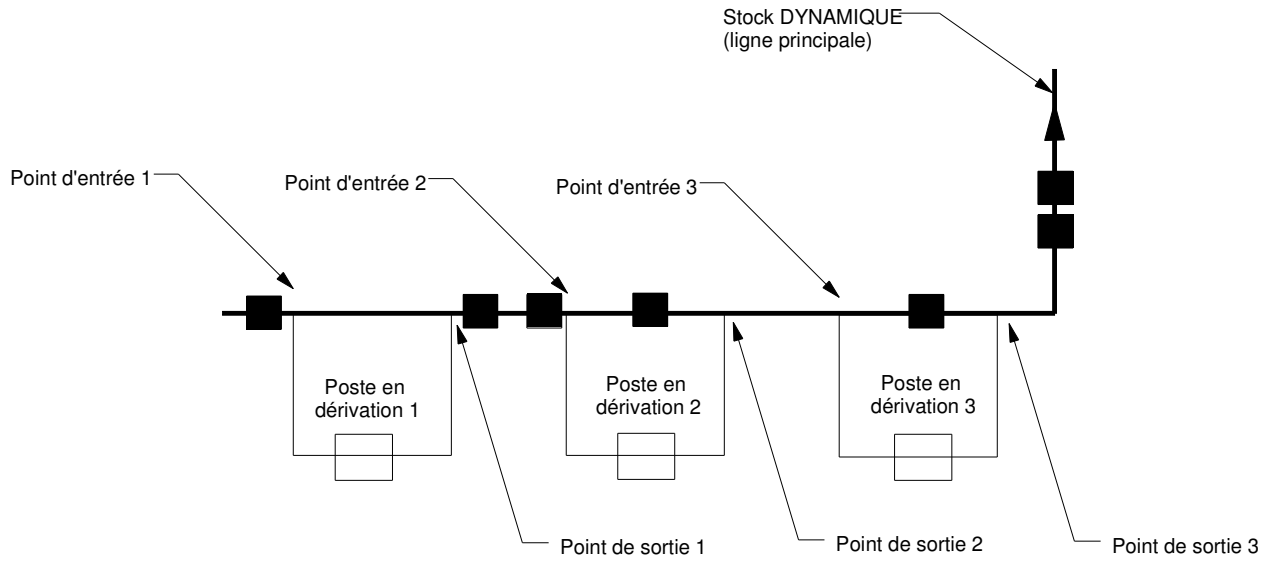


Figure 20 : Stocks/convoyeurs dynamiques de la ligne principale

Conclusion

Dans cette partie, on a étudié dans le détail les lignes d'assemblage à transfert asynchrone automatisé.

Pour ce type de ligne qui sert de support à cette recherche, une analyse fonctionnelle est proposée à partir du formalisme SADT. Cette analyse permet de faire ressortir non seulement les fonctions, mais aussi les données traitées par ces fonctions. On met aussi en évidence les échanges de données entre chaque fonction. Ce type d'analyse est utile à la compréhension de la structure physique et du fonctionnement d'une ligne d'assemblage, qui comme on l'a vu, est un système particulièrement complexe car composé de plusieurs équipements intégrés les uns avec les autres. Le concepteur d'une telle ligne trouvera aussi dans cette analyse un outil lui permettant de découper son travail à partir de ces fonctions afin de simplifier la conception de la partie opérative comme de la partie commande.

Le dernier chapitre sur l'aspect dynamique du flux de palettes permet de faire ressortir des particularités importantes et propres à ce type de ligne : le nombre fini de palettes que la ligne peut recevoir, l'utilisation de certains convoyeurs pour accumuler des palettes au même titre qu'un stock accumule des pièces devant une machine. On a identifié deux types de stock/convoyeur : les stocks statiques pour lesquels l'ordre de sortie est le même que l'ordre d'entrée (type FIFO), et les stocks dynamiques qui autorisent des permutations entre les palettes parce qu'ils possèdent plusieurs points d'entrée ou sortie.

La liaison entre ces deux types de stocks sur une ligne pose le problème de la gestion des priorités entre les palettes qui sortent d'un poste en dérivation et celles qui circulent sur le convoyeur central. Ce problème sera repris en détail dans la partie 3 " Présentation du système expérimental".

Bibliographie de la partie 2 : "Processus d'assemblage"

[LITTLE 94] : "Automated assembly scheduling : a review"

Computer Integrated Manufacturing Systems, Vol. 7 N°1 02/1994, pages 51 à 61

D. Little, A. Hemmings

[CETIM 95] : "Equipements d'assemblage"

Cetim information N°146, 1995

[DESCHAMPS 94] : "Gestion hiérarchisée de cellules flexibles d'assemblage"

Thèse de doctorat N°94526, Université de Toulouse, 1994

J.C. Deschamps

[OLIVIER 89] : "La périrobotique d'assemblage"

Editions Hermés, 1989

P. Olivier

[YING-CHIN 98] : "Machine layout with a linear single-row flow path in an automated manufacturing system"

Journal of Manufacturing Systems, Vol. 17 N°1, 1998, pages 1 à 22

[RAMPERSAD 95] : "A case study in the design of flexible assembly systems"

International Journal of Flexible Manufacturing Systems, Vol. 7 1995, pages 255 à 286

H.K. Rampersad

- PARTIE 3 -
***Présentation du système
expérimental***

Sommaire

Introduction	73
I- Architecture et décomposition fonctionnelle du système expérimental	74
II- Analyse du système informationnel.....	77
- II.1- Typologie du séquençement des opérations sur une ligne transfert	77
- II.2- Données de pilotage d'une ligne transfert avec postes en dérivation.....	79
<i>II.2-a) Paramètres intrinsèques à la palette (produit)</i>	<i>80</i>
<i>II.2-b) Paramètres intrinsèques au poste en dérivation</i>	<i>82</i>
- II.3- Organisation de la structure de commande.....	82
Conclusion	85

Introduction

Afin de mettre en évidence les problèmes liés aux processus d'assemblage multiproduit sur ligne transfert asynchrone, et notamment les problèmes de gestion de flux physique et informationnel, la suite de ce travail de recherche s'est appuyée sur un système de production réel.

En effet, compte tenu de la complexité des systèmes de production et de leur comportement stochastique, l'utilisation ou le développement de modèles théoriques pour les étudier s'avère incertaine si on ne les a pas validé par la pratique.

Par ailleurs, une petite enquête auprès d'intégrateurs ou d'utilisateurs de lignes transfert asynchrone fait ressortir qu'il existe effectivement un certain nombre de problèmes qui sont bien souvent contournés au détriment du rendement opérationnel de ces systèmes.

Les objectifs des expérimentations sur un système réel sont donc d'observer son comportement dynamique, d'identifier les problèmes et de collecter des données pouvant servir à l'élaboration et la validation d'un modèle.

Le système servant de support à ces expérimentations est la **Ligne Flexible Expérimentale** du département Organisation et Gestion de Production de l'IUT de Cergy-Pontoise (site d'Argenteuil).

Dans le chapitre 1, on détaille la structure physique de cette ligne en 3 modules fonctionnels qui permettent d'assurer les fonctions *CHARGER/DECHARGER*, *TRANSFERER* et *ASSEMBLER* vues dans la partie 2.

Le chapitre 2 décrit la structure informationnelle. On présente tout d'abord une typologie des différents séquencements d'opérations qu'il est possible d'envisager sur une ligne transfert synchrone ou asynchrone. Ce sont bien sûr les séquencements sur lignes à transfert asynchrone qui sont surtout détaillés. Pour une ligne comportant des postes de travail en dérivation telle que celle étudiée, on présente ensuite les paramètres utiles à la gestion de la dérivation des palettes/produits et l'algorithme d'accès sur un poste. Cela met en évidence l'intérêt d'un dispositif de codage des palettes pour assurer à la fois leur routage et leur traçabilité. On termine ce chapitre par la description de la structure de commande d'une telle ligne, qui est à la fois distribuée et hiérarchisée.

I- Architecture et décomposition fonctionnelle du système expérimental

Le composant principal de la ligne flexible est un transfert libre modulaire automatisé, dont la structure fait apparaître 3 types de modules fonctionnels (figure 1) :

- Module 1 : **Ligne de chargement / déchargement (x1)**;
- Module 2 : **Ligne principale (x1)**;
- Module 3 : **Poste en dérivation (x6)**;

- **La ligne de chargement/déchargement** (figure 1) remplit la fonction **CHARGER/DECHARGER**. Elle est constituée de 4 tronçons linéaires fermés servant de stockage aux palettes en attente de lancement ou de livraison. Un tronçon ouvert à une de ces extrémités sert de stockage aux palettes, ou plus précisément aux produits, présentant un défaut. Ce tronçon permet d'isoler ces palettes afin de ne pas les réintroduire dans le système sans prendre en compte leur défaut.

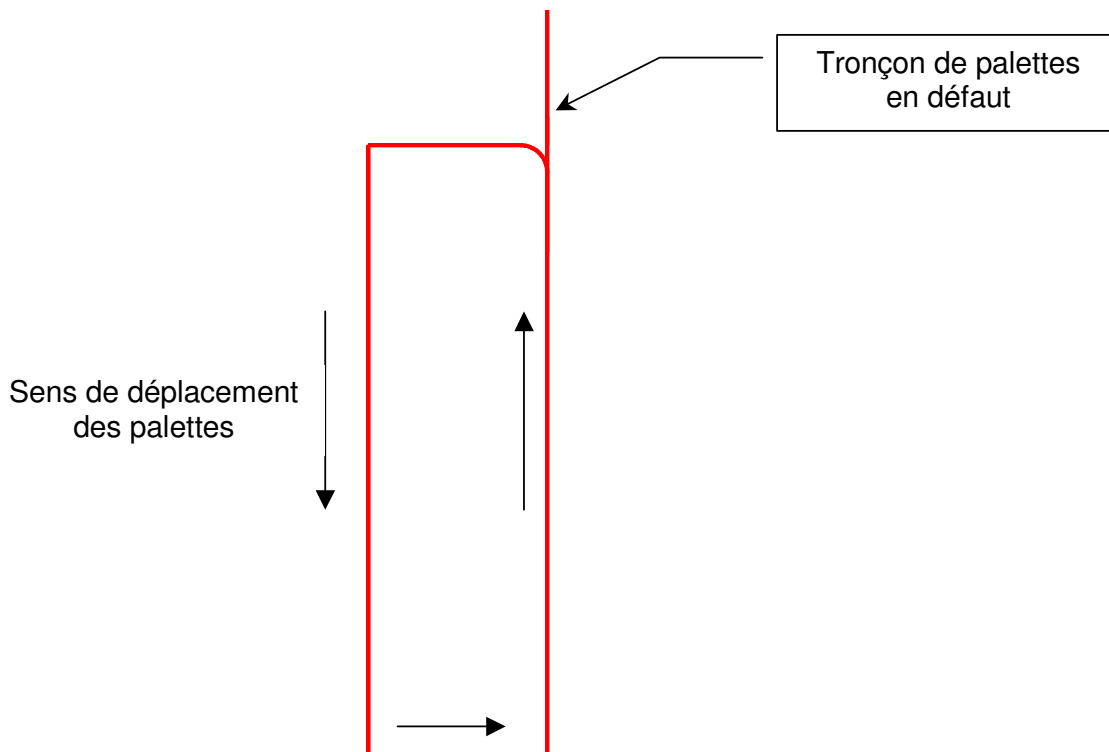


Figure 1 : Ligne de chargement/déchargement

- **La ligne principale** (figure 2) remplit la fonction **TRANSFERER**. Elle sert de stock dynamique et permet de distribuer les palettes sur les différents postes en dérivation. Elle est composée de 3 tronçons de distribution et d'un tronçon qui autorise le rebouclage éventuel des palettes sans passer par la ligne de chargement/déchargement.

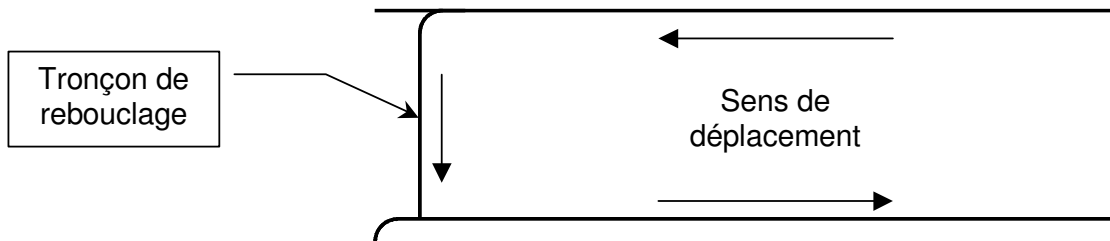


Figure 2 : Ligne Principale

- **Les postes en dérivation** (figure 3), au nombre de 6, assurent la fonction **ASSEMBLER**. Ils comportent un tronçon servant de stock statique amont (capacité physique = 6 palettes), un tronçon servant de stock statique aval (capacité physique = 6 palettes) et un tronçon sur lequel est implanté un dispositif d'indexage permettant de réaliser une opération d'assemblage.

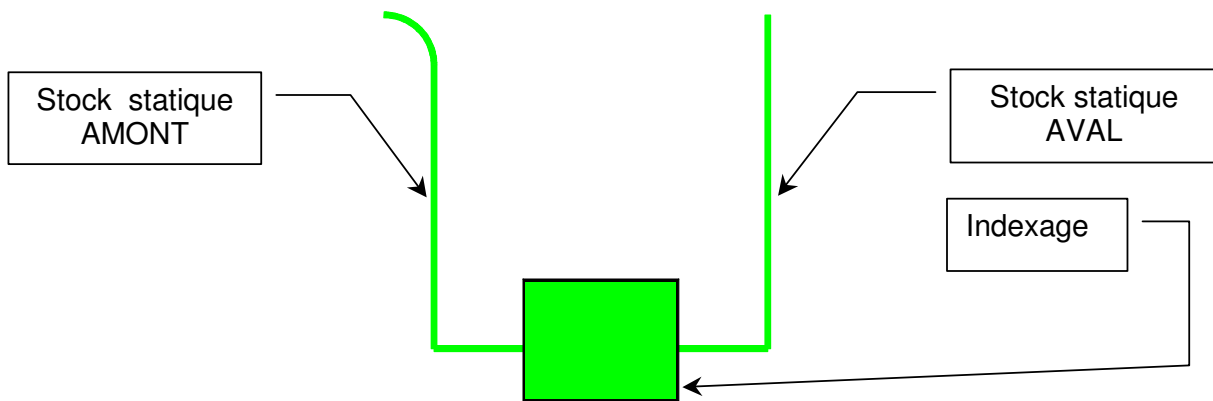


Figure 3 : Poste en dérivation

La liaison des différents modules permet d'obtenir l'architecture représentée sur la figure 4.

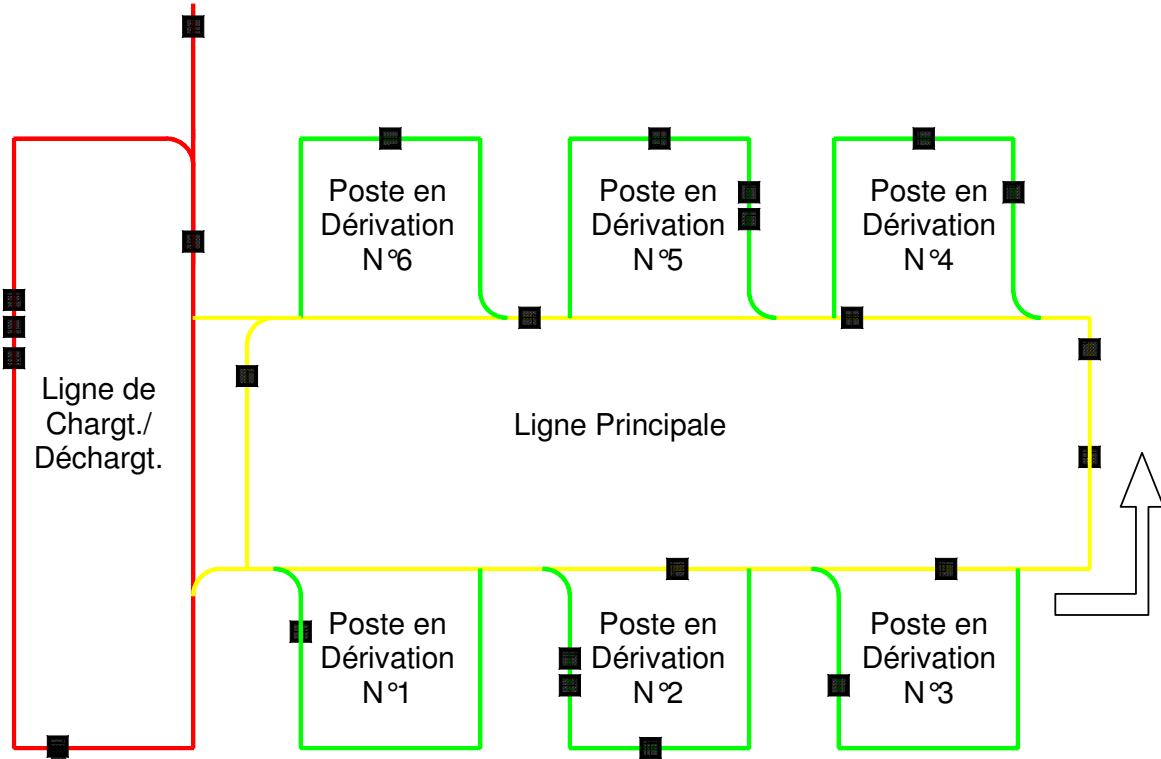


Figure 4 : Structure générale du système physique

Au maximum, 25 palettes sont disponibles pour effectuer la production. Si la quantité à produire est supérieure à 25, alors il faudra attendre le déchargement des palettes lancées dans la première série. Il n'est toutefois pas utile d'attendre le déchargement de la totalité des 25 palettes de la première série, le lancement des palettes de la série suivante pouvant s'effectuer au fur et à mesure du déchargement des palettes de la série précédente.

II- Analyse du système informationnel

□ II.1- Typologie du séquençement des opérations sur une ligne transfert

Plusieurs types de séquençements des opérations de production peuvent être réalisés sur une ligne transfert :

II.1-a) Séquençement de type " Transfert lié / Gamme liée" : ce type de séquençement correspond aux produits réalisés sur transfert lié (voir partie 1 chap. V.1). La gamme de production comporte des opérations non permutable, dont l'ordre ne peut pas être changé. Ce type de séquençement est caractéristique des lignes dédiées à un seul produit et ne fait pas l'objet de cette recherche.

Les autres types de séquençements présentés par la suite sont réalisables uniquement sur des transferts libres (asynchrones) comportant des postes en dérivation et permettant d'assembler plusieurs types de produits différents (mixed model assembly lines). Ils sont donc applicables au système expérimental.

II.1-b) Séquençement de type "Transfert libre / Gamme liée / Phases non redondantes" : Dans ce type de séquençement, les opérations se déroulent dans un ordre imposé et de façon non redondante. On peut distinguer dans ce type :

- les gammes où les opérations se déroulent dans l'ordre croissant de l'implantation des postes de travail sur le transfert. Dans ce type, l'ensemble des opérations peut s'effectuer sans faire reboucler les palettes sur la ligne principale;
- les gammes où les opérations se déroulent dans un ordre indépendant de l'implantation des postes de travail, mais avec la nécessité de faire reboucler les palettes sur la ligne principale pour réaliser toutes les opérations de la gamme.

Exemple :

a) On considère une ligne transfert identique à celle servant de support expérimental (6 postes en dérivation). Trois types de produits différents sont fabriqués, avec donc trois gammes différentes :

Gamme A : P1, P2, P3, P4, P5, P6

Gamme B : P1, P2, P3, P6

Gamme C : P3, P4, P6

Ces trois gammes ont un séquençement d'opérations qui se réalise dans l'ordre croissant d'implantation des postes en dérivation. Il n'y a pas de rebouclage ni de redondance des opérations. La gamme A correspond à un produit devant passer sur tous les postes en dérivation, les gammes B et C correspondent à des produits dont la réalisation ne nécessite pas le passage sur tous les postes.

b) Autre cas :

Gamme D : P3, P1, P2

Gamme E : P5, P6, P1

Gamme F : P2, P3, P6, P1

Ces trois gammes ont un séquençement d'opérations qui se réalise dans un ordre indépendant de l'implantation des postes en dérivation, sans redondance des opérations, mais avec la nécessité de reboucler sur la ligne principale.

II.1-c) Séquençement de type "Transfert libre / Gamme liée / Phases redondantes" :

Type identique au précédent, sauf que les produits passent sur certains postes plus d'une fois.

Exemple :

Gamme G : P1, P3, P4, P1, P5

Gamme H : P3, P4, P3, P4, P5, P6

Gamme I : P2, P5, P6, P1, P6

II.1-d) Séquençement de type " Transfert libre / Gamme libre / Phases non

redondantes" : Dans ce type, les opérations se déroulent dans un ordre non imposé et de façon non redondante. En fonction de la disponibilité des postes, l'ensemble des opérations pourra être effectué avec ou sans rebouclage des palettes sur la ligne principale.

Exemple :

Un produit doit passer sur les postes 1, 3 et 6. Toutes les combinaisons de séquençement de ces 3 opérations sont possibles : (P1, P3, P6), (P1, P6, P3), (PP6, P1, P3), etc....

Par la suite, on utilisera essentiellement des séquençements d'opérations correspondant à des gammes d'assemblage. Dans ce cas, l'ordre des opérations est la plupart du temps imposé.

C'est donc essentiellement le séquençement de type "**Transfert Libre / gamme liée / phases non redondantes**" avec opérations dans l'ordre croissant des postes qui sera exploité en expérimentation, car c'est le plus représentatif des processus d'assemblage par pièce maîtresse.

□ II.2- Données de pilotage d'une ligne transfert avec postes en dérivation

Les paramètres pris en compte pour la dérivation d'une palette sur un poste en dérivation sont donc :

- le routage de la palette : la palette doit-elle passer sur le poste considéré ?
- l'état du poste : disponible ou en arrêt ;
- la charge du stock statique amont : inférieure ou égale à sa capacité logique ;
- la réalisation des opérations antérieures ;
- l'absence de défaut sur le produit.

L'algorithme d'accès d'une palette sur un poste en dérivation prend en compte ces paramètres (figure 5).

Ces paramètres peuvent être classés en 2 catégories :

- les paramètres intrinsèques à la palette (produit) : le routage, la réalisation des opérations antérieures, l'absence de défaut, etc...
- les paramètres intrinsèques au poste en dérivation : la disponibilité, la charge du stock statique, ...

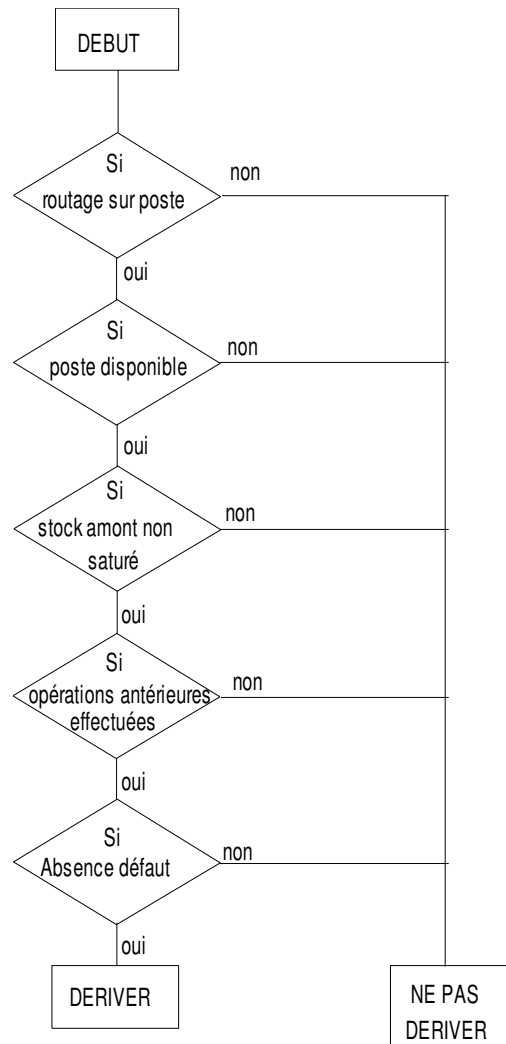


Figure 5 : Algorithme d'accès

II.2-a) Paramètres intrinsèques à la palette (produit)

Le principe du transfert libre et le séquençement de type "Transfert libre/ Gamme liée / Phases non redondantes" impliquent une indépendance totale de chaque palette. Il est donc nécessaire de disposer d'un système capable d'identifier chaque palette en fonction d'un code correspondant aux différents paramètres vus précédemment. Ces paramètres étant essentiellement binaires, le code peut être réalisé par un ensemble de bits pouvant prendre deux valeurs (0 et 1).

Pour le type de séquençement choisi, le code associé au routage doit comporter un nombre de bits égal au nombre de postes en dérivation : à chaque bit correspond un poste en dérivation. Par exemple, si une palette se présente à l'entrée d'une dérivation et

si le bit associé à ce poste est à **1**, alors la palette peut être dérivée. Dans ce cas, ce bit de routage doit être mis à **0**, de telle sorte que si la palette repasse ultérieurement devant ce même poste en dérivation, elle ne soit pas de nouveau dérivée (non redondance des opérations). Le code de routage doit donc être évolutif tout au long du processus de la palette. En fin de processus, tous les bits de ce code sont à **0**.

Le code associé à la réalisation des opérations antérieures ne comporte qu'un seul bit. Par exemple, si toutes les opérations antérieures par rapport à l'opération courante ont été réalisées, ce bit est à **1**. Si une des opérations précédentes n'a pu être réalisée, ce bit est à **0**. Dans ce cas, la palette effectue un rebouclage sur la ligne principale afin de se représenter devant les postes en dérivation où les opérations n'ont pas été réalisées. Ce rebouclage doit donc, par la même occasion, remettre le bit d'antériorité à **1**. Comme pour le code de routage, ce code doit être évolutif.

Si un défaut survient sur un produit lors d'une opération d'assemblage, les opérations postérieures ne doivent pas être réalisées, ce qui signifie que la palette ne doit plus être dérivée et doit retourner sur la ligne de déchargement où elle sera dérivée sur le tronçon de défaut (*voir chap. I*). Le code associé à ce paramètre ne comporte donc qu'un seul bit. Si par exemple ce bit de défaut est à **1**, la palette peut être dérivée.

En plus de ces 3 paramètres nécessaires au pilotage du flux physique, il peut être intéressant de compléter le code de base par d'autres informations utiles au suivi de la production, appelé encore "*traçabilité*":

- un identificateur de palette : numéro ou série de caractères alphanumériques;
- l'historique de la palette : ordre et dates de passage sur les différents postes;
- diverses données numériques permettant de vérifier la qualité du produit : mesures dimensionnelles, de position, de forces, etc....

La principale caractéristique du système retenu pour réaliser le codage et l'identification des palettes est donc la capacité d'informations traitées, chiffrée en bits ou en octets.

Actuellement, les dispositifs de type "*Étiquettes électroniques avec lecture/écriture des informations par induction électromagnétique*" correspondent à la principale solution industrielle pour le codage évolutif et l'identification des produits. Ils comportent des plots de lecture/écriture disposés aux points où il est nécessaire d'effectuer l'une ou l'autre de ces opérations, par exemple en entrée de poste en dérivation ou au niveau de l'indexage.

Les étiquettes électroniques sont fixées sur les palettes et sont en fait des mémoires dynamiques dont l'alimentation est réalisée par induction au passage devant un plot de lecture/écriture. Leur capacité varie de 64 à 256 Koctets.

II.2-b) Paramètres intrinsèques au poste en dérivation

Il y a essentiellement deux paramètres nécessaires au pilotage du flux physique et intrinsèques au poste en dérivation :

- *l'état du poste* : Si le poste est disponible, la palette peut être dérivée. En cas d'indisponibilité (arrêt ou panne), la palette ne peut pas être dérivée et le bit de réalisation des opérations antérieures doit changer d'état.

- *la charge du stock statique amont* : pour qu'une palette soit dérivée, la charge du stock statique amont doit être inférieure ou égale à sa capacité logique. Cette capacité logique, inférieure ou égale à la capacité physique est fixée par le niveau informationnel supérieur (supervision, voir chap. II.3). En cas de dépassement de la capacité logique, plus aucune palette ne peut être dérivée. Le bit de réalisation des opérations antérieures doit changer d'état.

Comme on le voit, ces paramètres intrinsèques au poste en dérivation modifient l'état d'un des paramètres intrinsèques à la palette (bit d'antériorité).

□ II.3- Organisation de la structure de commande

Plusieurs organisations de la structure de commande sont possibles :

- *structure centralisée,*
- *structure distribuée,*
- *structure distribuée et hiérarchisée* qui correspond à celle du système expérimental. Cette structure présente un agencement pyramidal sur plusieurs niveaux de l'autorité, ce qui a pour effet de limiter le nombre d'informations traitées par chaque niveau au strict nécessaire. Les échanges d'informations peuvent se faire entre des systèmes de commande de même niveau ou entre des niveaux différents. Une des caractéristiques fondamentales qui permet de distinguer les niveaux est le temps de réponse et la quantité d'informations accessibles.

Les figures 6 et 7 mettent en évidence les différents niveaux de la structure de commande et leur réalisation pratique par réseaux de communication. Dans la suite, nous nous limiterons aux niveaux 0 à 2 et à l'échange de données avec le niveau 3.

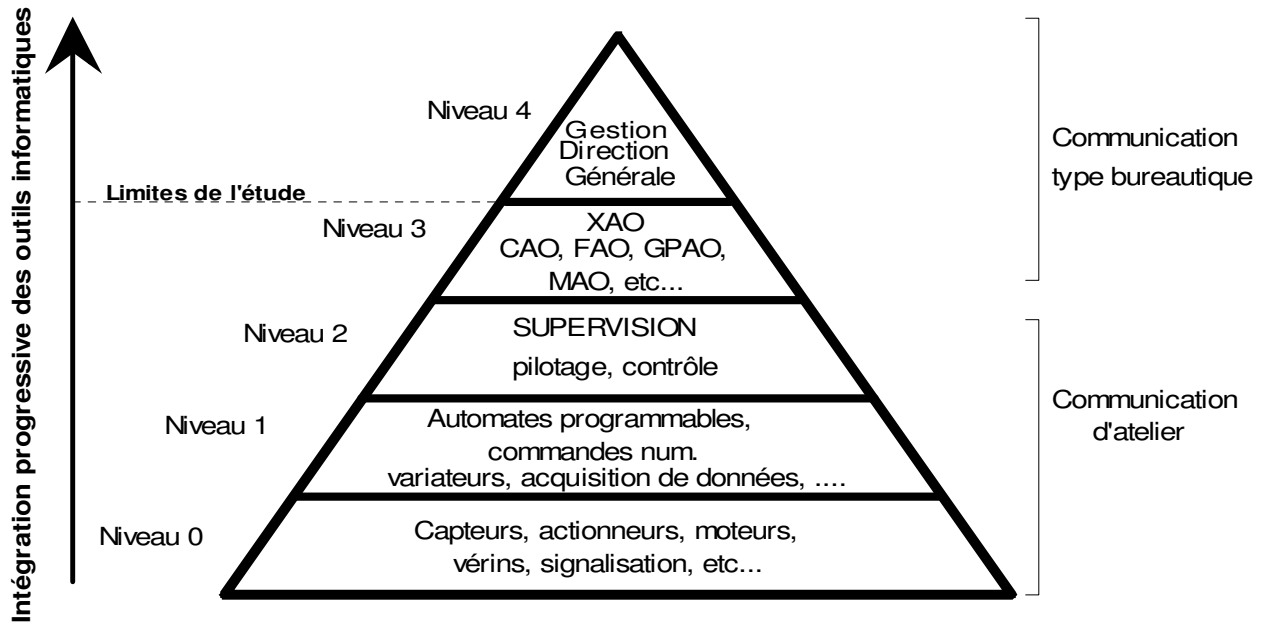


Figure 6 : Structure pyramidale et niveaux hiérarchiques de la commande

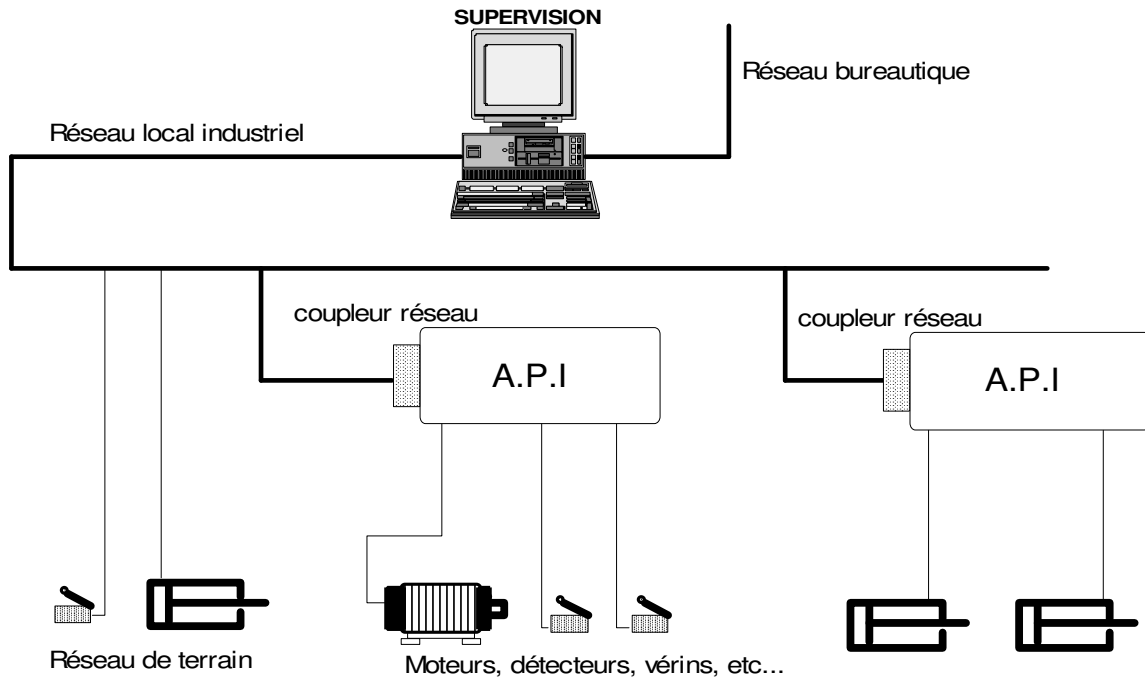


Figure 7 : Réalisation pratique et réseau local

En respectant la structure pyramidale, le système de commande de la ligne flexible expérimentale est implanté de la manière suivante (figure 8) :

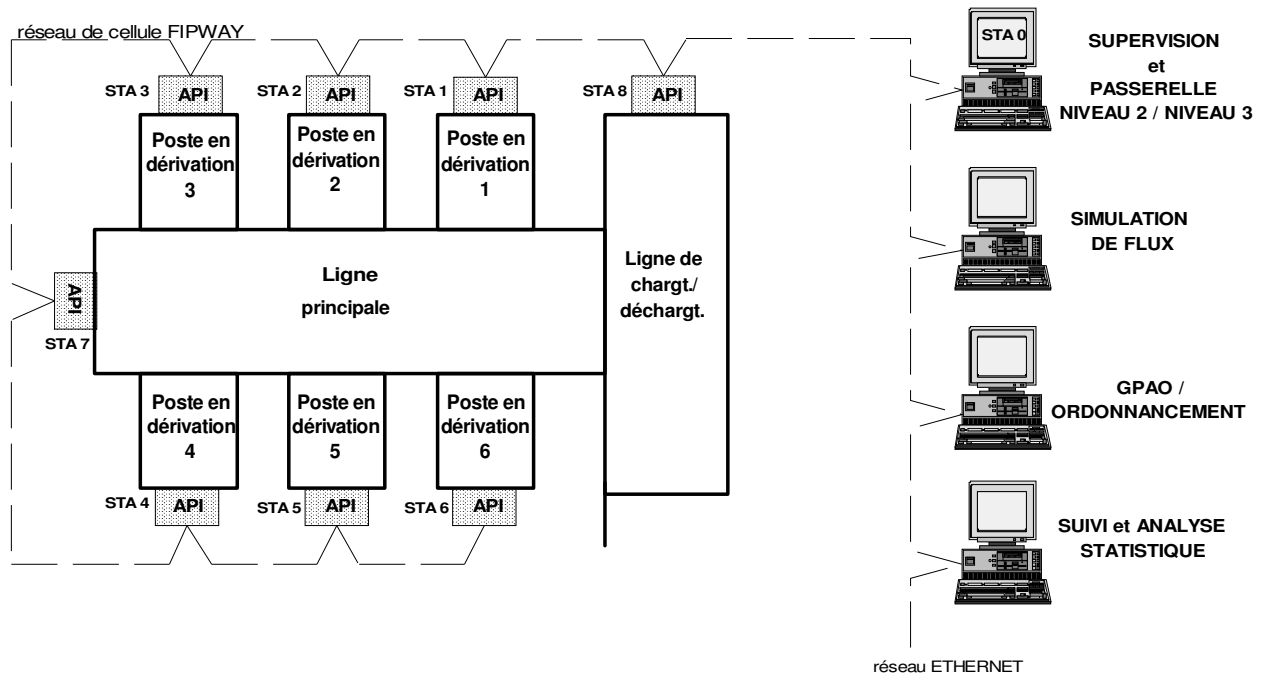


Figure 8 : Structure informationnelle générale du système expérimental

- **Au niveau 0**, on trouve tous les préactionneurs, actionneurs et détecteurs du système physique : contacteurs et moteurs d'entraînement des convoyeurs, distributeurs et vérins des butées, d'indexages et de dérivations, détecteurs de fins de course ou de palette, plots de lecture/écriture des codes de palettes,Ce niveau ne présente pas de liaison inter-niveau.

- **Au niveau 1**, on trouve les parties commandes de chaque module fonctionnel qui sont réalisées à partir d'automates programmables. Chaque poste en dérivation, la ligne principale et la ligne de chargement/déchargement sont commandés par leur propre automate programmable. La liaison inter-niveau entre ces différents automates est réalisée par un réseau local d'automates (type FIPWAY).

- **Au niveau 2**, un ordinateur relié au réseau d'automates assure la supervision de la ligne et la communication par échange de données avec le niveau supérieur, notamment l'ordonnancement et la GPAO, mais aussi des outils tels que la simulation de flux et le suivi statistique.

Conclusion

Cette partie a permis de présenter le système physique qui sert de support expérimental à cette thèse.

La structure retenue, avec des postes de travail implantés en dérivation, autorise une grande variété de routages des palettes. Après avoir détaillé les principaux séquencements des opérations, on précise que la suite du travail sera limité au séquencement de type "*Gamme liée / phases non redondantes*" qui est le plus représentatif des processus d'assemblage. En effet, celui-ci oblige à piloter le flux de palettes en respectant une contrainte d'antériorité (ou de précédence) sur les opérations de la gamme, ce qui est le plus souvent le cas en assemblage.

Les paramètres et l'algorithme d'accès d'une palette sur un poste sont ensuite expliqués, en précisant les paramètres intrinsèques à la palette et ceux intrinsèques au poste. La réalisation pratique de la prise en compte de ces paramètres nécessite d'utiliser un dispositif de codage individuel des palettes.

On termine cette partie en présentant la structure de commande et en faisant apparaître les outils informatiques intégrés dans cette structure au niveau 2.

- PARTIE 4 -
***Modélisation et simulation
de flux de production***

Sommaire

Introduction	90
I- Définitions.....	92
<input type="checkbox"/> I.1- Définition de la simulation de flux	92
<input type="checkbox"/> I.2- Classification des modèles de simulation	92
<input type="checkbox"/> I.3- Terminologie utilisée en simulation à événements discrets.....	96
II- Limites et objectifs de la simulation	98
<input type="checkbox"/> II.1- Ce que la simulation peut faire.....	98
<input type="checkbox"/> II.2- Ce que la simulation ne peut pas faire (seule).....	99
III- Résumé de quelques travaux de recherche utilisant la simulation	101
IV- Méthodologie de conduite d'une simulation	102
<input type="checkbox"/> IV.1- Analyse du problème	103
<input type="checkbox"/> IV.2- Modélisation et programmation.....	104
<input type="checkbox"/> IV.3- Expérimentations sur le modèle.....	106
<input type="checkbox"/> IV.4- Rapport et conclusions	108
V- Approfondissement de la démarche de modélisation	110
<input type="checkbox"/> V.1- Objectifs d'un Modèle Graphique de Connaissances	111
<input type="checkbox"/> V.2- Etat de l'art et pratiques industrielles sur les modèles graphiques.....	112
<i>V.2-a) Analyse de déroulement</i>	<i>112</i>
<i>V.2-b) Travaux de recherche.....</i>	<i>115</i>
<input type="checkbox"/> V.3- Proposition d'un Modèle Graphique de Connaissances	118
VI- Implémentation du modèle du système expérimental	126
<input type="checkbox"/> VI.1- Modèle géométrique du système physique	126
<input type="checkbox"/> VI.2- Modèle Graphique de Connaissances	127
<input type="checkbox"/> VI.3- Modèle d'action.....	129

Conclusion 133

Bibliographie 134

Introduction

Dans cette partie on va présenter et justifier l'utilisation d'un outil informatique permettant de simuler le comportement dynamique de la ligne expérimentale. Bien que l'on dispose du système réel, en grandeur nature, des expérimentations, mêmes simples, entraînent des temps de réponse qui peuvent s'avérer importants. De plus, la collecte de données peut présenter non seulement des difficultés au niveau du mesurage, mais aussi certaines imprécisions. Pour réduire, voire même supprimer ces problèmes, deux possibilités sont offertes :

- une modélisation mathématique,
- une modélisation simulateur.

C'est la deuxième possibilité qui est utilisée dans ce travail de recherche. On peut justifier ce choix par le fait que le modèle de simulation peut reproduire très précisément le comportement d'un atelier de production et est utilisé comme une expérience virtuelle. Il ne fournit pas directement une solution analytique au problème posé. Dans cette partie, et pour l'ensemble de ce travail, il est sous-entendu que la simulation dont on parle est à événements discrets (voir chap. 1).

Il faut aussi rappeler que la simulation de flux est un outil informatique qui est de plus en plus utilisé par les industriels et par les chercheurs. [LAW 91, page 2] cite plusieurs enquêtes (aux Etats-Unis) qui montrent la place de la simulation dans l'industrie : l'une d'elles indique que parmi 14 techniques utilisées, la simulation arrive en 2^{ème} position pour 84% des entreprises sondées. Pour les chercheurs, le principal intérêt est de pouvoir travailler sur un système de production virtuel, dont le comportement peut être très proche du système réel, à moindre coût et sans aucun risque. Dans le domaine de l'optimisation et de la prise de décision, les autres avantages de la simulation font que cet outil permet, depuis une dizaine d'années seulement, de mettre en œuvre des méthodes qu'il était inimaginable d'appliquer sur les systèmes réels ou sur des modèles mathématiques.

Dans le chapitre 1, on définit la simulation de flux à événements discrets ainsi que la terminologie utilisée dans ce domaine.

Le chapitre 2 permet de voir quelles sont les limites de la simulation de flux.

Des travaux de recherche utilisant la simulation de flux sont résumés dans le chapitre 3.

Une méthodologie de conduite d'une étude de simulation est proposée dans le chapitre 4. On présente les différentes étapes à suivre, et parmi elles, l'étape de modélisation qui est la plus délicate.

Dans le chapitre 5, l'étape de modélisation est approfondie et l'on montre l'intérêt de l'utilisation d'un modèle graphique de connaissance. Après un état de l'art des modèles

graphiques permettant de décrire et d'analyser les flux, on propose un formalisme qui nous semble bien adapté à la simulation de flux. Plusieurs modèles graphiques sont élaborés à partir de ce formalisme, appliqués aux différents modes de pilotage des flux vus dans la partie 1.

Ce formalisme est ensuite appliqué au système expérimental dans le chapitre 6. On présente en fin de chapitre le codage du modèle d'action avec le progiciel WITNESS.

I- Définitions

□ I.1- Définition de la simulation de flux

La simulation des flux de production intègre à la fois la construction d'un modèle et l'utilisation expérimentale de ce modèle pour étudier un problème.

Le modèle consiste en une représentation d'un système réel, capable de reproduire son fonctionnement.

La simulation est l'activation du modèle dans le temps, afin de connaître son comportement dynamique et de prédire son comportement futur [CLAVER 97].

La simulation est principalement utilisée pour étudier les flux physiques (pièces, matières, outils, etc...) et informationnels (Ordres de Fabrications, Kanban, etc...) dans l'atelier et les disponibilités des ressources (opérateurs, machines, convoyeurs, etc.).

Elle ne permet pas de trouver directement et de façon optimale des solutions à des problèmes de production. En cela, on peut l'assimiler à une simple boîte noire qui réagit aux consignes qu'on lui donne (variables d'entrée), mais incapable, seule, de déterminer la valeur optimale de ces consignes.

On l'utilise en général pour évaluer et comparer des scénarios possibles.

Ses capacités d'imitation et de prédiction permettent d'obtenir des renseignements sur les conséquences de changements ou de modifications dans l'atelier (au niveau physique ou décisionnel), avant que ceux-ci ne soient effectués.

Les modèles de simulation sont capables de décrire le système avec le degré de détail et de précision nécessaire qui convient à la résolution du problème posé.

Cette description inclut la partie physique de l'atelier , mais peut aussi inclure certains aspects du système de pilotage (gestion de production).

□ I.2- Classification des modèles de simulation

[LAW 91] propose une classification des modèles de simulation qui distingue les *Modèles Physiques* des *Modèles Logico-mathématiques*.

- **Les modèles physiques** sont ceux dans lesquels le système réel est représenté par une réplique ou une maquette, à une échelle différente et éventuellement à l'aide de matériaux différents. Ils sont utilisés à des fins d'entraînement :

simulateurs de vol, de conduite, maquettes de véhicules pour des essais aérodynamiques,...

- **Les modèles Logico-Mathématiques ou symboliques** sont définis par des relations logiques et quantitatives qui sont manipulées et changées pour voir comment le modèle du système réel réagit. Ils sont exécutés sur des ordinateurs. C'est exclusivement ce type de modèle qui sera utilisé dans la suite de ce travail.

Une autre distinction concerne la prise en compte d'aléas ou de variations aléatoires dans le modèle.

- Si le système est indépendant de l'influence de variables aléatoires ou imprévisibles, on utilise un **modèle déterministe**.

- Si les aléas jouent un rôle significatif dans le comportement du système (exemple typique : les pannes), on utilise un **modèle stochastique**. C'est ce type de modèle qui sera utilisé dans la suite de ce travail.

Une troisième classification distingue :

- **les modèles statiques**, pour lesquels le temps n'intervient pas.

Exemple : modèle comptable permettant de calculer le bénéfice en fin d'année à l'aide d'un tableur.

- **les modèles dynamiques**, pour lesquels le comportement est une fonction du temps.

Exemple : système de maintenance dans une usine.

Enfin, à l'intérieur des *modèles dynamiques*, on distingue :

- **les modèles à événements discrets** (ou discontinus) dans lesquels les changements d'état ne surviennent que lors d'événements tels le début ou la fin d'une opération, la mise en attente d'une pièce dans un stock, la libération d'une ressource, ... Dans une simulation à événements discrets, les flux essentiels que l'on examine sont composés d'éléments isolables que l'on peut dénombrer et identifier individuellement. Ces éléments sont couramment appelés "Entités" ou "Articles" (voir Chap. 1.3).

- **les modèles continus**, plus adaptés aux flux continus, qui utilisent des équations mathématiques pour prendre en compte les changements d'état qui s'effectuent de façon continue au cours du temps. Les valeurs des variables d'état sont recalculées régulièrement selon un pas d'horloge d'après ces équations.

Exemple : un réacteur chimique.

- **les modèles combinés** (ou mixtes), qui intègrent les deux aspects.

Exemple : industrie métallurgique ou agro-alimentaire.

Synthèse de cette classification :

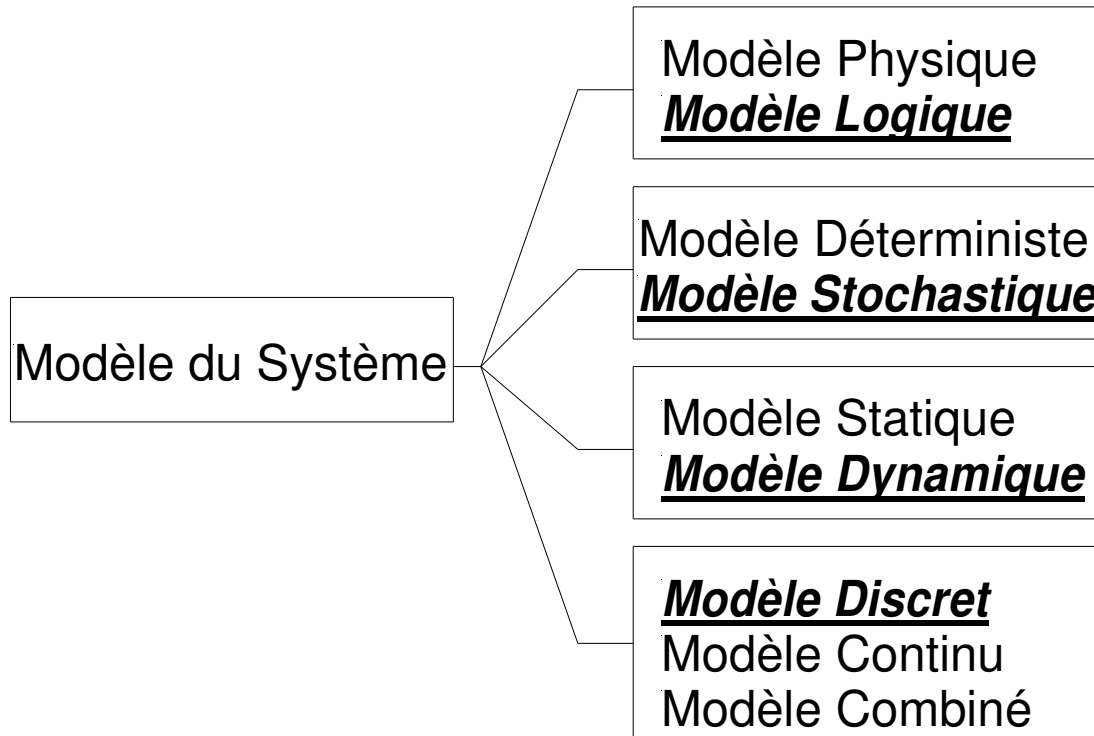


Figure 1 : Classification des modèles
(en souligné et italique, le type de modèle utilisé dans ce travail)

Dans la suite de ce travail, seuls les modèles à événements discrets seront utilisés.

Dans [CERNAULT 88], on trouve encore une autre classification de la modélisation pour la simulation de flux en fonction de l'approche :

- L'approche par événements :

C'est l'approche la plus générale. Elle consiste à rassembler tous les événements qui peuvent se produire et à décrire la logique des changements d'état.

On peut distinguer dans la logique des changements d'état :

- des règles liées aux modes opératoires sur le procédé de fabrication.

Exemple : fin de l'usinage d'une pièce.

- des règles liées à la conduite / gestion.

Exemple : Si le lot A est prêt avant le lot B, alors démarrer la fabrication du lot C sur le traitement thermique.

La difficulté des règles de conduite aux postes de fabrication est une des difficultés de cette approche.

Une fois le travail effectué pour la phase de modélisation, l'écriture du modèle consiste à programmer la logique des changements d'état. Le logiciel doit être en mesure de stocker la liste des événements créés. Le déroulement de la simulation consiste en la recherche, dans la liste, du prochain événement prévu.

L'échéancier est le module qui gère l'avance du temps au fur et à mesure que les événements apparaissent.

- L'approche par activités :

C'est une approche qui s'appuie sur un raisonnement naturel : un procédé est décrit comme l'enchaînement d'activités et d'attentes. On peut le regarder à travers les activités en indiquant les conditions nécessaires à leur début et à leur fin.

Les attentes débutent à la fin de chaque activité. Elles se terminent lorsque les conditions nécessaires à l'activité suivante sont réunies.

La modélisation consiste à programmer les conditions de déclenchement et de fin des activités. A chaque incrémentation du temps, on examine si les conditions de début et de fin des activités sont réalisées.

- L'approche par processus :

On parle de cette approche lorsque la modélisation consiste à rassembler des processus.

Les processus sont formés de séquences d'événements et d'activités : durée d'une activité, utilisation d'une machine, stockage de pièces, ...

Ces processus peuvent être paramétrés. Dans le logiciel, ils correspondent à des sous-programmes ou à des primitives avec lesquels on peut décrire le fonctionnement d'une installation. La puissance d'un logiciel sera liée aux processus mis à la disposition de l'utilisateur (manutention par robot, par convoyeur, ...).

Le graphique de la figure 2 [CETIM 89] donne un aperçu des principaux logiciels de simulation en faisant apparaître l'approche utilisée et le niveau de convivialité.

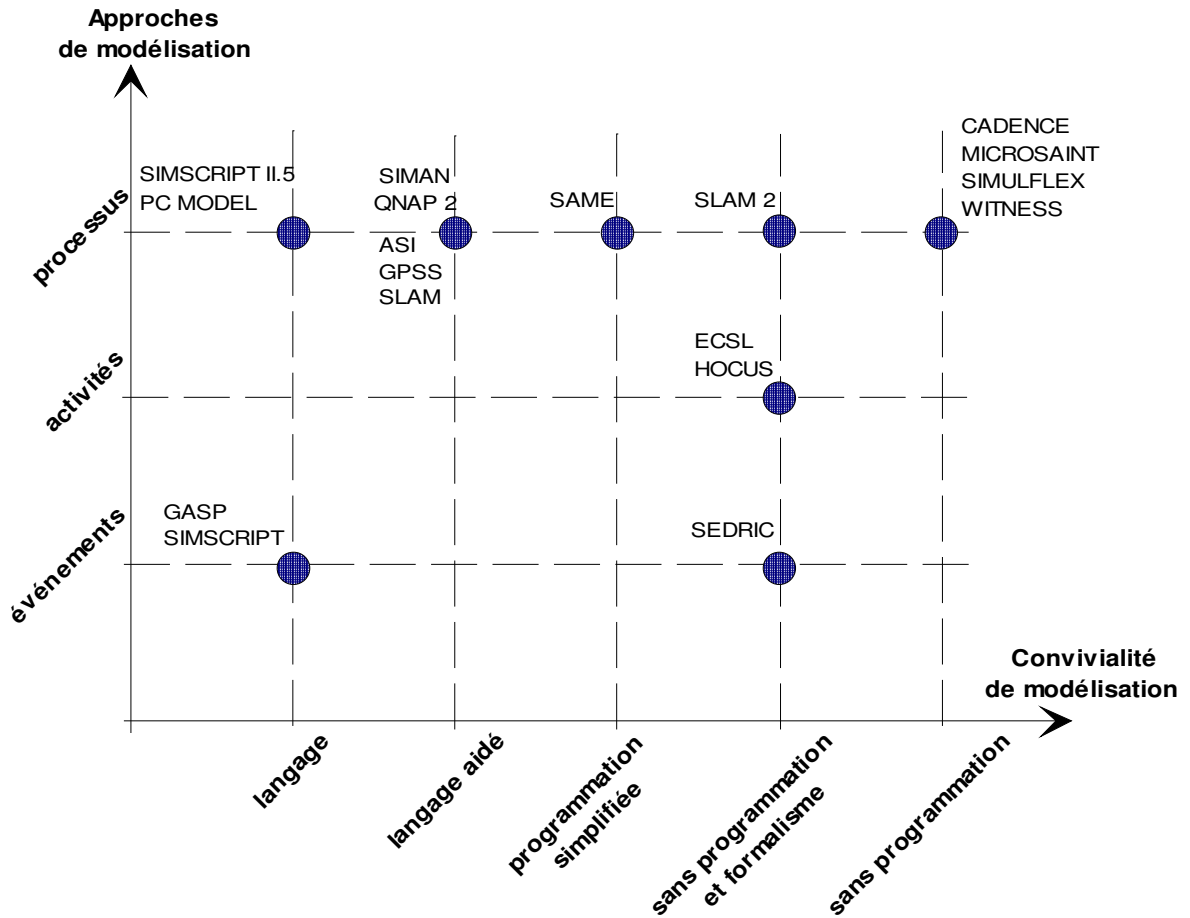


Figure 2 : Les approches de modélisation en fonction des progiciels

□ 1.3- Terminologie utilisée en simulation à événements discrets

Chaque progiciel de simulation de flux manipule des éléments de base qui sont :

- **Entité ou Article** : Un article est l'élément isolable composant le flux discret. Il se déplace individuellement à travers le modèle.

Exemple : Dans un atelier, les articles sont les pièces à fabriquer.

- **File d'attente ou Stock** : Les stocks sont des éléments qui permettent d'accumuler des articles entre deux machines. Les stocks peuvent être gérés de différentes façons : FIFO, LIFO, avec critère de priorité, en vrac,

- **Opération ou Machine** : C'est un élément qui prélève des articles d'un autre élément (un stock ou une autre machine), les traite pendant un temps de cycle qui correspond à un délai fixé dans une unité de temps appropriée.

Les articles, stocks et machines constituent l'essentiel des éléments physiques d'un modèle de simulation à événements discrets. Pour faire fonctionner le modèle ou pour prélever des informations, on a aussi besoin d'éléments logiques, qui se décomposent en deux catégories :

- **Les Variables** : Ce sont des valeurs qui caractérisent les éléments du système en entier. On peut y accéder depuis n'importe quel élément physique du modèle. L'analogie avec un langage de programmation structurée est la notion de *variable globale*.

- **Les Attributs** : Ce sont des variables spécifiques à un article donné et qu'il porte avec lui pendant sa durée de vie dans le modèle. C'est en quelque sorte la carte d'identité d'un article, dans laquelle on pourra mettre par exemple l'heure d'arrivée et l'heure de sortie d'un article, son temps de cycle sur une machine, le nombre de trous à percer sur une autre machine, etc.... Plusieurs articles différents ou identiques peuvent donc porter le même attribut, mais celui-ci pourra avoir des valeurs différentes.

Une autre particularité des modèles à événements discrets est le type d'expérimentation en simulation que l'on veut faire. Au moment où l'on lance une simulation (à $T = 0$), le modèle est vide et ne contient aucun article. Deux types d'expérimentation peuvent être programmés :

- **Expérimentation à Horizon Fini ou Terminatif** : Dans ce cas, la simulation s'exécute jusqu'à ce qu'une condition soit vérifiée. Par exemple un atelier qui fonctionne jusqu'à ce que tous les articles aient été traités. Au début de la simulation, l'atelier est vide, tout comme à la fin de la simulation.

- **Expérimentation à Horizon infini ou à Régime Stabilisé** : dans ce type d'expérimentation, on désire étudier le comportement de l'atelier uniquement en régime stabilisé, sans tenir compte de la période de montée en régime ou du vidage de l'atelier. Il n'y a donc pas forcément de condition d'arrêt liée à l'absence d'article, mais l'on définit une durée de simulation suffisamment significative pour laisser au système le temps de se stabiliser. Aucune méthode analytique n'a été trouvée dans la littérature pour déterminer la durée de la période de montée en régime. En pratique, pour la déterminer, les experts de la simulation recommandent de mesurer le niveau total d'encours dans le modèle ou le flux sortant (en

article/unité de temps). En effet, au début de la simulation, le niveau d'encours va monter progressivement puis se stabiliser. De la même façon, le flux sortant sera d'abord nul, puis va augmenter et se stabiliser entre deux valeurs acceptables.

Lors des expérimentations menées pour ce travail (voir la Partie 5), ces deux types de simulation seront réexaminées pour le cas étudié.

II- Limites et objectifs de la simulation

□ II.1- Ce que la simulation peut faire

Avant d'utiliser cet outil, il est important de savoir ce que l'on peut en attendre, notamment par rapport aux outils de modélisation mathématique.

Voici une liste non exhaustive des possibilités offertes à l'utilisateur de la simulation de flux :

- FOURNIR DES ESTIMATIONS REALISTES :

- du comportement attendu du système,
- des variations à l'intérieur du système.

- EVALUER LES EFFETS DES ACTIONS SUIVANTES :

- ajouter, déplacer ou supprimer des machines,
- modifier les flux,
- modifier les durées de processus et / ou de montage,
- introduire de nouveaux produits ou supprimer des produits existants,
- modifier les systèmes de manutention,
- modifier l'ordonnancement et la répartition des tâches,
- ajouter ou supprimer du personnel,
-

Cette liste montre bien l'intérêt qui peut être porté à cet outil informatique, non seulement auprès des industriels, mais aussi auprès des chercheurs pour vérifier leur théorie.

□ II.2- Ce que la simulation ne peut pas faire (seule)

- Elle ne peut pas optimiser la performance d'un système. Elle peut seulement donner des réponses à des questions du genre : " Qu'est-ce qui se passe si ... ?"

Comme on l'a déjà dit, une fois que le modèle est programmé et validé, la simulation fonctionne comme une boîte noire en fonction d'un scénario de fonctionnement. Elle ne fait donc que reproduire le comportement du système modélisé.

- Elle ne peut donner des résultats justes si les données sont imprécises.

Un point très important, et sur lequel on n'insiste jamais assez, concerne les données utilisées par la simulation. Il est indispensable de vérifier la validité de ces données sous peine d'obtenir des résultats sans commune mesure avec ceux obtenus sur le système réel. Cela est particulièrement critique quand le système simulé est d'un niveau de complexité ne permettant pas une vérification analytique des résultats fournis par la simulation. L'article de [BALLOT 97] qui compare la simulation de flux avec la recherche opérationnelle, indique que la qualité des résultats de la simulation dépend exclusivement de la qualité des règles de fonctionnement et des données introduites lors de la modélisation. Les outils de simulation permettent toujours d'obtenir un résultat, mais n'apportent rien quant à sa validité par rapport au système réel que l'on simule.

- Elle ne peut décrire les caractéristiques d'un système qui n'a pas été complètement modélisé.

La phase de modélisation se déroule généralement en plusieurs étapes. On commence par construire un modèle global du système, puis on l'affine progressivement, en suivant éventuellement une méthodologie [DUFRENE 94], en validant chaque étape avant de perfectionner le modèle. Le problème qui se pose lors de la modélisation est de savoir jusqu'à quel niveau de détail il faut aller pour que le modèle soit représentatif de la réalité. En ce sens, il est donc préférable de rechercher à affiner le modèle, plutôt que de se contenter d'un modèle trop général et imprécis.

- Elle ne peut résoudre des problèmes mais seulement fournir des indications à partir desquelles des solutions peuvent être déduites.

Lorsque le modèle est prêt à être utilisé pour la simulation, il faut encore se poser la question de savoir quelles sont les informations que l'on veut observer pendant ou à la fin de la simulation. Ces informations, appelées "Indicateurs" sont des variables du système sur lesquelles un traitement statistique est effectué. La majorité des progiciels proposent en standard des résultats de simulation sous la forme de rapports statistiques qui peuvent

être agrémentés de graphiques pour faciliter l'analyse. Les indicateurs "standard" concernent essentiellement les quantités et les durées relatives aux flux d'articles, ainsi que les taux d'utilisation, de panne, d'arrêt des machines et ressources. Les figures 3 et 4 présentent des exemples de rapports statistiques.

C'est uniquement à partir de l'observation du modèle et des indicateurs pendant et après la simulation que l'on peut tenter de résoudre le problème posé.

Nom	Qté. Entrée	Qté. Expédiée	Qté. Rebutée	Qté. Assemblée	Qté. Rejetée	En Cours	E. Cours Moy.	Tps Moy.
TETE	601	0	0	601	1	0	4	13.98
CORPS	603	0	0	601	0	2	3.6	12.53
VIS	1206	0	0	1202	0	4	2.46	4.28
ASSM	601	600	0	0	0	1	2.85	9.98

Figure 3 : Rapport statistique des Articles

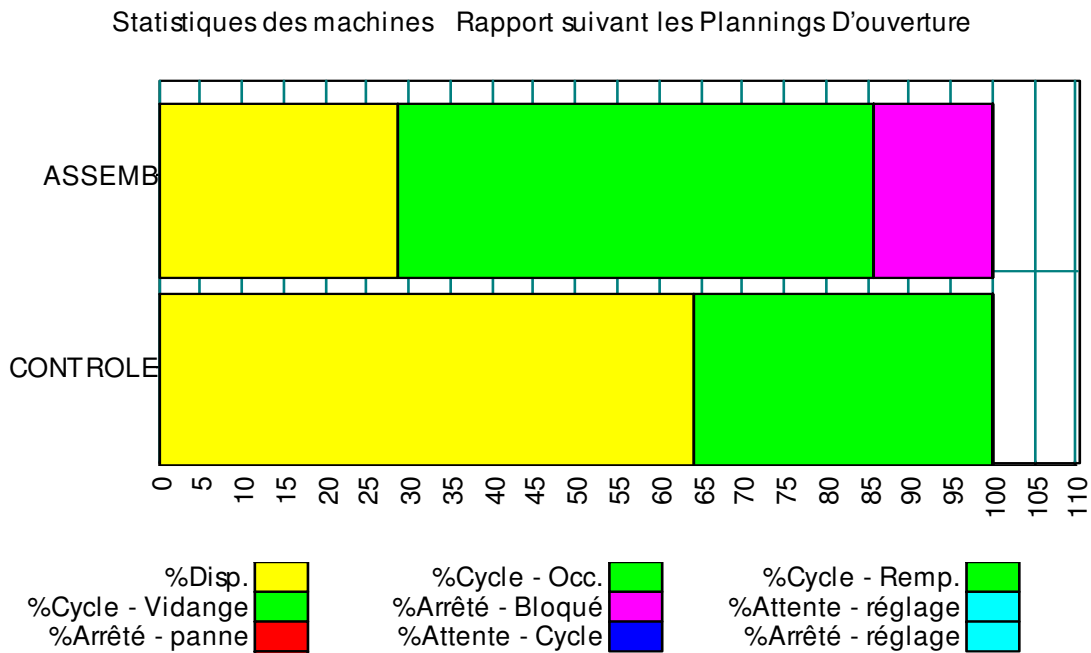


Figure 4 : Rapport statistique des Machines

III- Résumé de quelques travaux de recherche utilisant la simulation

Beaucoup de travaux de recherche s'appuient sur la simulation pour traiter des problèmes d'ordonnancement. En voici quelques exemples :

- [CAUX 95] pose le problème des dates d'introduction des pièces dans un système de traitement de surfaces. La simulation est utilisée pour évaluer l'ordonnancement proposé par une heuristique de placement des pièces et une optimisation par la méthode du Grimpeur (voir partie 7).
- [GOUJON 96] présente un étude destinée à optimiser la planification d'un chantier de moulage au sable. L'optimisation est réalisée par un couplage entre simulation et algorithmes stochastiques de type Recuit Simulé, Grimpeur et Kangourou. Les améliorations du temps de production sont considérables puisqu'elles atteignent 80% de gain par rapport à la solution initiale avec la méthode Kangourou.
- [BOUKACHOUR 96] et [BAVAY 96] proposent un système d'ordonnancement par simulation qui permet de planifier non seulement "au plus tôt", ce qui est la démarche classique, mais aussi "au plus tard". Ce mode d'ordonnancement est appelé ici "Retro-simulation" ou simulation arrière.

La simulation peut aussi être utilisée dans le domaine de la maintenance des systèmes de production, pour choisir des indicateurs ou pour mesurer les effets des pannes sur le fonctionnement global du système.

- [PERES 97] utilise la simulation de flux pour tester des indicateurs de performances prenant en compte les effets directs des actions de maintenance. La simulation doit pouvoir restituer les trois processus de base dont dépend l'évolution des systèmes de production : processus de sollicitation/réponse, processus de défaillance/réparation et processus d'inspection/révision. Il est précisé par ailleurs que les logiciels de simulation offrent la possibilité de prise en compte des défaillances, mêmes s'ils ne présentent pas de fonctionnalités dédiées à la maintenance.

Enfin, on peut citer quelques travaux récents qui utilisent à la fois le concept de modélisation "multi-agents" pour spécifier le système de conduite et la simulation pour tester les algorithmes de résolution des agents :

- [QUERE 97] simule le fonctionnement d'un convoyeur à entrées/sorties multiples dont le sous-système de décision est modélisé par l'approche multi-agents. Un agent est placé sur chaque poste de chargement du convoyeur et la simulation est utilisée pour évaluer les algorithmes de résolution des agents.

- [KIEFFER 98] étudie un processus de production complexe (fabrication de semi-conducteurs) qui comporte de nombreux bouclages de flux et qui pose le problème de la synchronisation de composants. La simulation est utilisée parce qu'aucun outil d'ordonnancement n'est capable de prendre en compte en un temps raisonnable le nombre important de tâches à traiter, les multiples rebouclages et les aléas des équipements. Quant à l'approche multi-agents, elle permet de décentraliser et de coordonner les centres de décision d'ordonnancement afin d'augmenter la réactivité du pilotage face aux aléas. Le simulateur est utilisé en phase d'exploitation de la ligne de production et permet d'affiner les temps de cycle. D'autre part, il peut être alimenté par des données de suivi de production et permettre ainsi d'indiquer aux exploitants les décalages de délai.

IV- Méthodologie de conduite d'une simulation

Dans ce chapitre, l'objectif est de proposer une méthodologie générale pour la réalisation d'un projet de simulation de flux. L'objectif de cette méthodologie est non seulement de fournir un fil conducteur à tout projet de simulation, mais aussi de s'assurer, avant le début du projet, que l'on est en possession de toutes les données utiles. Une étude de simulation de flux se déroule généralement en quatre macro-étapes (figure 5):

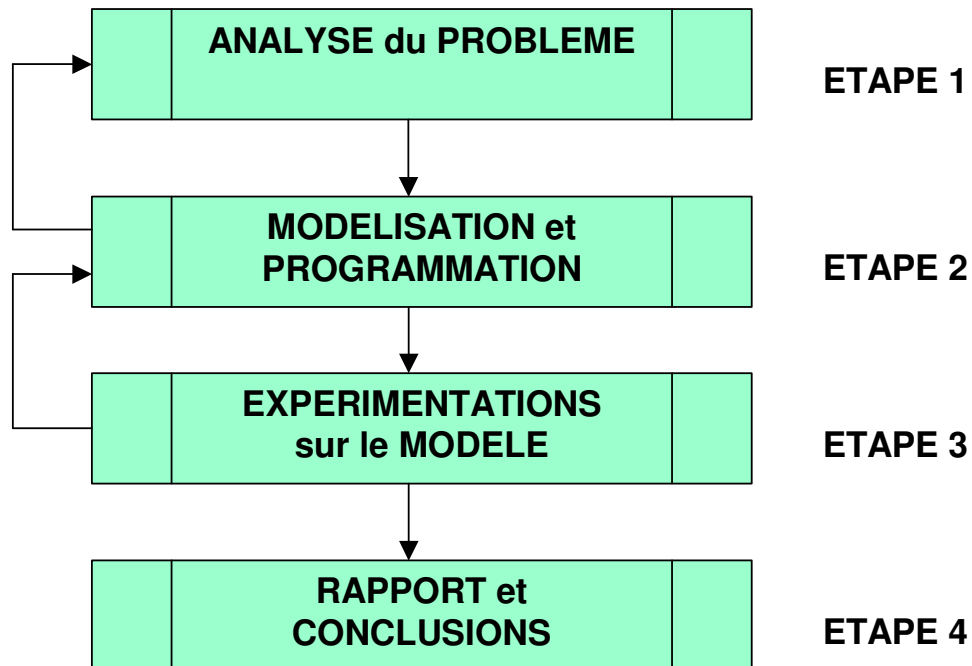


Figure 5 : Les 4 étapes d'une étude de simulation

Comme on le voit sur la figure 5, ces macro-étapes se déroulent de façon séquentielle, mais des rebouclages ou des retours en arrière sont possibles afin de corriger ou de compléter l'élaboration du modèle en fonction des objectifs fixés.

Chacune de ces macro-étapes va maintenant être détaillée en plusieurs étapes et l'on va faire ressortir plus précisément l'enchaînement du projet.

□ IV.1- Analyse du problème

L'analyse du problème est un préliminaire indispensable et d'une grande importance, puisque c'est dans cette étape que l'on doit définir précisément ce que l'on veut mettre en évidence avec la simulation, et quelle précision on attend. On détermine aussi les indicateurs de performance qui vont permettre de vérifier si l'on atteint les objectifs que l'on s'est fixé. Enfin, il faut pouvoir fournir des données numériques au modèle. Celles-ci sont relatives à tous les éléments utilisés dans la simulation, comme par exemple :

- **Données sur les articles à fabriquer** : gammes de fabrication (routage), loi d'arrivée dans le modèle (fréquence et taille de lot).
- **Données sur les moyens de production** : nombre et types de machines, lois de pannes, nombre et types de ressources complémentaires de production (outils, palettes, etc...)
- **Données sur les systèmes de manutention** : nombre et type de convoyeurs ou de chariots, capacité en nombre d'articles, vitesse de déplacement, etc...
- **Données sur les stocks et les magasins** : types et capacités, règles d'entrée et de sortie, etc...
- **données sur le personnel** : effectifs, compétences et horaires.

A ces données numériques, il convient d'en ajouter d'autres qui s'expriment sous forme logique et qui caractérisent les règles de pilotage de l'atelier. Il est à noter que bien souvent, l'objectif de la simulation est de tester certaines règles pour déterminer les plus pertinentes. A titre d'exemple :

- **Règles de lancement et de séquençement des articles en fabrication** : à la commande ou en fonction des prévisions, par lots ou à l'unité, périodique ou apériodique,...
- **Règles d'ordonnancement** des articles sur les différentes machines de l'atelier.
- **Règles d'affectation des ressources.**

Il est à noter que de nombreux chercheurs utilisent la simulation pour tester de nouvelles règles de pilotage, notamment dans le domaine de l'ordonnancement et de la planification à moyen terme [DRAKE 98], [BOSTEL 97],

Enfin, en complément aux données numériques et logiques, on doit disposer de documents graphiques, à la fois pour avoir une représentation géométrique (ou spatiale) du système simulé, mais aussi pour avoir une représentation des flux. On utilisera en guise de modèle géométrique un plan de masse de l'atelier à simuler. La macrographie des flux permet de décomposer le processus en actions élémentaires tels que opérations, stockages, transferts et contrôles. Dans le chapitre 6 de cette partie, un modèle graphique de connaissance sera développé afin de formaliser cette macrographie des flux.

La figure 6 montre comment l'analyse du problème peut être découpée en plusieurs étapes.

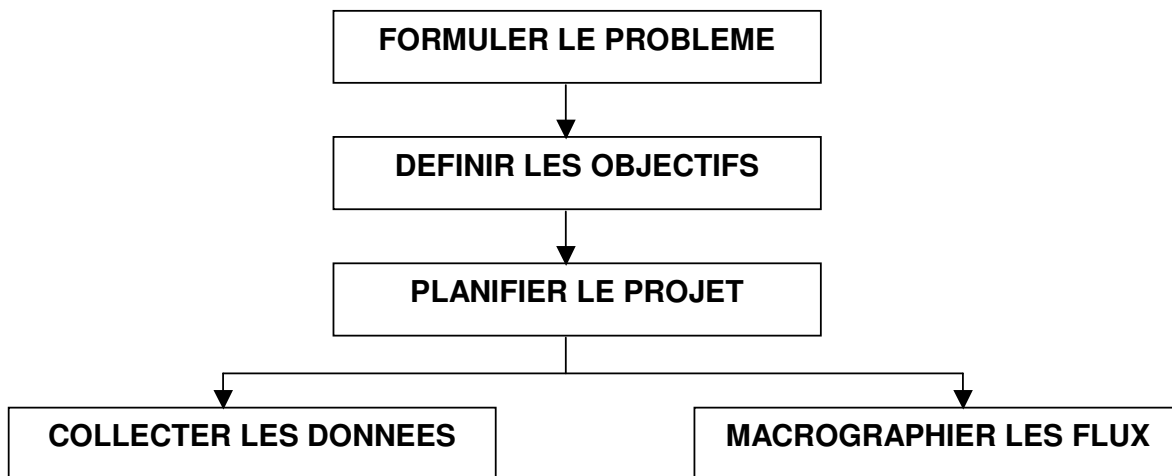


Figure 6 : Détail de la macro-étape "Analyser le problème"

□ IV.2- Modélisation et programmation

La construction du modèle permet de coder le modèle dans un langage informatique approprié. Cette étape est de plus en plus facilitée par l'évolution des progiciels dont la tendance est de substituer aux primitives d'un langage une interface graphique, interactive et conversationnelle. Le concepteur du modèle n'a plus besoin d'être un informaticien aguerri pour utiliser la simulation de flux. La figure 7 montre par exemple une fenêtre à renseigner pour modéliser une machine avec ce type de langage [LANNER 98].

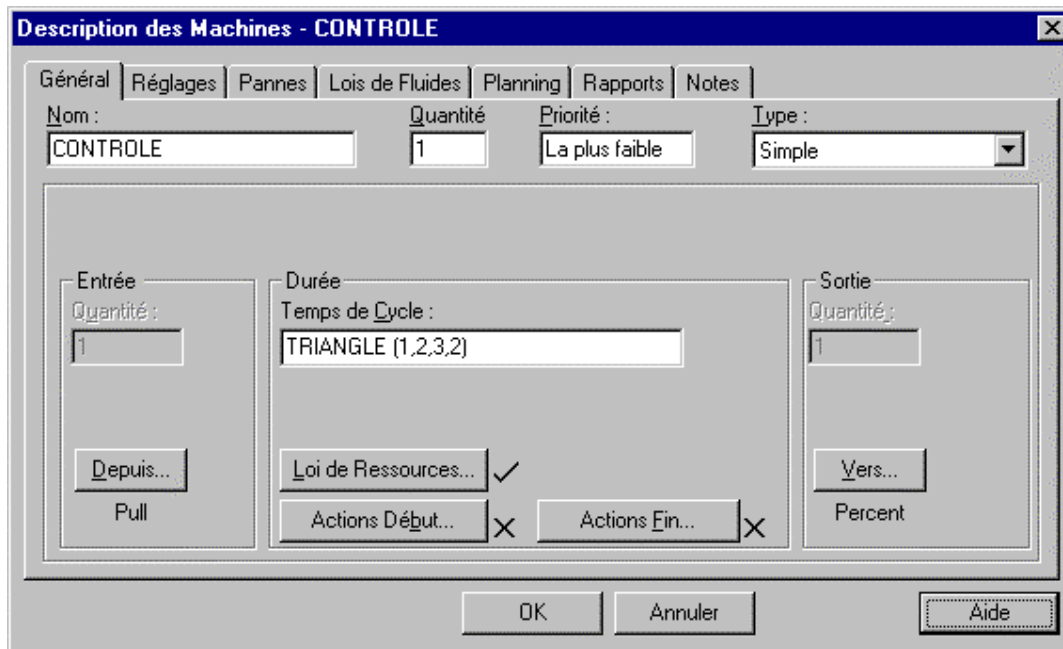


Figure 7 : Exemple de fenêtre de saisie d'un progiciel de simulation (Witness)

Une fois le modèle réalisé, il faut vérifier si les règles logiques qui décrivent le flux sont bien programmées et correspondent à ce qui est demandé. On exécute donc des tests de simulation uniquement pour vérifier le bon fonctionnement du modèle. Si des différences apparaissent, il faut pouvoir "tracer" tous les événements du modèle et vérifier leur cohérence par rapport aux données. La plupart des progiciels disposent d'une fonction qui permet de visualiser le fonctionnement du modèle événement par événement (figure 8).

```

9.00 : Article_A(s) arrive(nt)
11.79 : La machine PREPARATION quitte l'état Occupée
Article_B en sortie de l'élément PREPARATION vers Stock_2
Article_B en entrée de l'élément PEINTURE à partir de Stock_2
Article_A en entrée de l'élément PREPARATION à partir de Stock_1
Ressource obtenue par CONTROLE
12.00 : Article_B(s) arrive(nt)
Article_B en sortie vers Stock_1
12.00 : Article_A(s) arrive(nt)
13.00 : temps mis à jour
13.39 : La machine CONTROLE quitte l'état Occupée
Article_A en sortie de l'élément CONTROLE vers SHIP
Ressource obtenue par PEINTURE
13.89 : La machine PEINTURE essaye de quitter l'état Réglage pour le
réglage n° 1
Ressource obtenue par PREPARATION

```

Figure 8 : Exemple de visualisation de la trace des événements

Cette étape doit se terminer par une validation qui consiste par exemple à comparer les résultats fournis par le modèle aux résultats du système réel si celui-ci existe. Les rapports statistiques peuvent aider à cette validation entre résultats simulés et résultats réels.

La visualisation graphique et dynamique du modèle est un atout considérable pour en faciliter sa vérification et sa validation, puisque l'on peut suivre tous les changements d'état qui interviennent sur le modèle.

L'enchaînement des étapes de la modélisation et de la programmation est présenté sur la figure 9.

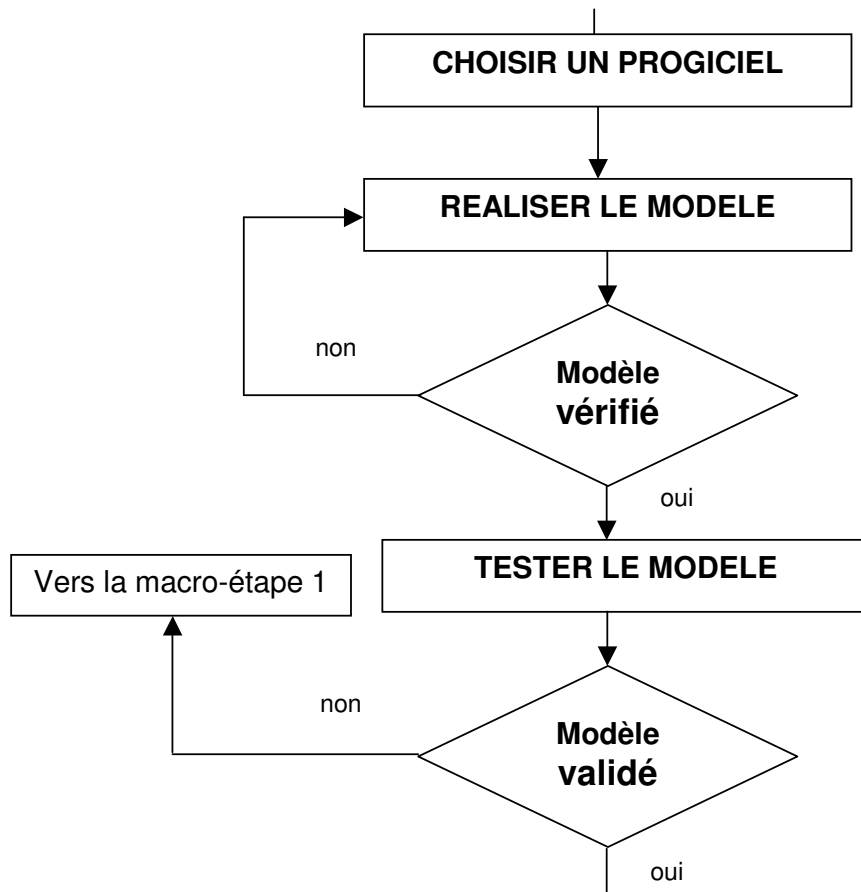


Figure 9 : Détail de la macro-étape "Modéliser et Programmer"

□ IV.3- Expérimentations sur le modèle

L'exploitation de la simulation est l'étape où l'on utilise le modèle comme support expérimental pour évaluer le comportement dynamique du système. Il faut bien sûr avoir

défini les données sur lesquelles on va pouvoir agir pour atteindre les objectifs que l'on s'est fixé dans la 1ère étape. Un scénario ou une expérimentation se caractérise donc par un jeu de données qui varie à chaque itération du processus expérimental. Des méthodes telles que les Plans d'Expériences [PILLET 92] peuvent être utilisées pour organiser les scénarios, réduire leur nombre et interpréter les résultats.

[DELAMARE 93] propose justement une utilisation combinée de la simulation de flux et des Plans d'Expérience (méthode Tagushi) pour améliorer l'organisation d'atelier dans l'industrie du bois. On trouve aussi dans cet article une représentation fonctionnelle (par diagrammes SADT) de la méthodologie de conduite d'une simulation.

On montrera dans la partie 7 (optimisation de la production) que l'utilisation des plans d'expériences n'est pas adaptée à des problèmes d'optimisation complexes, où le nombre de solutions est très grand. Dans ce cas, la simulation est couplée à un générateur de solutions et les résultats sont ensuite analysés par un algorithme d'optimisation. Cette approche fait partie de ce travail de recherche et est présentée dans la Partie 8.

Enfin, il faut être capable d'interpréter les résultats fournis par la simulation. Cela suppose la maîtrise de notions de statistique telles que l'*intervalle de confiance*, la *moyenne arithmétique ou temporelle*, l'*écart type* et éventuellement l'*analyse spectrale* [CERNAULT 88], [LAW 91].

La figure 10 présente le détail de cette macro-étape d'expérimentations sur le modèle.

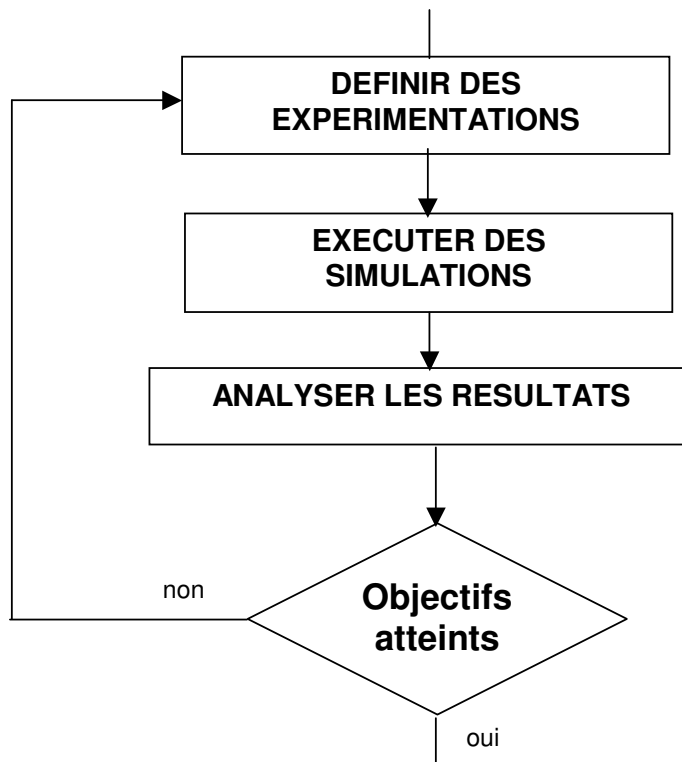


Figure 10 : Détail de la macro-étape "Expérimentations sur le modèle"

□ IV.4- Rapport et conclusions

Cette dernière macro-étape est importante vis-à-vis du demandeur de l'étude de simulation. En effet, celui-ci n'est pas forcément un expert dans ce domaine et il va falloir présenter les résultats de l'étude pour qu'ils soient compréhensibles. Entre autre, il faut tenir compte des remarques suivantes :

- Ne pas « noyer » le demandeur dans des tableaux de chiffres insignifiants pour lui.
- Expliquer les indicateurs utilisés.
- Justifier les hypothèses simplificatrices.
- Utiliser un grapheur pour présenter les résultats.
- Présenter les différentes solutions donnant le même résultat et proposer des critères de choix.
- Suggérer de nouvelles pistes d'étude.

Cette macro-étape est détaillée dans la figure 11 :

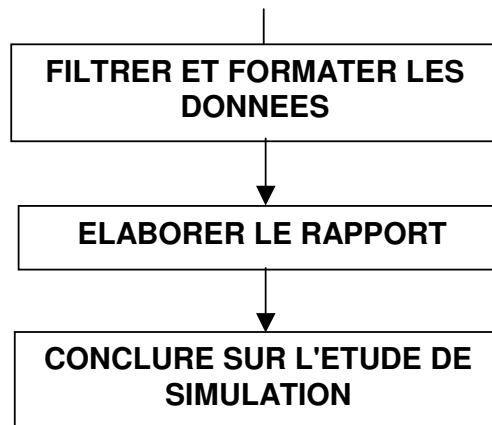


Figure 11 : Détail de la macro-étape " Rapport et conclusion"

Dans ce chapitre, une méthodologie de conduite d'un projet de simulation de flux a été proposée sous forme de macro-étapes, elles-mêmes découpées en plusieurs étapes et rassemblées sous la forme d'un organigramme. La figure 12 est la représentation complète de cet organigramme, étape par étape, en faisant apparaître les limites de chaque macro-étape.

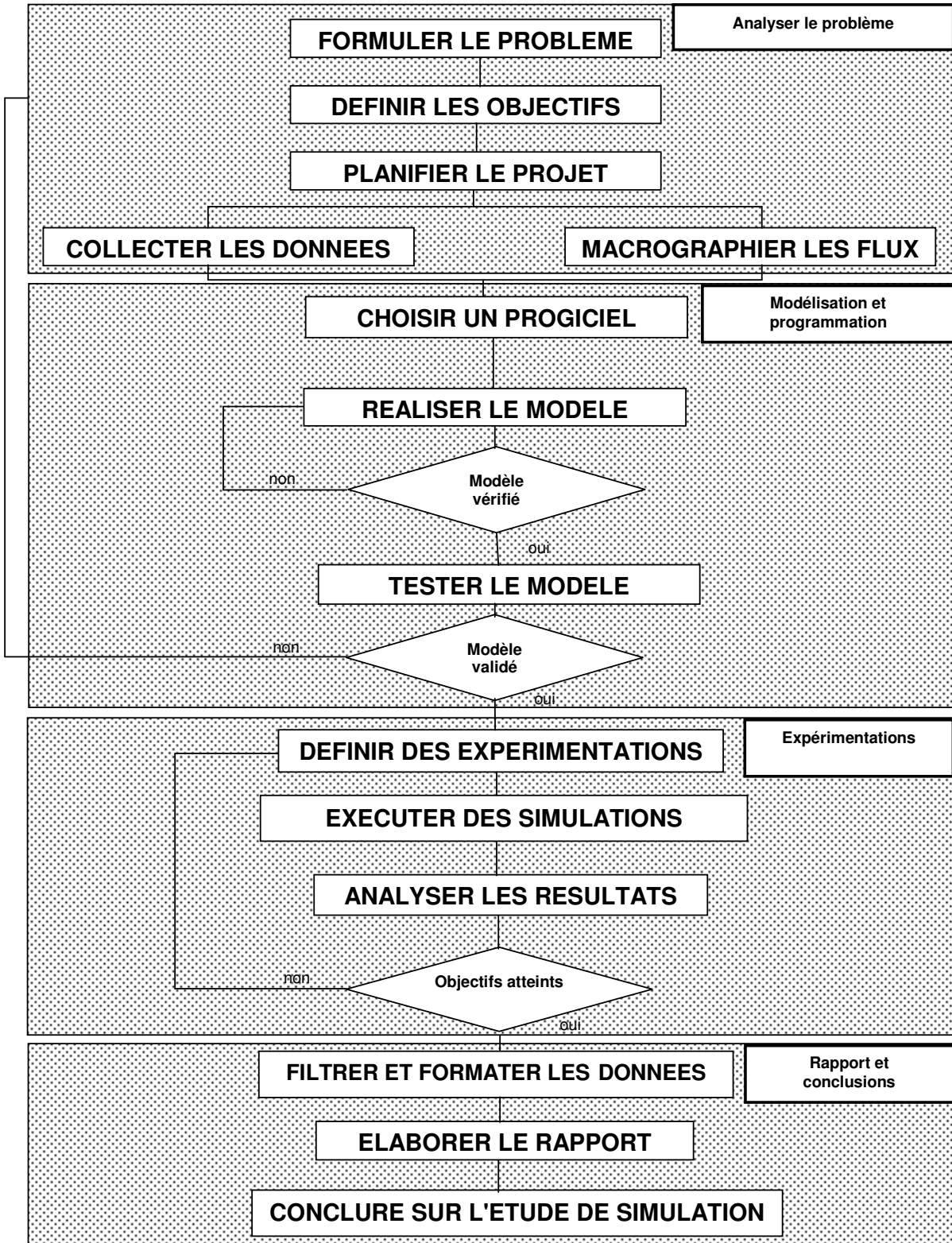


Figure 12 : Organigramme général de la méthodologie

V- Approfondissement de la démarche de modélisation

Dans le chapitre précédent, une méthodologie générale de conduite d'un projet de simulation a été présentée. Si les étapes d'*Analyse du problème*, d'*Expérimentations* et de *Conclusions* ne présentent pas de difficultés particulières, il n'en est pas de même pour l'étape de *Modélisation*. Bien qu'il ait été dit que la construction du modèle soit de plus en plus facilitée par l'évolution des progiciels, le concepteur d'un modèle de simulation de flux peut difficilement passer directement du système réel, existant ou à concevoir, au modèle informatique. Un certain nombre d'étapes intermédiaires sont nécessaires pour mener à bien la construction du modèle.

Les travaux de [QUERE 97] et de [HAMAIDI 97] démontrent le souci d'approfondir la démarche de conception d'un modèle, en découpant cette conception en plusieurs sous-modèles.

Le travail de [HAMAIDI 97], inspiré de la modélisation d'entreprise IDEF3 et CIMOSA [VERNADAT 96], privilégie une approche multi-modèles, qui décompose le modèle global de l'atelier en trois sous-modèles :

- *Le modèle des flux* : il décrit la circulation des flux physiques et informationnels dans le système de production.
- *Le modèle du système physique* : il décrit la géométrie du système physique et représente les ressources du système de production ainsi que leur capacité et leur localisation dans l'espace.
- *Le modèle de conduite* : il décrit les lois de commande et les règles d'allocation des ressources du système physique aux tâches en cours d'exécution.

La méthodologie présentée par [QUERE 97] décompose le processus de modélisation en quatre phases consécutives et éventuellement itératives (figure 13) :

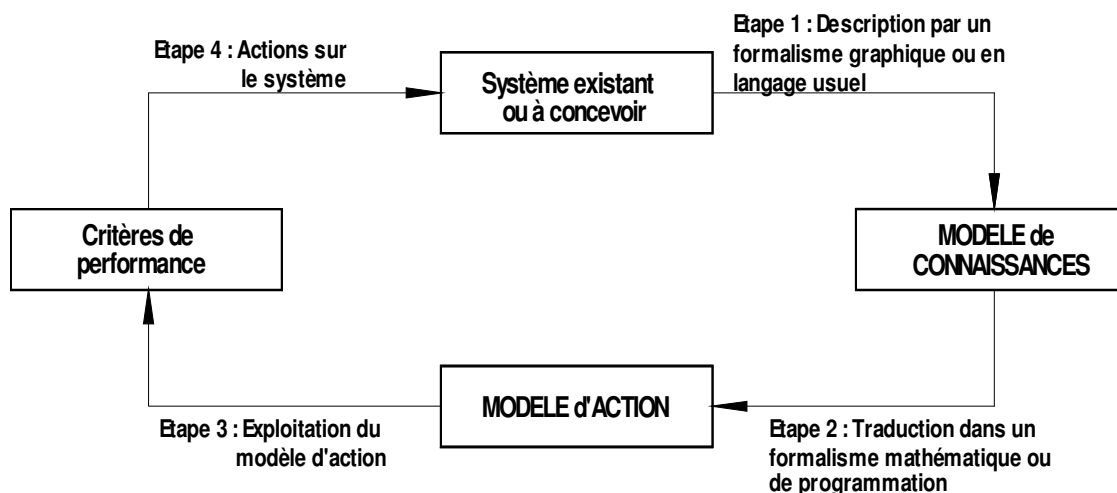


Figure 13 : Processus de modélisation

Cette méthodologie axée sur la modélisation est assez proche de la démarche générale présentée dans le chapitre IV et sur la figure 12. L'aspect qui semble important ici est de passer par un modèle intermédiaire de connaissances, dans un formalisme graphique par exemple, avant de passer au modèle d'action qui est traduit dans le langage du progiciel utilisé.

□ V.1- Objectifs d'un Modèle Graphique de Connaissances

La contribution qui peut être apportée aux idées présentées dans les deux travaux précédents est de créer un *Modèle Graphique de Connaissances (M.G.C)*, préliminaire à la modélisation informatique, sur lequel apparaît à la fois le modèle des flux et le modèle de conduite (le modèle géométrique pouvant être superposé au modèle des flux ou présenté séparément).

Les objectifs d'un tel modèle sont multiples :

- Structurer et rassembler les données utiles à la construction du modèle informatique.
- Servir d'outil de communication entre les différents acteurs (opérateurs, techniciens de maintenance, ingénieurs, responsables, ...) à la fois pour la conception d'un système de production mais aussi pour son exploitation.
- Permettre de représenter aussi bien les flux physiques qu'informationnels.
- Identifier et définir les règles et lois utiles au pilotage du système.
- Servir de support à la réalisation du modèle informatique.

L'une des caractéristiques essentielles de ce modèle est de devoir être construit à partir d'objets graphiques symbolisant les éléments du système réel.

Avant de présenter le formalisme graphique proposé pour ce modèle, un résumé des formalismes déjà utilisés dans l'industrie ou développés dans des travaux de recherche est abordé dans le chapitre suivant.

□ V.2- Etat de l'art et pratiques industrielles sur les modèles graphiques

V.2-a) Analyse de déroulement

Ce formalisme est très utilisé dans l'industrie, en tant que complément graphique à une gamme ou pour décrire un processus de production. Il permet d'analyser l'enchaînement des actions qui composent un processus. Ces actions sont décomposées en :

- **OPERATIONS**, destinées à modifier les caractéristiques de la matière.

Exemple d'opérations : usinage, formage, assemblage, etc....

Les opérations d'un processus de production industrielle font presque toujours appel à des machines qui peuvent être automatisées ou pas.

- **TRANSFERTS**, destinés à modifier la localisation de la matière.

- **STOCKAGES** , destinés à laisser s'écouler le temps sans agir sur les caractéristiques ni sur la localisation de la matière.

- **CONTROLES**, destinés à contrôler la qualité du produit, généralement par comparaison ou par mesurage,

- **ATTENTES**, destinées à découpler les opérations qui ont des temps opératoires différents.

Des symboles, plus ou moins conventionnels, sont utilisés pour permettre une visualisation de ces différentes actions (figure 14).






Opération	
Transfert	
Stockage	
Contrôle	
Attente	

Figure 14 : Symboles d'une analyse de déroulement

La liaison de ces différents symboles permet d'obtenir un graphique des flux. La figure 15 est un exemple de l'utilisation de ce formalisme.

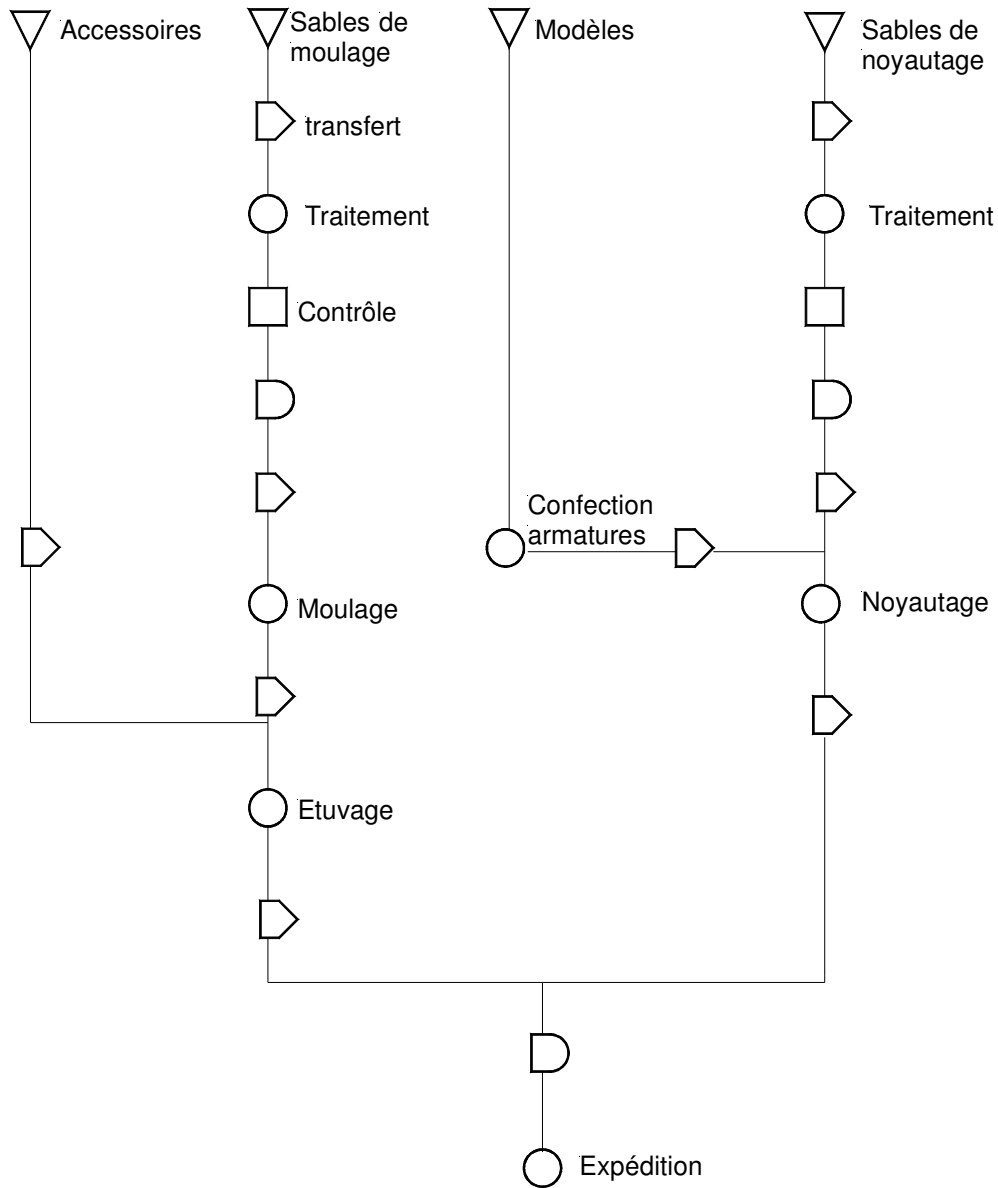


Figure 15 : Exemple d'analyse de déroulement

Un autre forme rencontrée est la matrice de déroulement qui permet de rassembler des informations quantitatives en plus des actions du processus (figure 16).

▽	▹	○	◐	□	Distance	Temps	Quantité	Poids	Déroulement
●									Sortie magasin matière
	●				70 m	0,3 h	150	300 kg	vers sciage
		●				0,05 h/p			Sciage
			●			1,4 h			Attente stabilisation
	●				25 m	0,12 h	150	300 kg	vers fraisage
		●				0,08 h/p			Fraisage
			●		200 m	0,9 h	100	200 kg	vers MMT
				●		0,5 h/p			Mesurage
	●				50 m	0,25 h	100	200 kg	vers montage

Figure 16 : Exemple de matrice de déroulement

Ce formalisme présente les avantages d'être très répandu dans l'industrie, d'utiliser des symboles simples à mémoriser et d'un niveau de détail pas trop important, de mettre en évidence les actions qui ne correspondent pas à de la valeur ajoutée (toutes celles qui ne sont pas des *OPERATIONS*). On obtient un modèle graphique synthétique et général des flux physiques.

Parmi les inconvénients, on peut justement regretter de ne pas voir apparaître les flux informationnels. D'autres part, ce modèle est trop général et ne regroupe pas toutes les données et règles utiles à la construction du modèle informatique.

V.2-b) Travaux de recherche

La thèse de [PARIS 93] propose le développement d'un formalisme à partir des travaux de [BOURRIERES 90] et [BARAKAT 91]. Ceux-ci utilisent quatre symboles de base pour la réalisation d'un schéma opérationnel (figure 17).

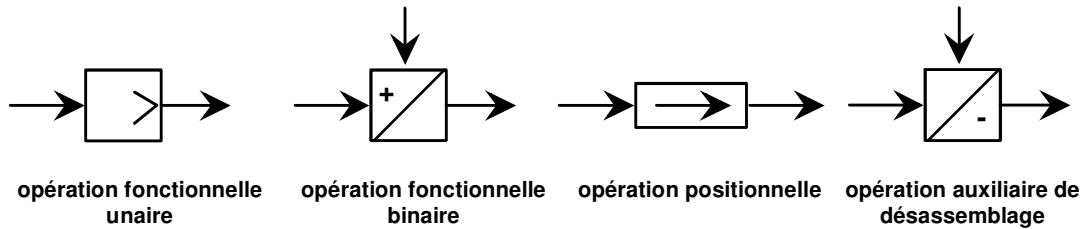


Figure 17 : Les quatre symboles de base de [BOURRIERES 90]

La figure 18 montre l'utilisation de ce formalisme pour représenter un processus d'assemblage.

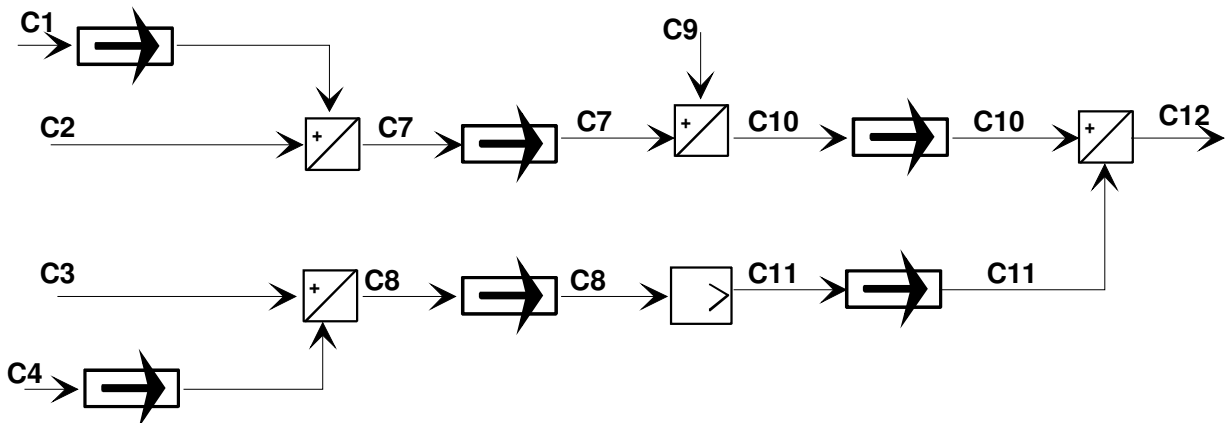


Figure 18 : Exemple d'un processus d'assemblage

A partir de ce formalisme, [PARIS 93] propose une extension destinée aux lignes d'assemblage automatisées, avec un système de convoyeurs. La figure 19 présente les principaux symboles de ce formalisme, et la figure 20 représente une application similaire à celle du système expérimental qui sert de support à nos travaux de recherche.

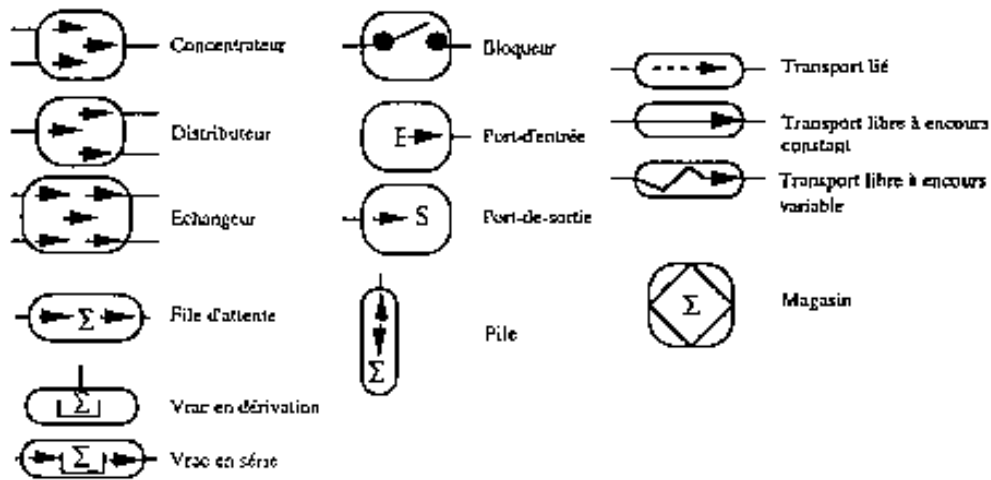


Figure 19 : Les symboles de base de [PARIS 93]

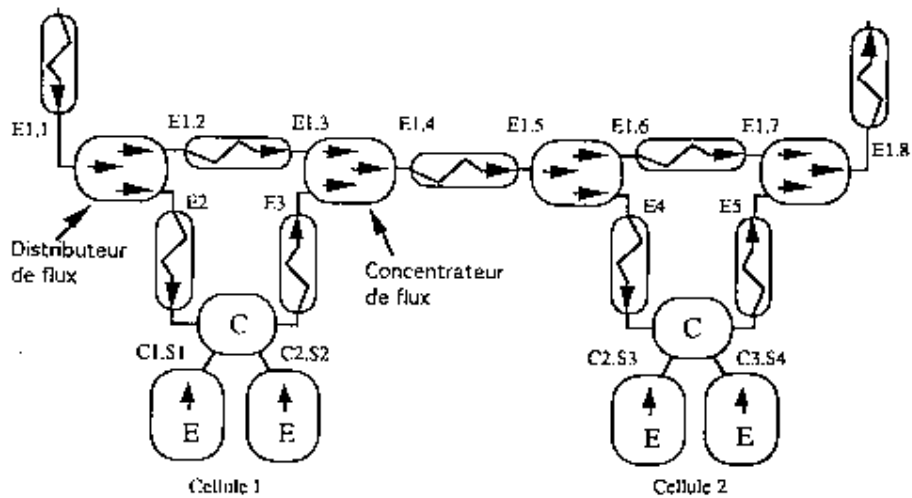


Figure 20 : Exemple de modélisation d'une ligne avec 2 postes en dérivation

Ce formalisme est d'une précision beaucoup plus importante que celui utilisé pour l'analyse de déroulement. D'autre part, il est d'une compréhension facile et d'une similitude assez proche avec les éléments réels qui composent le système de production. On dispose de beaucoup de détails sur le sous-système physique, mais par contre on n'a aucune information sur les sous-systèmes logiques et décisionnels.

Parmi les autres formalismes rencontrés dans la littérature, l'un d'eux est particulièrement prisé par les chercheurs. Il s'agit des Réseaux de Petri (RdP), pour lesquels les travaux et applications sont très nombreux. Dans le domaine de la production, l'ouvrage de [PROTH

95] présente les différentes utilisations des Réseaux de Pétri pour la conception et la gestion des systèmes de production.

Le formalisme des RdP utilise uniquement 3 symboles de base (figure 21) :

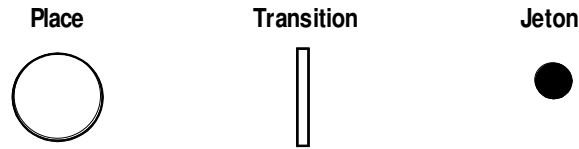


Figure 21 : Les trois symboles de base des RdP

Le fonctionnement des RdP est basé sur une dynamique qui autorise le franchissement des transitions sous certaines conditions de présence de jetons dans les places d'entrée.

La figure 22 présente une modélisation d'une ligne d'assemblage avec les RdP [PROTH 95].

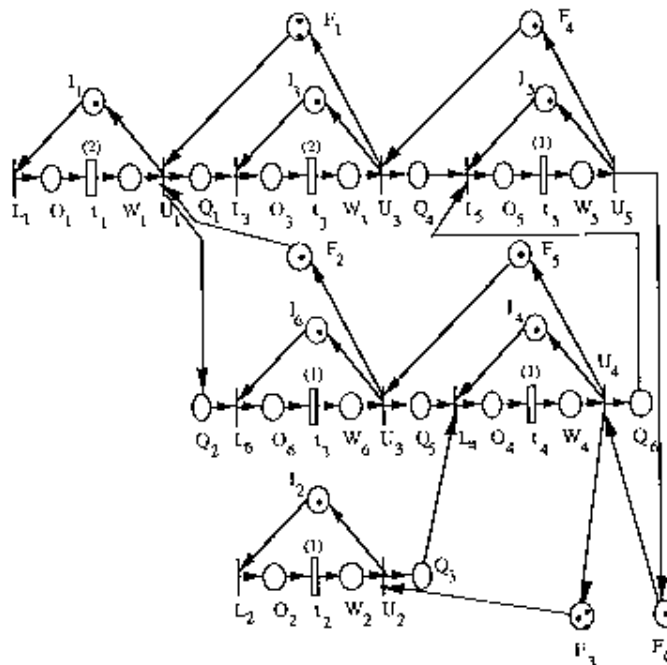


Figure 22 : Exemple de modélisation d'une ligne d'assemblage

La puissance de ce formalisme n'est plus à démontrer, notamment pour la description des règles de pilotage. Par rapport au problème que l'on s'est posé, le reproche que l'on peut lui faire est d'être d'un niveau d'abstraction trop important, notamment pour répondre à

l'objectif de communication entre les différents acteurs. D'autre part, un modèle en Réseaux de Pétri est encore trop éloigné du modèle informatique et si cette modélisation était utilisée, il faudrait sûrement encore une étape intermédiaire avant de passer au modèle d'action.

D'autres formalismes sont encore rencontrés pour la simulation de flux, et notamment les diagrammes d'action ou actigrammes [HAMAIDI 97]. Mais aucun ne répond à tous les objectifs que l'on s'est fixé dans le chapitre V.1.

Pour cette raison, nous allons présenter dans le chapitre suivant un formalisme issu de plusieurs travaux auxquels nous avons apporté des éléments permettant de l'utiliser comme Modèle Graphique de Connaissances.

□ V.3- Proposition d'un Modèle Graphique de Connaissances

Le livre de [de ROSNAY 75] et le formalisme proposé par [JULIEN 92] sont à la base de ce travail. On rappelle que l'objectif est de disposer d'un modèle graphique de connaissances servant d'interface entre le système réel et le modèle informatique.

L'un des apports de ce travail est d'assimiler les éléments utilisés dans une simulation de flux (voir chapitre I.3) à des objets génériques, disposant d'attributs et capables de réaliser des opérations (figure 23). Les trois objets de base que l'on utilisera dans les simulations à événements discrets sont les Machines, les Stocks et les Convoyeurs. Ces trois objets sont utilisés pour agir sur un autre objet qui correspond aux Articles.

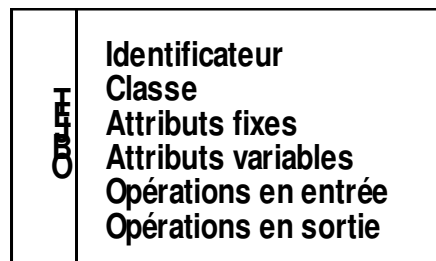
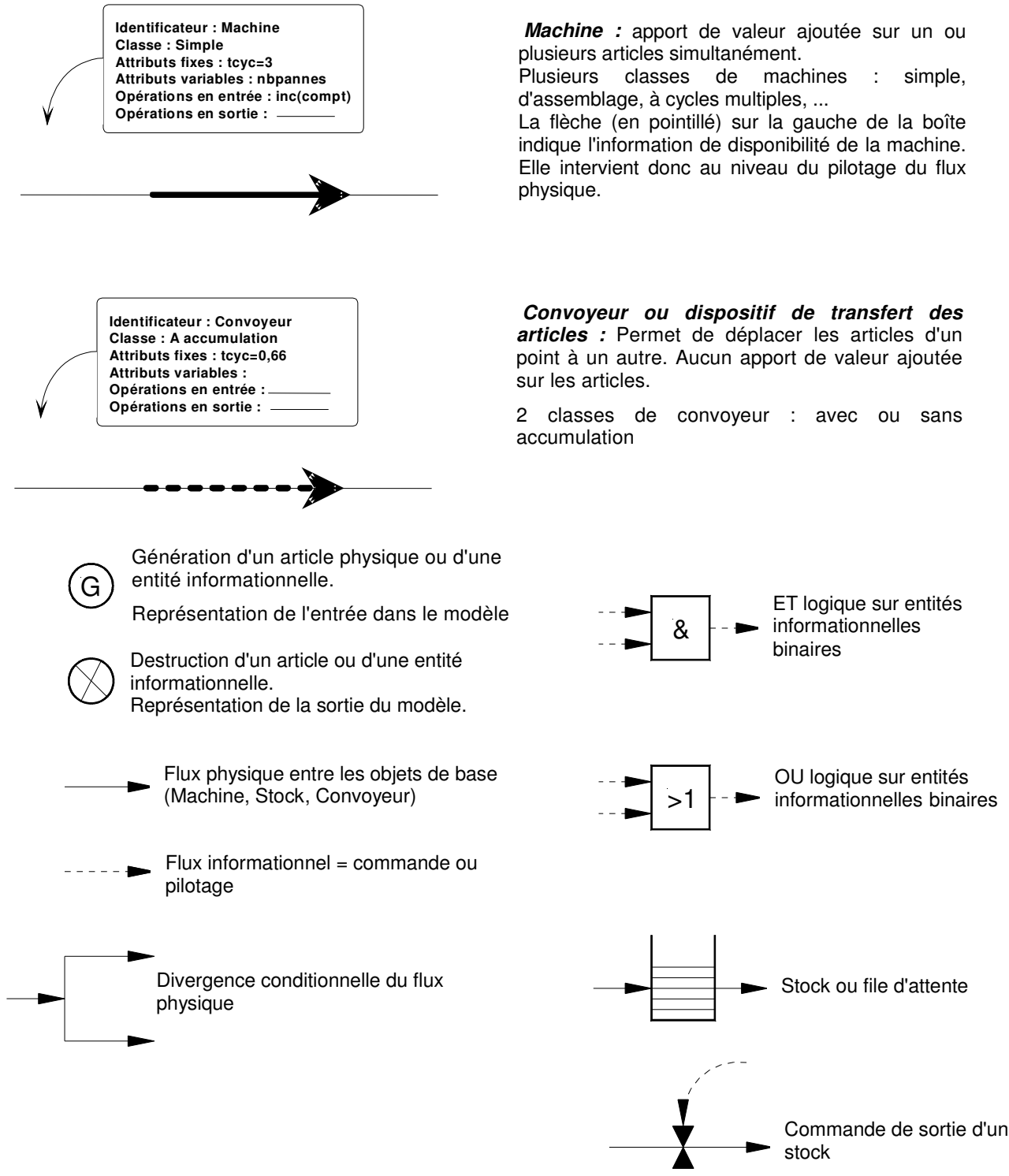


Figure 23 : Format des objets de base

Les symboles graphiques utilisés permettent non seulement de représenter les objets, mais aussi le flux physique et le flux informationnel (figure 24) :



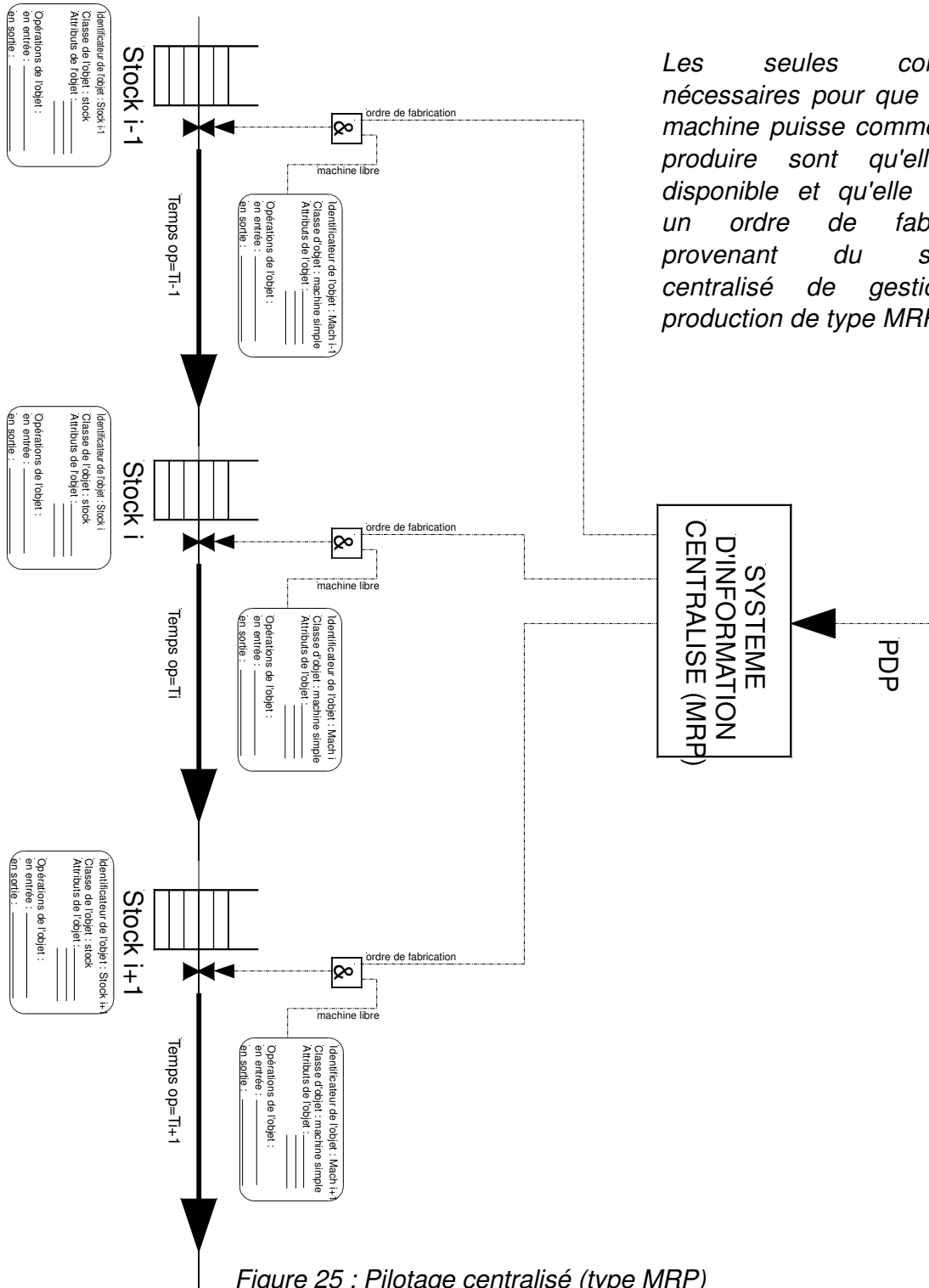
Machine : apport de valeur ajoutée sur un ou plusieurs articles simultanément. Plusieurs classes de machines : simple, d'assemblage, à cycles multiples, ... La flèche (en pointillé) sur la gauche de la boîte indique l'information de disponibilité de la machine. Elle intervient donc au niveau du pilotage du flux physique.

Convoyeur ou dispositif de transfert des articles : Permet de déplacer les articles d'un point à un autre. Aucun apport de valeur ajoutée sur les articles.

2 classes de convoyeur : avec ou sans accumulation

Figure 24 : Symboles graphiques utilisés

Les figures 25 à 26 présentent les Modèles Graphiques de Connaissances obtenus avec ce formalisme, appliqués aux modes de pilotage des flux (Partie 1 chapitre IV).



Les seules conditions nécessaires pour que chaque machine puisse commencer à produire sont qu'elle soit disponible et qu'elle reçoive un ordre de fabrication provenant du système centralisé de gestion de production de type MRP.

Figure 25 : Pilotage centralisé (type MRP)

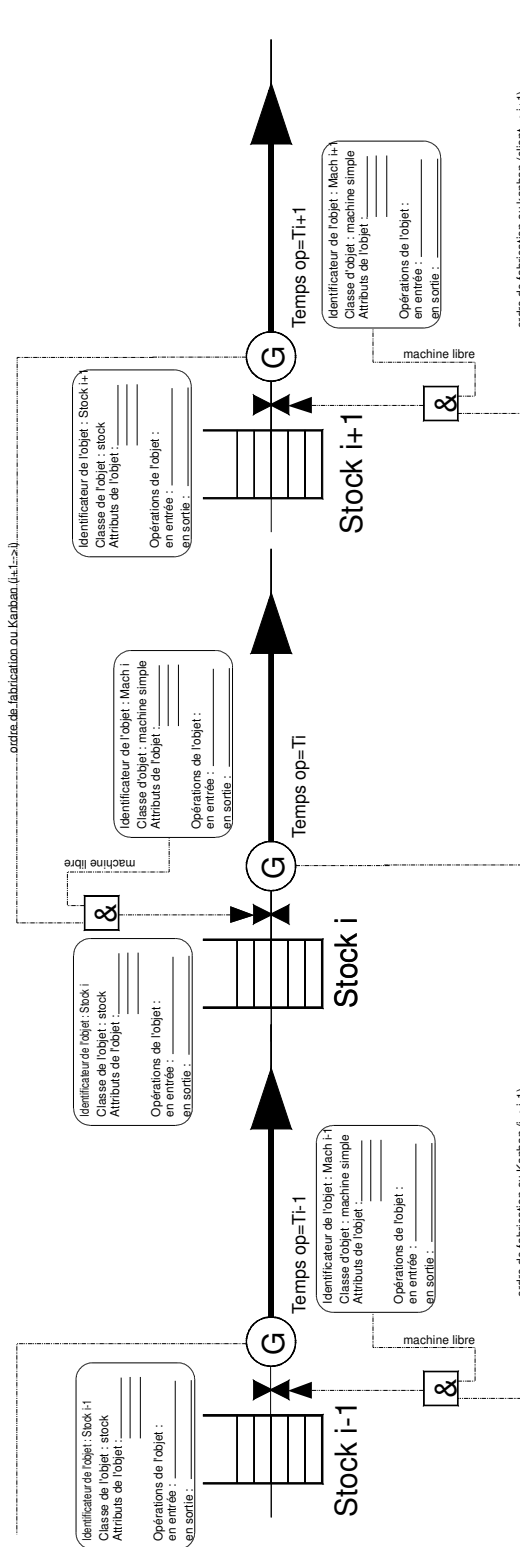


Figure 26 : Pilotage par l'aval et décentralisé (type Kanban)

Contrairement au pilotage centralisé, les conditions nécessaires pour qu'une machine travaille sont qu'elle soit disponible et que la machine située en aval ait "consommé" un article. La consommation d'un article par la machine "client" déclenche donc, sous la forme d'une étiquette appelé Kanban par les Japonais, un ordre de fabrication de la machine amont. Si la machine aval ("client") ne travaille pas, la machine "fournisseur" ne peut plus travailler, car elle ne reçoit plus d'ordre de fabrication, et cela même si elle est disponible. Ce mode de pilotage a principalement pour effet de limiter les encours entre chaque machine.

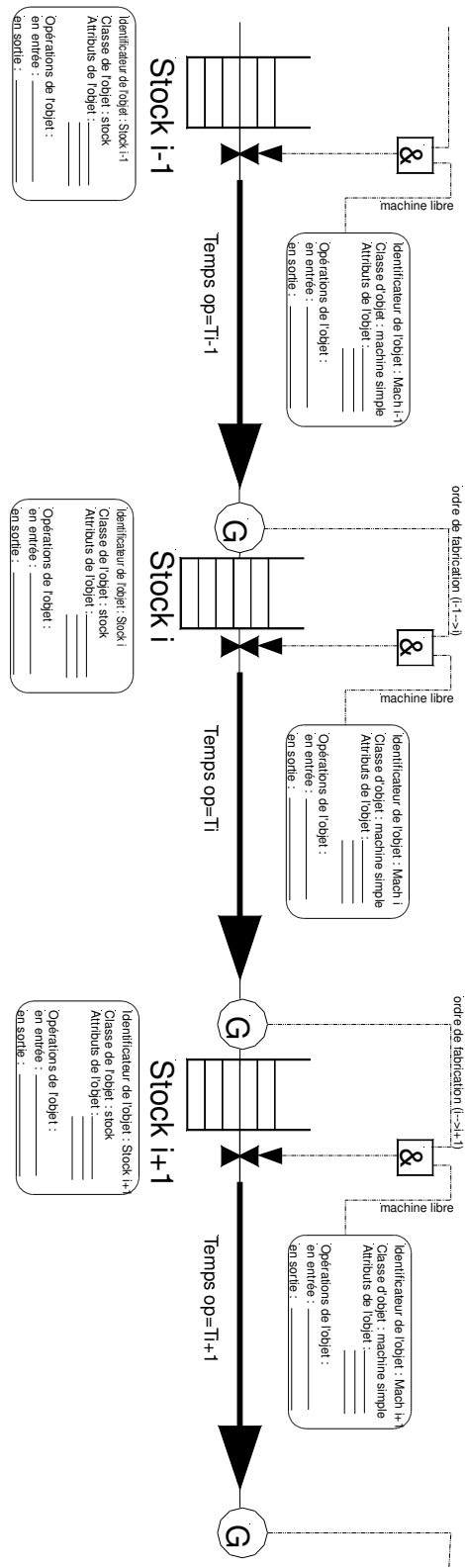


Figure 27 : Pilotage par l'amont et décentralisé

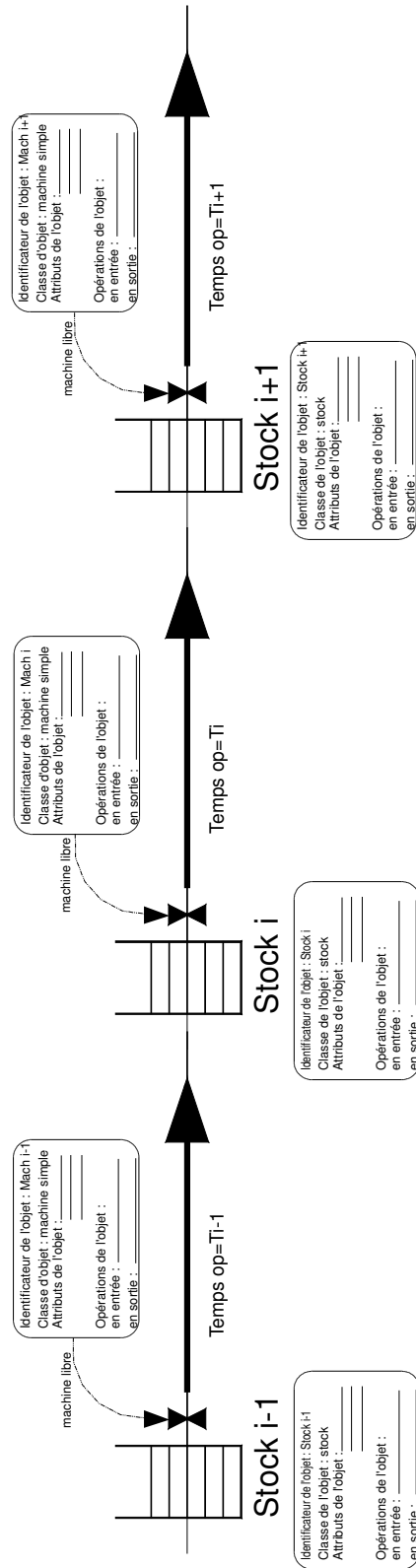


Figure 28 : Pilotage synchronisé

Comme on peut le voir sur ces différents exemples, le Modèle Graphique des Connaissances est très imagé par rapport au système réel. On retrouve, sur le même modèle, à la fois une modélisation du flux physique, du flux informationnel, des données utiles à la simulation et le modèle de conduite peut s'y superposer. Sa représentation est suffisamment simple et claire pour pouvoir servir de document de travail à tous les acteurs d'une simulation. La figure 30 montre un autre exemple de Modèle Graphique des Connaissances, le modèle géométrique du système physique est donné (figure 29) pour permettre de visualiser l'implantation des différents éléments.

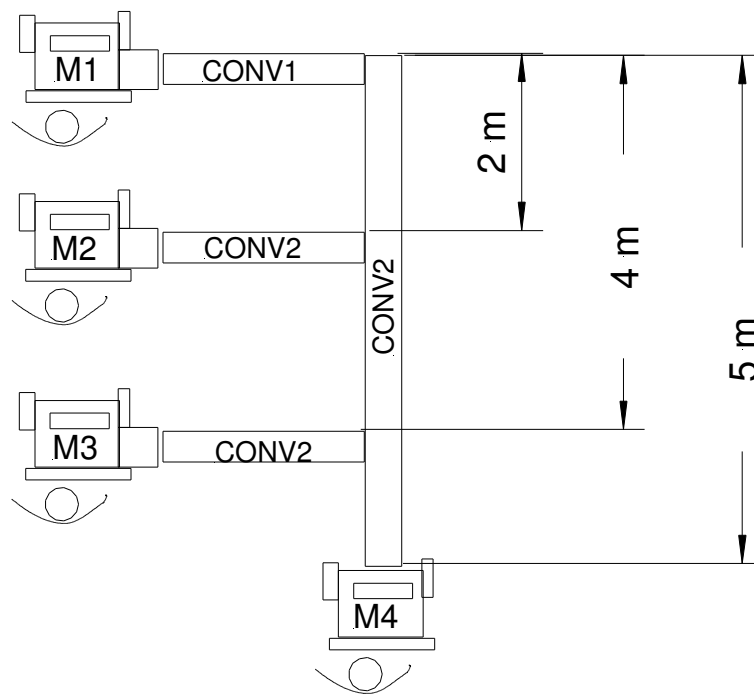


Figure 29 : Modèle géométrique (implantation des éléments)

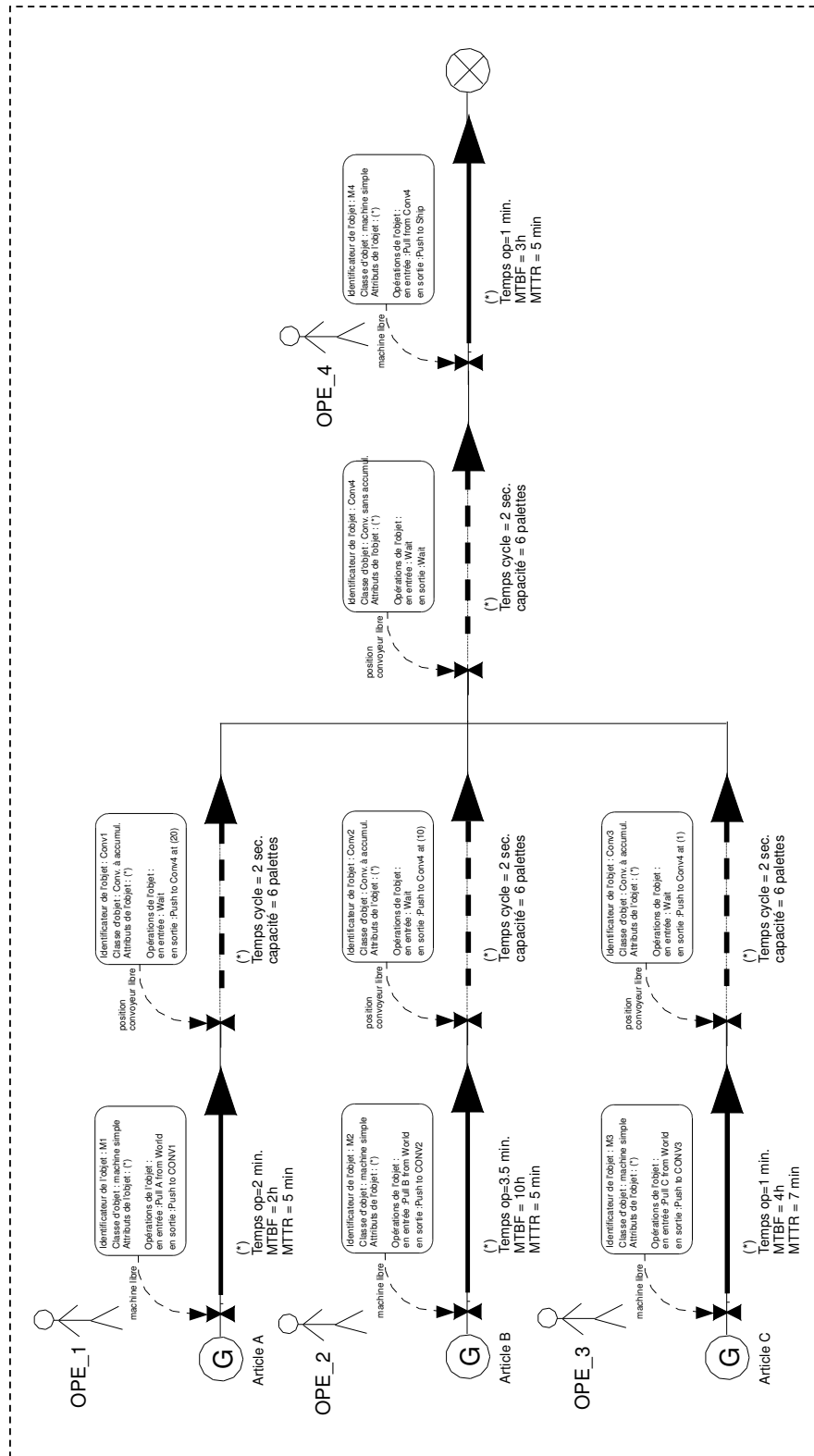


Figure 30 : Modèle Graphique de Connaissances

Le Modèle Graphique de Connaissances du système expérimental est présenté avec ce formalisme dans le chapitre suivant.

VI- Implémentation du modèle du système expérimental

□ VI.1- Modèle géométrique du système physique

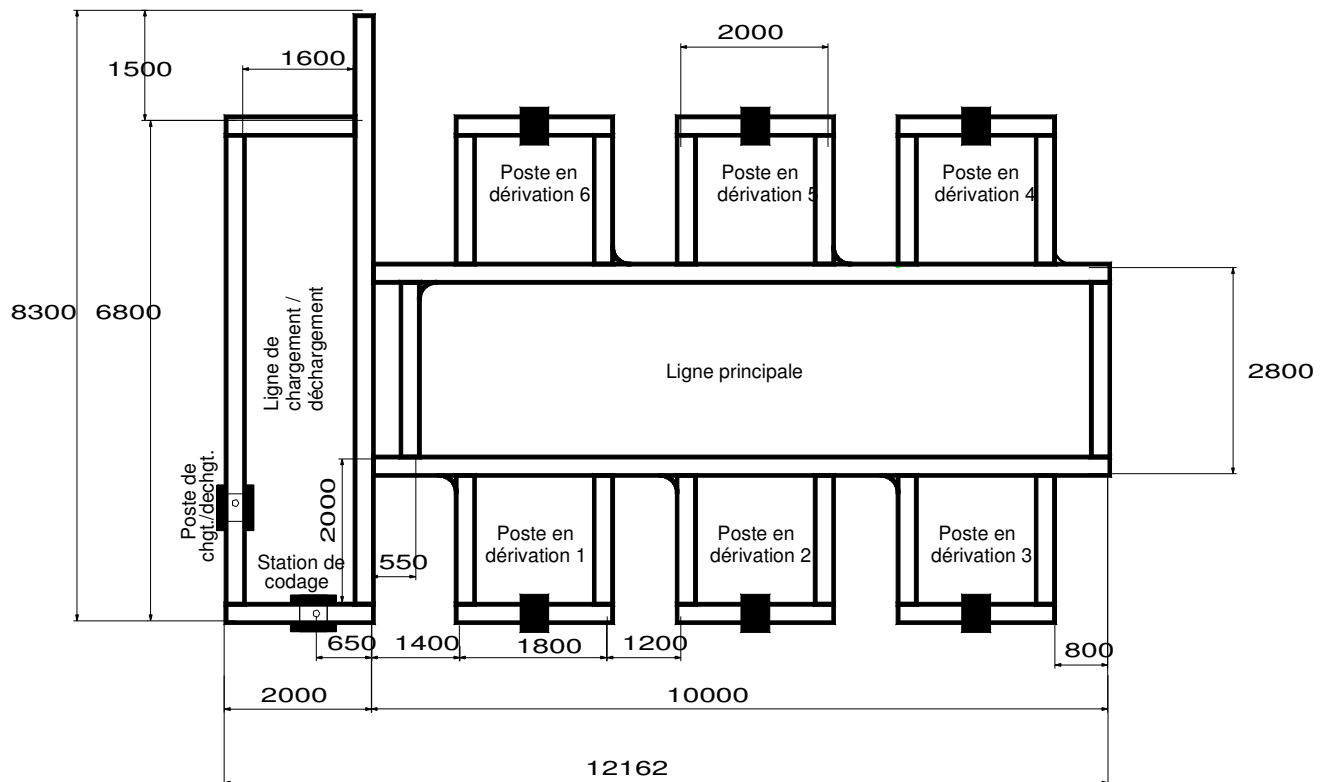


Figure 31 : Modèle géométrique du système expérimental

□ VI.2- Modèle Graphique de Connaissances

L'élément fonctionnel le plus significatif sur cette ligne expérimentale est le poste en dérivation, que l'on retrouve six fois à l'identique (voir partie 3 "Système expérimental"). C'est donc le Modèle Graphique de Connaissances d'un poste en dérivation qui va être détaillé, sachant que sa représentation est identique pour l'ensemble des postes en dérivation de la ligne. Avant de passer à cette modélisation, il est utile de représenter le modèle géométrique d'un poste en dérivation seul, pour préciser quelles en sont les limites physiques (figure 32).

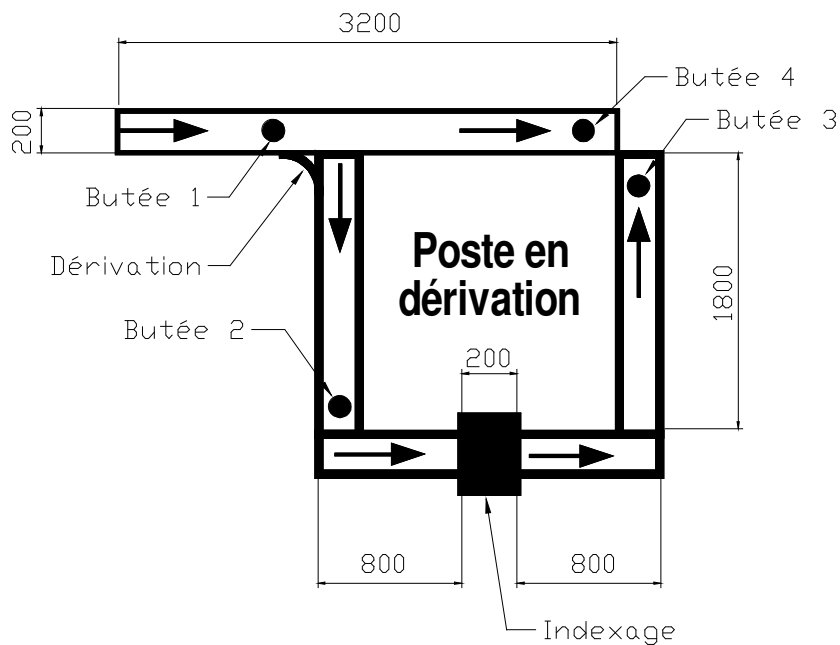


Figure 32 : Modèle géométrique d'un poste en dérivation

On rappelle que l'objectif du Modèle Graphique de Connaissances est de disposer d'un document sur lequel toutes les données nécessaires à la réalisation du modèle informatique de simulation sont réunies.

L'une des premières étapes de la réalisation de ce modèle est l'identification des objets. Cette étape nécessite une analyse du système réel qui va permettre au concepteur du modèle de "filtrer" les éléments physiques réels pour passer aux objets du modèle graphique. Il est donc important dans cette étape de décider quels sont les éléments qui

doivent être modélisés, et avec quel niveau de précision. Malgré tout, comme cela a déjà été montré, ce processus de réalisation du modèle peut être itératif, avec d'abord un dégrossissage, puis un affinage progressif de la modélisation.

Pour le modèle d'un poste en dérivation, les objets identifiés lors de cette étape sont les suivants :

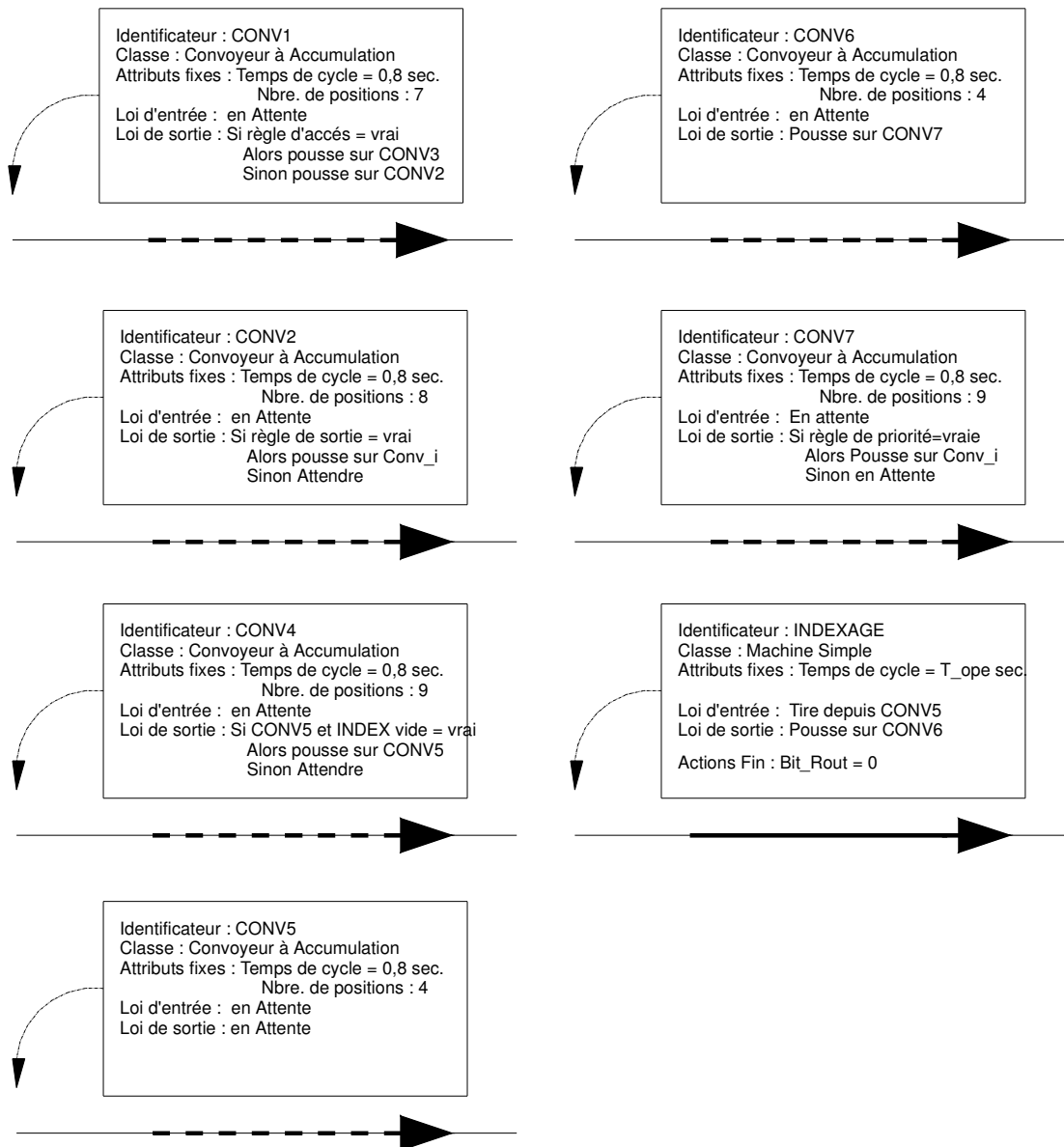


Figure 33 : Objets du modèle de connaissance d'un poste en dérivation

Les objets de base qui sont retenus pour la réalisation du modèle sont essentiellement des convoyeurs à accumulation et une machine (pour modéliser l'indexage). Une des simplifications de cette étape est de ne pas modéliser les butées parce qu'elles ne semblent pas intervenir de façon significative dans le fonctionnement général d'un poste en dérivation. Si l'on avait choisi de les modéliser, on aurait utilisé un objet de classe "Machine" pour le faire. Autre simplification, le tronçon de convoyeur de la ligne principale (long. 3200) est divisé en deux convoyeurs CONV1 et CONV2, car les lois d'entrée et de sortie sont actives uniquement aux extrémités des convoyeurs.

La figure 34 présente le Modèle Graphique de Connaissances d'un poste en dérivation.

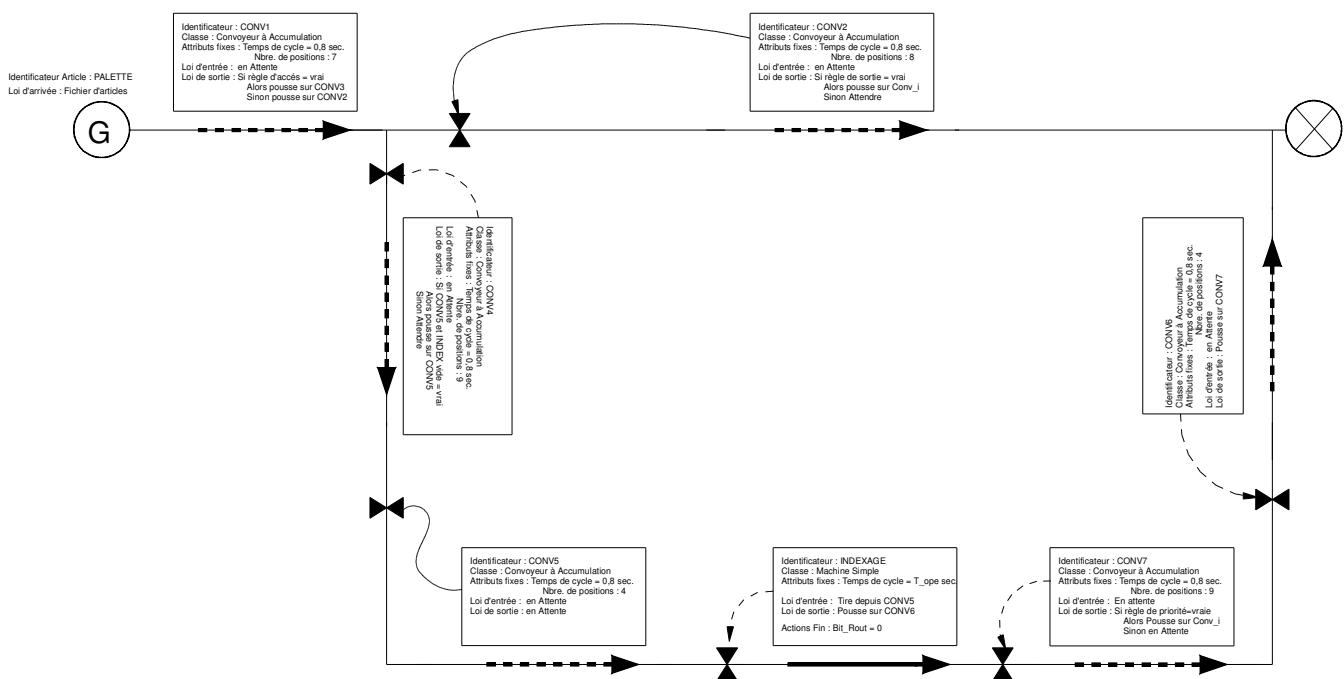


Figure 34 : Modèle Graphique de Connaissances d'un poste en dérivation

□ VI.3- Modèle d'action

Le codage du modèle d'action est réalisé avec le progiciel WITNESS 9™ [LANNER 98]. Le principal intérêt de ce progiciel est de disposer d'une interface graphique de programmation et de visualisation dynamique (voir les figures 35 et 36). D'autre part, il a été choisi parce qu'il peut échanger des données de type objet (au sens de la programmation orientée objet) par le biais des liens O.L.E Automation de Microsoft (Object Linked Embedding) avec d'autres applications fonctionnant sous Windows 95™.

Cette fonctionnalité est utilisée principalement dans la partie "Choix et intégration d'un outil d'optimisation" (voir Partie 8) .

L'implémentation du modèle informatique est obtenu avec un minimum de programmation, mais à partir de la déclaration et de la description sous forme interactive des différents objets du modèle. Toutefois, seule la programmation des règles du sous-modèles de pilotage nécessite l'utilisation d'un langage algorithmique interprété avec des primitives habituelles (IF..THEN..ELSE, FOR..TO..NEXT, WHILE.., etc...). Cette programmation est réalisée par des lois ou des actions d'entrée et de sortie au niveau des objets de type Machine, Convoyeur et Article. Les lois permettant de décrire le flux et les actions sont utilisées pour modifier ou accéder à des valeurs de variables logiques prédéfinies ou définies par l'utilisateur.

Malgré son interface conversationnelle, Witness génère automatiquement un code en langage WCL (Witness Command Language) qui permet d'avoir une trace écrite de la construction du modèle informatique. Une partie de ce code correspondant au Modèle Graphique de Connaissances est présentée sur la figure 37.

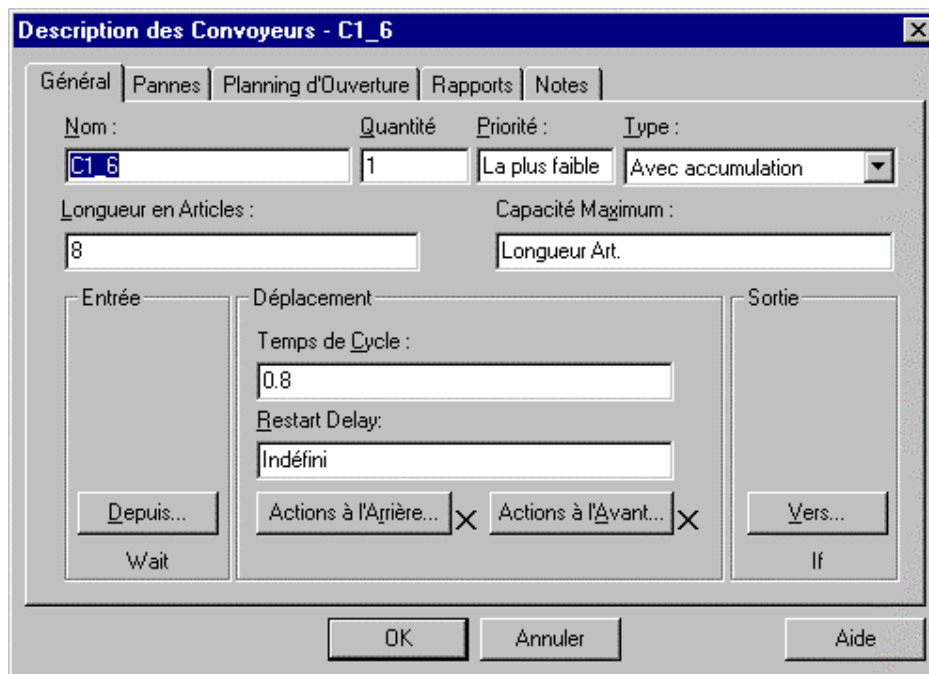


Figure 35 : Exemple d'une fenêtre de description d'un convoyeur

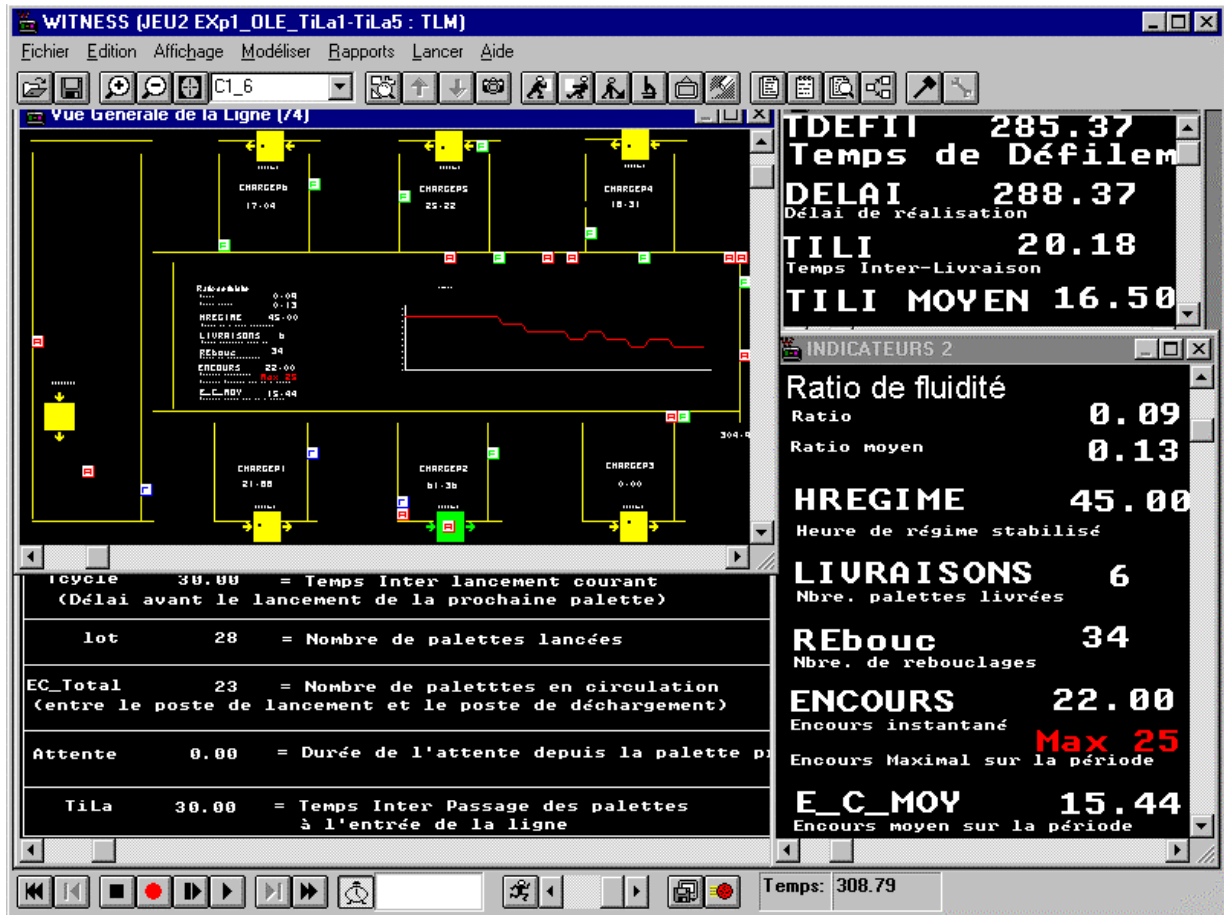


Figure 36 : Visualisation graphique et dynamique de la simulation

<p>NOM DE LA MACHINE: INDEX_1;</p> <p>TYPE: Simple;</p> <p>DESCRIPTION_REGLAGE</p> <p>Réglage n°: 1</p> <p>* Mode: Changement Valeur;</p> <p>* Expression: ICON;</p> <p>* Temps Réglage: 0.0;</p> <p>* Description: Réglage 1</p> <p>* Poste n°: 1;</p> <p>ACTIONS, Début</p> <p>Début</p> <p>Sreg1 = Sreg1 + 1</p> <p>Fin Actions</p> <p>FIN DESCRIPTION_REGLAGE</p> <p>TEMPS DE CYCLE: T_OPE (1);</p> <p>ACTIONS, Fin</p> <p>Début</p> <p>bit_rout (1) = 0</p> <p>!calcul charge</p> <p>CHARGE P1 = PUTIL (INDEX_1,2)</p> <p>!calcul de la charge totale en sec.</p> <p>STCYC1 = STCYC1 + T_OPE (1)</p> <p>Fin Actions</p> <p>LOI D'ENTREE: PULL from C5_1 at Front;</p> <p>LOI DE SORTIE: PUSH to C6_1 at Rear;</p> <p>FIN INDEX_1</p>	<p>NOM DU CONVOYEUR: C1_1;</p> <p>TYPE: A_accumulation;</p> <p>LONGUEUR ART.: 4;</p> <p>CAPACITE MAX: 4;</p> <p>LOI D'ENTREE: Wait;</p> <p>LOI DE SORTIE: IF bit_rout (1) = 1 AND bit_def = 1 AND bit_ante = 1 AND NPARTS (C4_1) < P1_AM</p> <p>PUSH to C4_1 at Rear</p> <p>ELSE</p> <p>PUSH to C2_1 at Rear</p> <p>ENDIF;</p> <p>TEMPS DE CYCLE: 0.8;</p> <p>PANNES: Aucune;</p> <p>FIN C1_1</p>
--	--

Figure 37 : Description d'objets en langage WCL

Conclusion

Cette partie a permis de présenter la simulation de flux, outil informatique qui va être utilisé pour analyser et optimiser le fonctionnement de la ligne expérimentale. En effet, plutôt que de mener des expériences longues et coûteuses sur la ligne réelle, la simulation permet de réaliser les mêmes expériences, mais de façon virtuelle, sur un ordinateur, sans aucun risque et avec des temps de réponse très rapides. Cet outil nous permettra, dans les prochaines parties de ce travail, de tester un très grand nombre de paramètres et de valeurs de ces paramètres en un minimum de temps.

Bien sûr, avant de lancer des simulations, il est nécessaire de construire un modèle de simulation du système étudié. Dans cette partie, deux contributions sont apportées :

- La première est d'ordre méthodologique, et permet d'organiser en quatre macro-étapes, elles-mêmes décomposées en étapes, la réalisation d'une étude de simulation. Cette démarche est présentée de façon graphique sous la forme d'un organigramme.
- La deuxième est d'ordre analytique, puisqu'elle a pour but de proposer la construction d'un Modèle Graphique de Connaissances, préliminaire à la construction du modèle d'action, qui est lui directement lié au progiciel utilisé. L'objectif du Modèle Graphique de Connaissances est de réaliser une modélisation préalable en faisant apparaître les différentes données, de nature numérique ou logique, ainsi que la décomposition en objets de simulation : machines, convoyeurs, articles, stocks, etc... Ceci permet de créer un document de synthèse sur lequel seront rassemblées toutes les données utiles à la construction du modèle d'action. L'expérience prouve en effet que la construction d'un modèle nécessite la collecte et la synthèse d'informations issues de multiples sources. Le Modèle Graphique de Connaissances est donc un outil de synthèse et de communication entre les différents acteurs d'une étude de simulation.

Dans le chapitre VI, on présente la réalisation pratique du Modèle Graphique de Connaissances de la ligne expérimentale et du modèle d'action réalisé avec le progiciel WITNESS. Ce modèle d'action a été construit, puis vérifié et validé en le comparant au système réel. Dans la partie 5, nous allons donc définir les simulations qui seront réalisées avec ce modèle.

Bibliographie de la partie 4 : "Modélisation et simulation de flux"

[LAW 91] : "Simulation Modeling and Analysis"

Editions McGraw-Hill, 2nd edition, 1991

A.M. Law, W.D. Kelton

[CLAVER 97] : "Gestion de flux en entreprise"

Editions Hermes, 1997

J.F. Claver, J. Gélinier, D. Pitt

[CETIM 89] : "La simulation de flux"

Editions du CETIM, 1989

[BALLOT 97] : "La simulation industrielle : aide réelle ou virtuelle à la prise de décision ?"

Revue Française de Gestion Industrielle, Vol. 16, N°1, 1997, pages 21 à 38

E. Ballot

[DUFRENE 94] : "A discrete event simulation methodology for the design of flexible manufacturing systems"

Summer Computer Simulation Conference, San Diego, California, July 18-20 1994

D. Dufrene, A. Gharbi, J.P. Kieffer, J.Y. Rebouché, L. Villeneuve

[CERNAULT 88] : "La simulation des systèmes de production"

Editions Cepadues, 1988

A. Cernault

[DRAKE 98] ; "Simulation as a planning and scheduling tool for flexible manufacturing systems"

On Line Ed. <http://tamcam.tamu.edu/pubs/ws95f.htm>

G.R. Drake, J.S. Smith, B.A. Peters

[HAMAIIDI 97] : "Modélisation du flux de production dans les procédés batch"

Actes du 2ème congrès franco-québécois de Génie Industriel, Albi 1997

L. Hamaidi, F. Vernada

[BOSTEL 97] : "Un modèle de simulation pour l'ordonnement des mouvements des équipements dans un chantier de transbordement rapide fer-fer"

Actes du 2ème congrès franco-québécois de Génie Industriel, Albi 1997

N. Bostel, P. Dejax

[QUERE 97] : "Modélisation multi-agents et ordonnancement dynamique d'un convoyeur à entrées / sorties multiples"

Actes du 2ème congrès franco-québécois de Génie Industriel, Albi 1997

A. Quéré, M. Belkhiter, M. Gourgand

[DELAMARE 93] : "Utilisation combinée des techniques de simulation de processus et des plans d'expériences pour améliorer l'organisation des ateliers"

RAPA, Vol. 6, N°2, 1993, pages 185 à 207

A. Delamare, J. Ciccotelli, P. Martin, P. Charpentier

[VERNADAT 96] : ""Enterprise modeling and integration : principles and applications"

Editions Chapman & Hall, London, 1996

F. Vernadat

[PARIS 93] : "Modélisation et simulation orientées objets pour la conception et l'exploitation d'ateliers flexibles d'assemblage automatique réactifs aux aléas de qualité"

Thèse de l'UFR de sciences et techniques de l'Université de Franche-Comté, 1993

J.L. Paris

[BARAKAT 91] : "Contribution à la modélisation et à la simulation orientées objet des systèmes flexibles de production"

Thèse de l'UFC de Besançon, 1991

O. Barakat

[BOURRIERES 90] : "Contribution à la modélisation intégrée des systèmes robotisés d'assemblage"

Thèse d'état de l'UFC de Besançon, 1990

J.P. Bourrières

[PROTH 95] : "Les réseaux de Petri pour la conception et la gestion des systèmes de production"

Editions Masson, 1995

J.M. Proth, X. Xie

[de ROSNAY 75] : "Le Macroscopie"

Editions du Seuil

J. de Rosnay

[JULIEN 92] : "Un modèle conceptuel pour la simulation orientée processus"

Congrès FRANCOSIM 1992

B. Julien

[PERES 97] : "La maintenance dans l'évaluation des performances des systèmes de production : une approche par la simulation"

Actes du 2ème congrès franco-québécois de Génie Industriel, Albi 1997

F. Pérès, E. Caillaud, D. Noyes

[CAUX 95] : "Couplage méthodes d'ordonnancement-simulation pour l'ordonnancement de systèmes industriels de traitement de surfaces"

Recherche Opérationnelle, 0399-0559, 04-95

C. Caux, G. Fleury, M. Gourgand, P. Kellert

[GOUJON 96] : "Planification pour les fonderies de type sable : une méthode combinant optimisation stochastique et simulation"

RAIRO-APII-JESA Vol. 30 N°6/1996 pages 963 à 987

J.Y. Goujon, P. Lacomme

[BOUKACHOUR] : "Simulation arrière pour l'ordonnancement d'atelier"

RAIRO-APII-JESA Vol. 30 N°7/1996, pages 989 à 1003

J. Boukachour, T. Galinho, J.P. Pécuchet

[BAVAY 96] : "Rétro-simulation pour l'ordonnancement coordonné de l'ensemble des ateliers d'une usine"

RAIRO-APII-JESA Vol. 30 N°7 / 1996, pages 1005 à 1020

R. Bavay, G. Bel, M. Corrège

[CAPRIHAN 97] : "Impact of routing flexibility on the performance of an FMS - A simulation study"

International Journal of Flexible Manufacturing Systems, 9, 1997, pages 299 à 300

R. Caprihan, S. Wadhwa

[KIEFFER 98] : "Amélioration de la réactivité en gestion d'atelier par la simulation discrète dans un contexte de coopération entre donneur d'ordres et sous-traitant"

Revue Française de Gestion Industrielle, Vol. 17, N°1, 1998, pages 31 à 47

J.P. Kieffer, J. Arékion, K. Ayadi, D. Baudoin

[LANNER 98] : "Manuel d'utilisation du logiciel Witness version 9"

Distribué par SERETE, 1998

Lanner Group

- PARTIE 5 -
***Elaboration du plan expérimental
sur le modèle de simulation***

Sommaire

Introduction	142
I- Hypothèses générales	144
II- Hypothèses propres.....	145
<input type="checkbox"/> II.1- 1 ^{ère} hypothèse : Variété des PRODUITS lancés	145
<input type="checkbox"/> II.2- 2 ^{ème} hypothèse : HORIZON de FABRICATION	146
<input type="checkbox"/> II.3- 3 ^{ème} hypothèse : SEQUENCEMENT du LANCEMENT	147
<input type="checkbox"/> II.4- 4 ^{ème} hypothèse : SEQUENCEMENT des OPERATIONS de fabrication	149
<input type="checkbox"/> II.5- 5 ^{ème} hypothèse : NOMBRE de PALETTES DISPONIBLES.....	149
III- Elaboration du plan expérimental complet.....	151
<input type="checkbox"/> III.1- Définition du plan expérimental.....	151
<input type="checkbox"/> III.2- Récapitulation de la typologie des hypothèses propres.....	151
<input type="checkbox"/> III.3- Structuration du plan et nombre d'expérimentations.....	152
IV- Variables de pilotage de la ligne.....	153
<input type="checkbox"/> IV.1- 1 ^{ère} variable : CADENCE de LANCEMENT	154
<input type="checkbox"/> IV.2- 2 ^{ème} variable : CAPACITE des STOCKS statiques AMONT.....	155
<i>IV.2-a) Capacité identique sur tous les postes</i>	<i>156</i>
<i>IV.2-b) Capacité différente d'un poste à l'autre</i>	<i>156</i>
<input type="checkbox"/> IV.3- 3 ^{ème} variable : CAPACITE des STOCKS statiques AVAL	156
<i>IV.3-a) Capacité identique sur tous les postes</i>	<i>156</i>
<i>IV.3-b) Capacité différente d'un poste à l'autre</i>	<i>156</i>
<input type="checkbox"/> IV.4- 4 ^{ème} variable : REGLE de PRIORITE en SORTIE de POSTE	157
<input type="checkbox"/> IV.5- 5 ^{ème} variable : NOMBRE de PALETTES à UTILISER.....	160
<input type="checkbox"/> IV.6- Niveau d'influence estimé sur les variables de sortie et valeurs :	160
V- Sélection des expérimentations et des variables d'entrée	161
<input type="checkbox"/> V.1- Expérimentations retenues	161
<i>V.1-a) Jeu N°1 (Lancement MONOPRODUIT)</i>	<i>162</i>
<i>V.1-b) Jeu N°2 (Lancement MULTIPRODUIT).....</i>	<i>163</i>
<input type="checkbox"/> V.2- Variables de pilotage testées à chaque expérimentation	166

VI- Mesures des variables de sortie (ou indicateurs de performance) réalisées à chaque simulation	167
<input type="checkbox"/> VI.1- Temps de défilement de chaque palette Td (en sec)	167
<input type="checkbox"/> VI.2- Délai de Fabrication d'une campagne D (en sec)	168
<input type="checkbox"/> VI.3- Nombre de réglages par poste de travail Nr(i)	168
<input type="checkbox"/> VI.4- Ratio de fluidité de chaque palette Rf	168
<input type="checkbox"/> VI.5- Encours de palettes sur la ligne entre le lancement et le déchargement Ec.	168
Conclusion	170

Introduction

Dans cette partie, nous allons définir le plan des expérimentations effectuées sur le modèle de simulation de flux de la ligne étudiée. Nous identifierons aussi les variables d'entrée ou variables de pilotage de la ligne, c'est à dire les paramètres sur lesquels une variation peut entraîner une modification du comportement dynamique. Ce comportement dynamique est mesuré à partir de variables de sortie, ou indicateurs de performance, qui caractérisent essentiellement le flux physique.

Ces expériences ont donc pour objectif de recueillir des données sur le fonctionnement de la ligne de production expérimentale, dans plusieurs situations différentes. Dans la partie 6 "Analyse des résultats des simulations", ces données seront analysées afin d'identifier quelles sont les variables d'entrée qui produisent un effet sur les variables de sorties et quelles valeurs les variables d'entrée doivent prendre pour minimiser ou maximiser les variables de sortie. Nous proposerons ensuite une optimisation à base de règles logiques. Dans la partie 7 et la partie 8, nous réaliserons encore des expériences en simulation, à partir des mêmes hypothèses et avec les mêmes variables, pour optimiser le fonctionnement à l'aide de méthodes algorithmiques.

Les hypothèses expérimentales peuvent être classées en deux catégories :

- les hypothèses générales, valables pour toutes les expériences, qui sont présentées dans le chapitre 1;
- les hypothèses propres à chaque expérience, dont notamment l'hypothèse de variété des produits lancés, qui sont présentées dans le chapitre 2;

A partir de ces hypothèses, le plan d'expérimentations complet est établi dans le chapitre 3, à partir de la combinaison de ces hypothèses. L'hypothèse de variété des produits est choisie pour classer les expériences en deux jeux d'essais :

- le jeu N°1, qui comporte des expériences avec une hypothèse de lancement de type "monoproduit";
- le jeu N°2 qui comporte des expériences avec une hypothèse de lancement de type "Multiproduit";

Dans le chapitre 4, on identifie les variables de pilotage de la ligne et on définit leurs niveaux de variation. Certaines de ces variables sont quantitatives et discrètes, d'autres

sont qualitatives. Dans un premier temps, on tente d'estimer grossièrement et intuitivement leur influence sur les variables de sortie.

Dans le chapitre 5, on sélectionne, à partir du plan d'expérimentations complet, les expériences qui seront réalisées. En effet, le grand nombre d'hypothèses et de variables de pilotage ne permet pas de réaliser toutes les combinaisons. On choisit donc celles qui nous semblent être les plus représentatives. On classe les expérimentations en deux jeux d'essais. Pour chaque expérimentation, on précise les variables d'entrée ainsi que leurs valeurs et on en déduit le nombre de simulations à lancer.

Les données recueillies pendant et après chaque simulation sont répertoriées en variables de sortie (indicateurs de performances). Elles sont présentées dans le chapitre 6.

I- Hypothèses générales

Les hypothèses suivantes sont vraies pour toutes les expériences réalisées :

- Chaque poste de travail ne peut être que dans deux états : disponible ou occupé. On considère donc qu'aucune panne ou perturbation ne peut survenir. Cette hypothèse peut paraître réductrice, mais la réalité industrielle montre que sur ce genre de système, les temps inter-pannes (MTBF) sont assez importants et que la durée des pannes (MTTR) est en général très courte. De plus, en cas de panne, la structure physique et informationnelle de tels systèmes fait que les effets peuvent être considérablement atténués, simplement en "basculant" la production d'un poste en panne sur un autre poste. C'est un des principes de la flexibilité.
- Pas de préemption : chaque opération commencée sur un poste de travail est terminée sans interruption.
- Chaque poste de travail ne peut traiter qu'une palette à la fois.
- Les temps opératoires de chaque poste de travail peuvent être déterministes (aucune variation) ou stochastiques (suivant une loi de distribution aléatoire). Le premier cas correspond aux postes de travail dont l'opération est réalisée par une machine automatique, par exemple un robot, alors que le second cas correspond aux postes de travail manuels, pour lesquels des variations de la durée des opérations peuvent ne pas être négligeables.
- Les Temps de réglage ne sont pas pris en compte, mais on relèvera par contre le nombre de réglages par poste de travail. On entend par réglage les opérations qui permettent de passer d'un produit à un autre : changement d'outillage, changement de posture, changement de programme, etc... Cette hypothèse n'est prise en compte que pour les lancements de type "Multiproduit".
- Chaque produit (palette), une fois qu'il est lancé, ne peut pas être livré tant que toutes les opérations n'ont pas été achevées. Aucun ordre d'annulation n'est possible.
- Chaque produit (palette) ne peut être traité que par un poste de travail à la fois. Les produits sont assemblés suivant le principe de la Pièce Maîtresse (voir Partie 2, chap. 1.3) qui fait que les composants sont ajoutés au produit sur un poste de travail à la fois.
- Aucune date de livraison n'est exigée. On ne cherche pas à rentrer dans un délai imposé, mais plutôt à déterminer un délai que l'on pourra garantir.
- La disponibilité des palettes est limitée par le nombre de palettes en réserve.
- Chaque palette peut être utilisée indifféremment pour n'importe quel type de produit (il n'y a pas de palette dédiée à un produit). En fait, chaque palette possède un posage qui sert d'interface de liaison avec le produit. Ce posage est spécifique

à un produit, et le temps de changement de posage, si cela est nécessaire, est inclus dans le temps de déchargement des produits finis.

- Le lancement d'un produit est conditionné par la disponibilité d'une palette et par le Temps Inter-Lancement.

II- Hypothèses propres

Ces hypothèses sont rangées en 5 classes :

- Hypothèse de variété des produits lancés;
- Hypothèse d'horizon de fabrication;
- Hypothèse de séquençement de lancement;
- Hypothèse de séquençement des opérations;
- Hypothèse de réserve de palettes disponibles.

□ II.1- 1^{ère} hypothèse : Variété des PRODUITS lancés

II.1-a) Lancement de type "MONOPRODUIT" : Un seul Ordre de Fabrication (O.F) est lancé correspondant à une seule référence de produit à fabriquer, en quantité quelconque, finie ou infinie (voir hypothèse "Horizon de fabrication").

Exemple : Lancement d'un lot de 160 produits de référence BTMEC03

Dans ce cas, on peut considérer que la ligne est dédiée à un seul type produit.

II.1-b) Lancement de type "MULTIPRODUIT" (Mixed models): Plusieurs ordres de fabrication sont lancés séquentiellement (voir hypothèse "Séquençement") correspondant chacun à des références de produits différents, avec des quantités quelconques.

□ II.2- 2ème hypothèse : HORIZON de FABRICATION

II.2-a) Horizon de fabrication fini : On considère que le lancement d'une campagne de production est réalisé uniquement quand la campagne précédente est complètement terminée (livrée), ce qui signifie que la ligne de production est vide au moment du lancement. On obtient ainsi un fonctionnement cyclique dont la période est constituée d'une phase de remplissage ou de montée en régime (appelée aussi "période de chauffe dans certains cas), d'une phase de fonctionnement à régime nominal ou stabilisé, et d'une phase de vidage ou fin de production (voir figure 1). La période peut par exemple correspondre à une journée de travail : on démarre le matin avec une ligne vide et on termine le soir en vidant la ligne.

C'est en général le niveau des encours ou le flux sortant de chaque référence qui permet

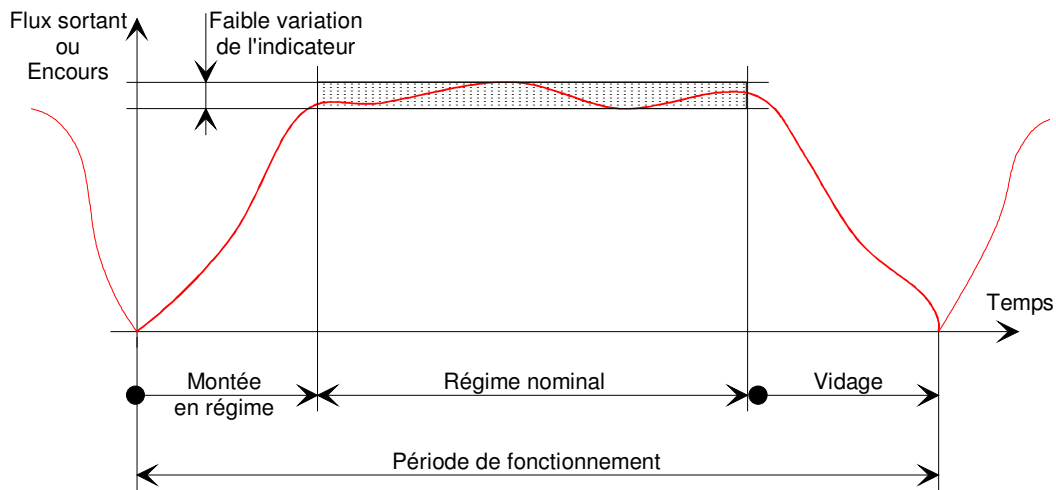


Figure 1 : Les différents régimes de fonctionnement

de déterminer dans quelle phase se trouve l'atelier.

II.2-b) Horizon de fabrication infini : Contrairement au type précédent, le flux de production va être ininterrompu : le lancement des produits se fait de façon continue, sans attendre le vidage de la ligne.

Si l'on est dans le cas d'un lancement de type "Monoproduit", cela signifie que la quantité à fabriquer est infinie.

Si l'on est dans le cas d'un lancement de type "Multiproduit", les tailles de lot de chaque OF sont finies, mais le nombre d'Ordres de Fabrication est infini. Dans ce cas, une autre classification est possible : Fabrications **REPETITIVES** ou **NON REPETITIVES**. Dans le premier cas, on fabrique un ensemble fini de références, dans le second cas, le nombre de références n'est pas fini, ce qui signifie que de nouvelles références peuvent être

fabriquées à chaque lancement. La taille de lot de chaque OF peut être différente à chaque lancement d'une même référence.

Exemple :

On fabrique 4 références différentes dont le séquençement (voir hypothèse "Séquençement") est le suivant (fabrications répétitives):



Figure 2 : Fabrication répétitive de références (tailles de lot variables)

On peut noter que dans le cas d'un horizon infini, on pourra éventuellement fonctionner avec un séquençement périodique ou apériodique des ordres de fabrication.

□ II.3- 3^{ème} hypothèse : SEQUENCEMENT du LANCEMENT

Cette hypothèse ne s'applique bien sûr qu'aux productions de type "Multiproduit" où il est nécessaire de définir une séquence ou un ordonnancement pour le lancement des différents Ordres de Fabrication.

II.3-a) Séquençement imposé sans découpage des Ordres de fabrication : Toutes les références identiques sont regroupées dans le même ordre de fabrication et la séquence des ordres de fabrication est imposée. Dans le cas d'un horizon de fabrication infini avec un séquençement périodique des ordres de fabrication comme dans le cas d'un horizon de fabrication fini, on ne doit donc avoir qu'un seul ordre de fabrication par référence et par période.

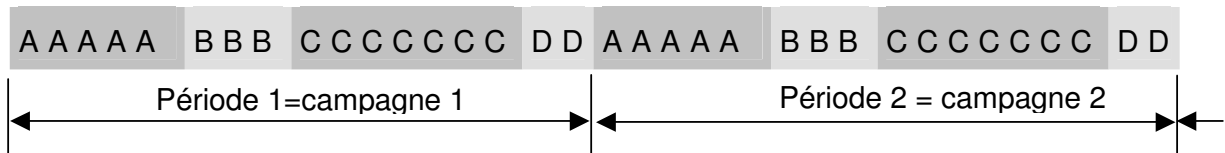


Figure 3 :Séquençement imposé sans découpage des OF par période

Remarque : Chaque période regroupant un ensemble d'OF est appelé "Campagne".

Cette hypothèse n'est pas significative pour les lancements de type "Multiproduit" non répétitifs, sans séquençement périodique des ordres de fabrication et à horizon infini.

II.3-b) Séquençement libre sans découpage des ordres de fabrication : Comme dans la classe précédente, à chaque campagne les références identiques sont regroupées dans le même ordre de fabrication, mais la séquence (ou l'ordonnancement) des ordres de fabrication est quelconque.

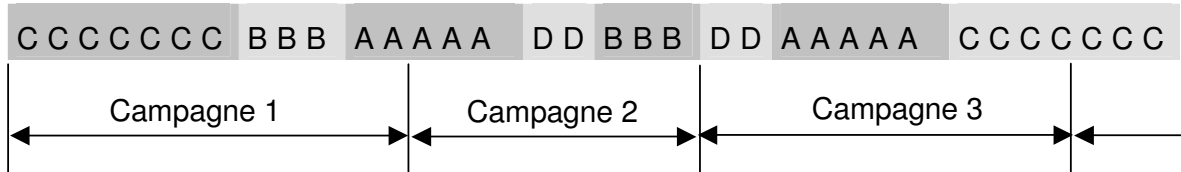


Figure 4 : Séquençement libre sans découpage des OF

II.3-c) Séquençement avec découpage des ordres de fabrication : Dans cette classe, à chaque campagne, on peut diviser chaque ordre de fabrication en plusieurs ordres de tailles variables, et les ordonnancer de façon quelconque.

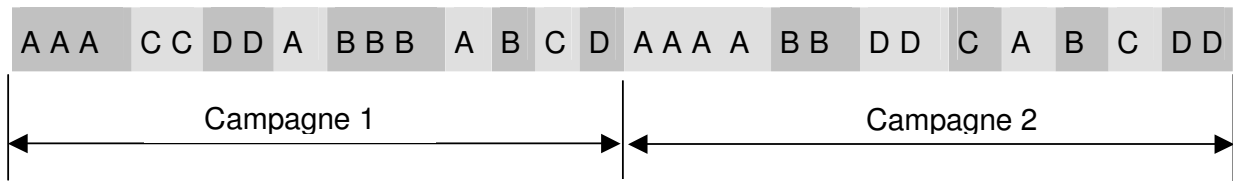


Figure 5 : Séquençement avec découpage des OF

On peut remarquer que dans le cas d'une fabrication à horizon infini, les deux classes précédentes ne font pas nécessairement apparaître de périodicité dans le séquençement.

□ II.4- 4^{ème} hypothèse : SEQUENCEMENT des OPERATIONS de fabrication

II.4-a) Opérations dans un ordre imposé sans redondance : La fabrication d'une référence doit être réalisée dans un ordre bien précis, sans permutation possible et sans redondance d'opération.

Exemple : OPé 1--> OPé 2 --> OPé 3 --> OPé 4 --> OPé 5

II.4-b) Opérations dans un ordre imposé avec redondance : Identique à la classe précédente, mais certaines opérations sont répétées.

Exemple : OPé 1--> OPé 3 --> Opé 1 --> OPé 3 --> Opé 4

II.4-c) Opérations dans un ordre quelconque sans redondance : identique à la classe a) mais les opérations peuvent être réalisées dans n'importe quel ordre.

Exemple : OPé 3 --> OPé 1 --> OPé 2 --> OPé 5 --> OPé 4

II.4-d) Opérations dans un ordre quelconque avec redondance : identique à la classe précédente, mais possibilité de répéter certaines opérations.

Exemple : OPé 2--> OPé 1 --> Opé 2 --> OPé 5--> Opé 2

□ II.5- 5^{ème} hypothèse : NOMBRE de PALETTES DISPONIBLES

Soit :

- **R_p**, le nombre de palettes disponibles pour faire une campagne, que l'on peut appeler "Réserve de palettes",

- **Q_c** la quantité totale de produits à lancer dans une campagne,

La comparaison entre ces deux variables fait apparaître plusieurs classes :

II.5-a) Réserve de palettes supérieure ou égale à la quantité de produits à lancer :

Dans ce cas, le lancement des palettes va se faire en respectant la cadence de lancement qui a été définie (voir typologie des paramètres de pilotage).

$$R_p \geq Q_c$$

II.5-b) Réserve de palettes inférieure à la quantité de produits à lancer : Cette situation peut provoquer des perturbations dans la cadence de lancement qui a été programmée. En effet, dès que les R_p premières palettes d'une campagne sont lancées, la $(R_p+1)^{i\text{ème}}$ palette doit attendre le retour et la disponibilité d'une des R_p premières palettes. Si ce temps d'attente est inférieur à la cadence de lancement programmé, aucune perturbation n'est créée. Par contre, si l'attente dépasse la cadence de lancement, un décalage se produit pour toutes les palettes qui restent à lancer. La cadence de lancement est toujours fixée relativement à la dernière palette lancée.

$$R_p < Q_c$$

Le graphique de la figure 6 présente un cas où la réserve de palette R_p est de 8, la cadence de lancement demandée (le Temps inter-Lancement $TiLa$) est de 3 sec.

Les 8 premières palettes sont donc lancées sans attente, à la cadence qui est demandée. A partir de la 9ème, le lancement est fait à condition qu'une palette soit de retour (livraison).

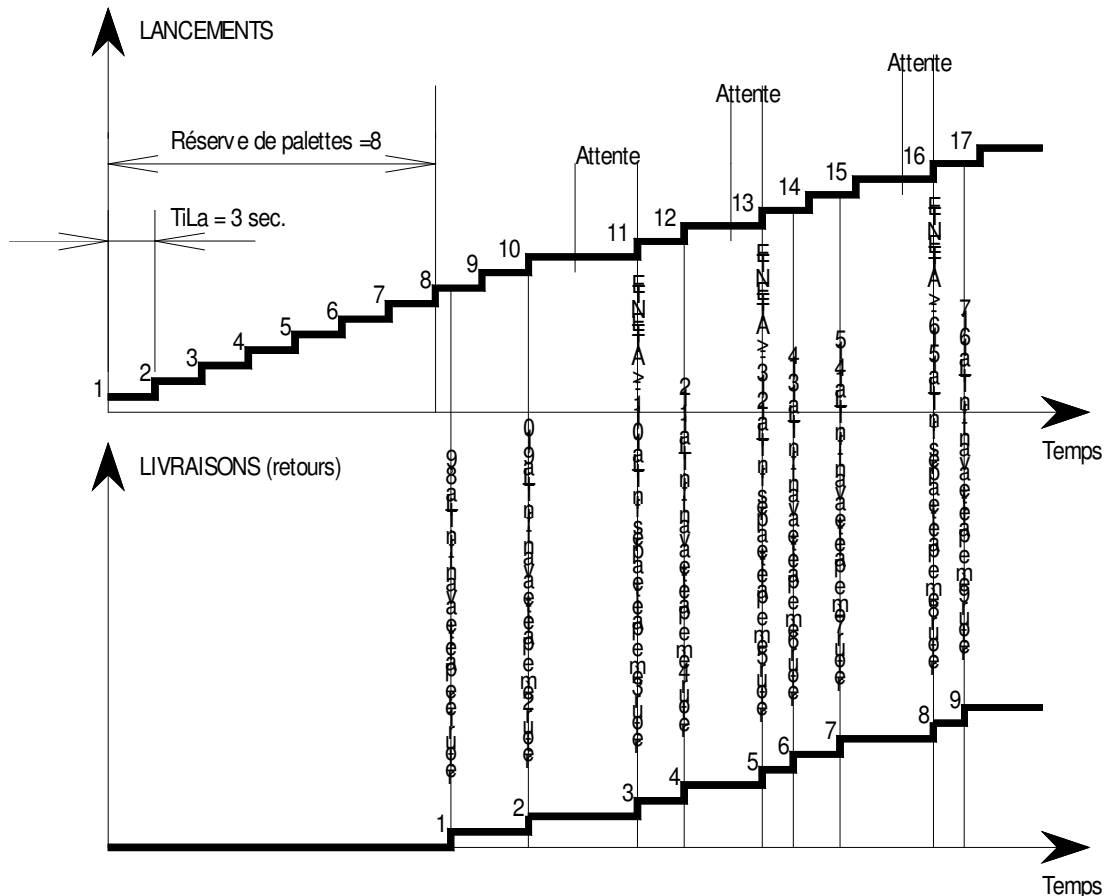


Figure 6 : Exemple : Réserve de palettes $R_p < Quantité à lancer Q_c$

III- Elaboration du plan expérimental complet

□ III.1- Définition du plan expérimental

On définit une expérimentation comme étant une combinaison de chaque hypothèse propre.

Le plan expérimental complet est le nombre total d'expérimentations qu'il est possible d'obtenir en combinant toutes les hypothèses.

A chaque expérimentation, les hypothèses générales et propres étant fixées, on exécutera plusieurs simulations. Chaque simulation correspondra à une combinaison des valeurs des variables d'entrées.

En résumé :

A partir des hypothèses, on crée des expérimentations et à partir de chaque expérimentation, on génère des simulations en fixant les valeurs des variables d'entrée.

□ III.2- Récapitulation de la typologie des hypothèses propres et estimation du niveau de contrainte

On propose dans ce chapitre de récapituler les différentes hypothèses propres au système étudié et d'indiquer si chacune d'elle est très contraignante (++) ou pas contraignante (0).

Une hypothèse est estimée contraignante si elle implique, à priori, une limitation des performances du système ou la nécessité de choisir judicieusement les valeurs des variables de pilotage.

Légende des niveaux de contrainte :

(++) = Contrainte FORTE, (+) = Contrainte FAIBLE, (0) = Contrainte NULLE ou très faible

1- VARIETE des PRODUITS lancés :

a- Lancement de type "MONOPRODUIT" (++)

b- Lancement de type "MULTIPRODUIT" (++)

2- HORIZON de FABRICATION :

a- Horizon de fabrication fini (++)

b- Horizon de fabrication infini (++)

3- SEQUENCEMENT du LANCEMENT :

a- Séquencement imposé sans découpage des Ordres de Fabrication (++)

b- Séquencement libre sans découpage des Ordres de Fabrication (+)

c- Séquencement avec découpage des Ordres de Fabrication (0)

4- SEQUENCEMENT des OPERATIONS de FABRICATION :

a- Opérations dans un ordre imposé sans redondance (++)

b- Opérations dans un ordre imposé avec redondance (++)

c- Opérations dans un ordre quelconque sans redondance (+)

d- Opérations dans un ordre quelconque avec redondance (0)

5- NOMBRE de PALETTES DISPONIBLES :

a- Réserve de palettes supérieure ou égale à la quantité de produits à lancer (0)

b- Réserve de palettes inférieure à la quantité de produits à lancer (++)

□ III.3- Structuration du plan et nombre d'expérimentations

L'hypothèse N°1 (variété des produits lancés) semble être la plus significative. Nous proposons donc d'établir un éclatement des expérimentations suivant deux jeux d'expérimentations correspondant chacun à un type de lancement :

Jeu d'expérimentations N°1 : Lancements de type "MONOPRODUIT"

Jeu d'expérimentations N°2 : Lancements de type "MULTIPRODUIT"

La combinaison de toutes les hypothèses permet d'obtenir le nombre total d'expérimentations :

$$2 \times 2 \times 3 \times 4 \times 2 = 96 \text{ expérimentations}$$

Exemples d'expérimentations obtenues :

Jeu d'expérimentation N°1 : Lancement MONOPRODUIT

Horizon INFINI / Séquencement IMPOSE SANS DECOUPAGE des OF / ORDRE IMPOSE des OPERATIONS SANS REDONDANCE / $R_p < Q_c$

Jeu d'expérimentation N°2 : Lancement MULTIPRODUIT

Horizon INFINI / Séquencement IMPOSE SANS DECOUPAGE des OF / ORDRE IMPOSE des OPERATIONS SANS REDONDANCE / $R_p < Q_c$

IV- Variables de pilotage de la ligne

Les variables de pilotage du processus (variables d'entrée) sont les variables sur lesquelles on peut agir pour optimiser le fonctionnement de la ligne suivant une fonction objectif. La fonction objectif est une relation faisant intervenir une ou plusieurs variables de sortie ou indicateurs de performance.

Le principe de l'optimisation de cette ligne repose uniquement sur la détermination d'une combinaison des variables de pilotage permettant de minimiser ou de maximiser la fonction objectif. On fait donc l'hypothèse qu'aucune modification technique ou structurelle ne peut être réalisée sur le système physique, même si l'on peut penser que cela améliorerait encore la réponse de la fonction objectif.

Pour chaque expérimentation, les hypothèses générales et propres étant fixées, on effectuera plusieurs simulations ou "Runs" en faisant varier à chacune d'elles, une ou plusieurs variables de pilotage.

L'un des points délicats de l'optimisation est justement de déterminer quelles sont les variables significatives pour la fonction objectif. Pour notre problème, on se base sur les variables qui sont à la disposition du responsable de fabrication quand il doit préparer le lancement d'une campagne de production sur la ligne :

- Le Temps inter-Lancement (TiLa);
- La capacité des Stocks statiques Amont (StAm);
- La capacité des Stocks statiques Aval (StAv);
- Les règles de priorité en sortie de poste (Prio);
- Le nombre de palettes à utiliser (Np).

Pour chacune de ces variables, plusieurs classes sont possibles.

□ IV.1- 1^{ère} variable : CADENCE de LANCEMENT

C'est l'intervalle de temps qui sépare le lancement de deux produits successifs. On l'appellera le Temps inter-Lancement (TiLa), exprimé en secondes.

IV.1-a) Cadence de lancement constante entre chaque produit et pour tous les ordres de fabrication : Dans cette classe, on ne distingue pas les différents ordres de fabrication qu'il peut y avoir dans une production "Multiproduit". La variable TiLa est unique (figure 7).

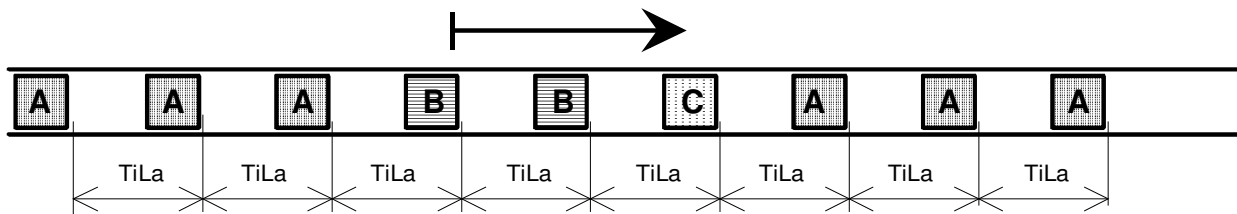


Figure 7 : Variable Tila Constante entre chaque palette

IV.1-b) Cadence de lancement constante sur chaque ordre de fabrication, mais différente d'un ordre de fabrication à l'autre : Cette classe ne s'applique qu'aux productions "Multiproduits". On peut faire apparaître éventuellement une autre variable qui est l'intervalle de temps qui sépare la dernière palette d'un ordre de fabrication à la première palette de l'ordre suivant. Dans cette classe, la variable TiLa est multiple, mais homogène. Elle sera identifiée par un indice (figure 8).

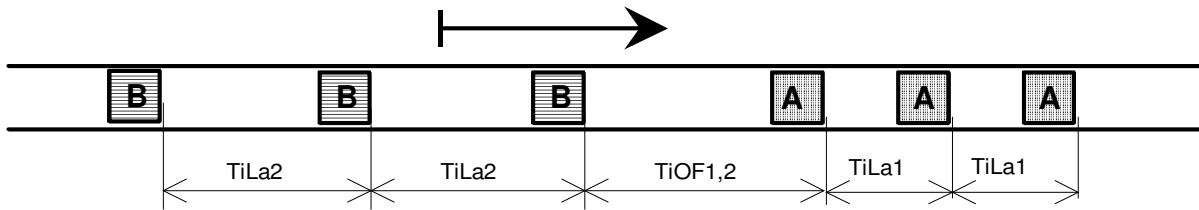


Figure 8 : Variable $TiLa$ constante sur chaque OF

IV.1-c) Cadence de lancement variable entre chaque produit : Dans cette classe, la cadence de lancement n'est plus considérée comme un paramètre dépendant de l'ordre de fabrication. La cadence peut-être différente d'un produit à l'autre, même pour des produits de même référence appartenant au même ordre de fabrication (figure 9).

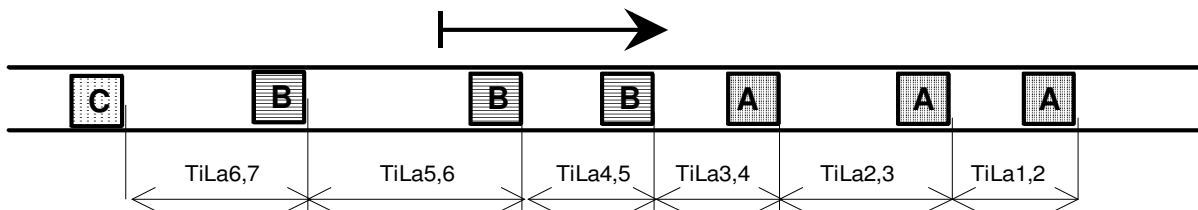
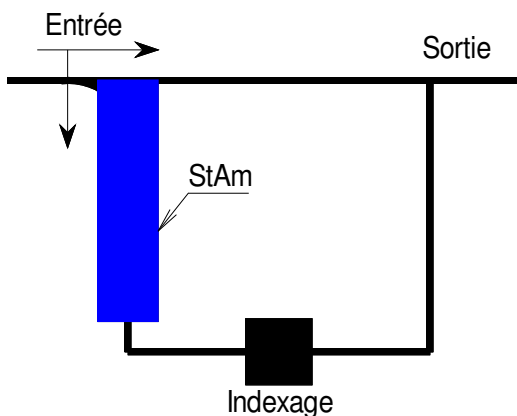


Figure 9 : $TiLa$ variable entre chaque palette

□ IV.2- 2^{ème} variable : CAPACITE des STOCKS statiques AMONT



Cette variable est appelée $StAm(i)$ avec $i = N^{\circ}$ du poste.

Elle peut être unique ou multiple.

Elle est prise en compte dans la règle d'accès d'une palette sur un poste (voir partie 3 chap. II.2)

Figure 10 : Variable $StAm(i)$

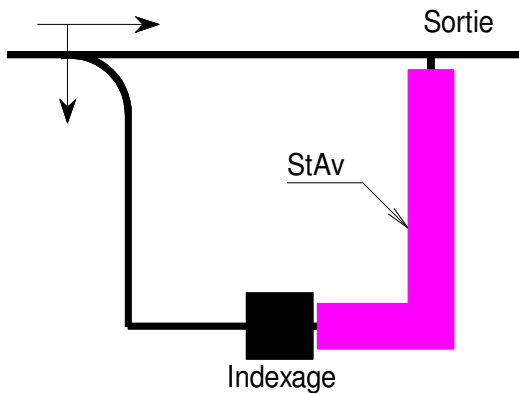
IV.2-a) Capacité identique sur tous les postes

$StAm(1)=StAm(2)=StAm(3)=\dots=StAm(6)$ = constante pour tout les postes i.

IV.2-b) Capacité différente d'un poste à l'autre

$StAm(1)\langle\rangle StAm(2)\langle\rangle StAm(3)\langle\rangle \dots \langle\rangle StAm(6)$

□ IV.3- 3ème variable : CAPACITE des STOCKS statiques AVAL



Cette variable est appelée $StAv(i)$ avec $i = N^\circ$ du poste.

Elle peut être unique ou multiple.

Elle est prise en compte dans certaines règles de priorité en sortie de poste d'une palette (voir variable "Règles de priorité en sortie de poste").

Figure 11 : Variable $StAv(i)$

IV.3-a) Capacité identique sur tous les postes

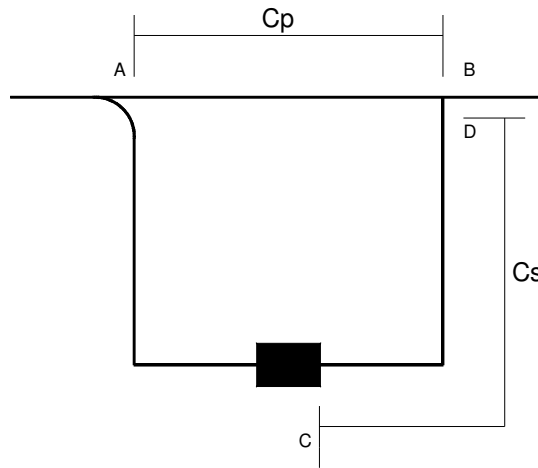
$StAv(1)=StAv(2)=StAv(3)=\dots=StAv(6)$ = constante pour tous les postes i.

IV.3-b) Capacité différente d'un poste à l'autre

$StAv(1)\langle\rangle StAv(2)\langle\rangle StAv(3)\langle\rangle \dots \langle\rangle StAv(6)$

□ IV.4- 4^{ème} variable : REGLE de PRIORITE en SORTIE de POSTE

IV.4-a) Règle 1₁ : Prise en compte du nombre de palettes en attente



Cp = nombre de palettes sur la ligne principale entre A et B
 Cs = nombre de palettes sur le poste en dérivation entre C et D
 CapaS = capacité du stock aval
 CapaP = capacité du tronçon de la ligne principale entre A et B

Règle de priorité : Si [(Cs >=Cp) ou (Cs=CapaS)] et [Cp<=CapaP]
 alors priorité à la palette sortant du poste en dérivation
 Sinon priorité à la palette sur la ligne principale
 Si [Cp>=CapaP]
 Alors BLOCAGE des palettes à l'entrée du poste (point A)

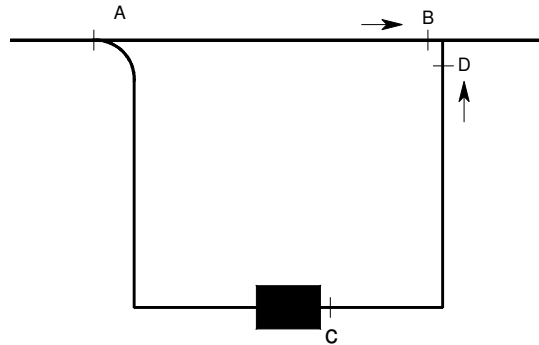
Figure 12 : Règle de priorité N^o1.1

IV.4-b) Règle 1₂ : avec les mêmes paramètres que la règle 1₁ :

Règle de priorité : Si [(Cs < Cp) ou (Cs=CapaS)] et [Cp<=CapaP]
 alors priorité à la palette sortant du poste en dérivation
 Sinon priorité à la palette sur la ligne principale
 Si [Cp>=CapaP]
 Alors BLOCAGE des palettes à l'entrée du poste (point A)

Figure 13 : Règle de priorité N^o1.2

IV.4-c) Règle 21 : Priorité à la première palette à l'intersection et priorité au poste en dérivation



point B : dB = détection des palettes venant de la ligne principale
point D : dD = détection des palettes venant du poste en dérivation

Règle de priorité : Si dD
Alors priorité à la palette en sortie de dérivation
Sinon priorité à la palette sur la ligne principale

Règle de priorité : Si dB
Alors priorité à la palette sur la ligne principale
Sinon priorité à la palette en sortie de poste en dérivation

Figure 14 : Règle de priorité N°2.1

IV.4-d) Règle 22 : Priorité à la première palette à l'intersection et priorité à la ligne principale

Règle de priorité : Si dB
Alors priorité à la palette sur la ligne principale
Sinon priorité à la palette en sortie de poste en dérivation

Figure 15 : Règle de priorité N°2.2

IV.4-e) Règle 3 : Priorité aux palettes en retard

Règle identique à la règle 2, avec en plus la prise en compte d'une information binaire de retard de la palette. Dans ce cas, le retard ne peut provenir que d'une palette circulant sur la ligne principale (la palette qui sort du poste a subi toutes les opérations alors que celle qui circule sur la ligne principale est peut être en rebouclage).

Règle de priorité : Si dB et retard Alors Priorité à la palette sur ligne principale Sinon priorité à la palette en sortie de poste en dérivation

Figure 16 : Règle de priorité N°3

IV.4-f) Règle 4 : Prise en compte d'un numéro d'ordre de priorité

Règle identique à la règle 2, mais chaque palette est affectée d'un numéro d'ordre de priorité (de 1 à n), avec par exemple : numéro 1 = priorité maximale, numéro n = priorité minimale. La priorité se fait donc en comparant les numéros d'ordre des 2 palettes présentes à l'intersection.

En cas d'égalité des numéros d'ordre, on peut trancher en prenant une des règles précédentes ou simplement en laissant la priorité au poste en dérivation ou à la ligne principale.

Exemple avec priorité au poste en dérivation en cas d'égalité des numéros d'ordre :

Numéro d'ordre (Pd) = Numéro d'ordre de la palette présente à l'intersection en sortie de Poste en Dérivation.

Numéro d'ordre (Lp) = Numéro d'ordre de la palette présente à l'intersection en sortie de Ligne Principale.

Règle de priorité : Si dD et Numéro d'ordre(Pd) \geq Numéro d'ordre(Lp) Alors Priorité à la palette en sortie de poste en dérivation Sinon priorité à la palette sur ligne principale

Figure 17 : Règle de priorité N°4

□ IV.5- 5^{ème} variable : NOMBRE de PALETTES à UTILISER

Cette variable (N_p) possède autant de valeurs qu'il y a de palettes en réserve (R_p). Il faut toutefois décider si l'on utilise toutes les palettes ($N_p = R_p$) ou si l'on limite le nombre de palettes ($N_p < R_p$). On peut donc distinguer ces deux classes :

IV.5-a) $N_p = R_p$: Toutes les palettes en réserve sont utilisées.

IV.5- b) $N_p < R_p$: On n'utilise pas toutes les palettes en réserve, bien que le nombre de produits (palettes) à lancer (Q_c) est supérieur au nombre de palettes en réserve.

□ IV.6- Niveau d'influence estimé sur les variables de sortie et valeurs

Légende du niveau d'influence :

(++) = influence forte (+) = influence faible (0) = influence nulle ou négligeable

1- CADENCE de LANCEMENT ($TiLa$) :

Influence : (++) si horizon fini, (0) si horizon infini

Valeurs possibles entre chaque palette lancée :

$$1 \leq TiLa \leq 20$$

2- CAPACITE des STOCKS AMONT ($StAm$) :

Influence : (++) si horizon fini, (+) si horizon infini

Valeurs possibles sur chaque poste en dérivation :

$$1 \leq St Am(i) \leq 6$$

3- CAPACITE des STOCKS AVAL ($StAv$) :

Influence : difficile à estimer

Valeurs possibles sur chaque poste en dérivation :

$$1 \leq St Av(i) \leq 6$$

4- REGLES de PRIORITE en SORTIE de POSTE (Prio) :

Influence : difficile à estimer.

Nombre de règles possibles à chaque sortie de poste : 6

5- NOMBRE de PALETTES à UTILISER (Np) :

Influence : (+)

Valeurs possibles : 25 palettes en réserve (Rp)

V- Sélection des expérimentations et des variables d'entrée

Dans le chapitre 3, on a déterminé le plan expérimental complet, c'est à dire issu de la combinaison de toutes les hypothèses. On a obtenu 96 expérimentations possibles, mais nous allons maintenant sélectionner celles qui seront utilisées en simulation. Pour ce qui est des variables d'entrée, dans le chapitre 4 nous avons identifié cinq variables, chacune avec plusieurs valeurs. De la même façon que pour les expérimentations, toutes les variables et toutes les valeurs ne seront pas testées. En effet, bien que destinées à la simulation, on a intérêt à limiter le nombre d'expérimentations et de variables pour ne pas avoir trop de cas à traiter. La sélection va se faire en fonction d'un critère de choix permettant de ne retenir que les expérimentations et les variables significatives, c'est à dire celles qui correspondent aux cas les plus critiques ou les plus fréquemment rencontrés en situation réelle.

□ V.1- Expérimentations retenues

On a déjà défini deux jeux d'expérimentations qui correspondent chacun à une situation significative, lancement MONOPRODUIT et lancement MULTIPRODUIT.

Hypothèses propres communes aux deux jeux d'expérimentations :

- Nombre de palettes en réserve : **Rp = 25**, inférieur ou égal à la quantité à lancer.
- Séquencement du lancement imposé sans découpage des ordres de fabrication.
- Séquencement des opérations dans un **ordre imposé sans redondance** avec $Pi[j]$ = opération sur le poste i ($i = 1$ à 6) pendant un temps de cycle de j secondes.

Les expérimentations sélectionnées seront donc les suivantes, avec les hypothèses propres spécifiques à chaque expérimentation :

V.1-a) Jeu N°1 (Lancement MONOPRODUIT)

- *Expérimentation N°1 :*

Horizon de fabrication FINI, correspondant à un lancement de 25 produits.

- *Expérimentation N°2 :*

Horizon de fabrication INFINI, prise de mesures entre la 50^{ème} et la 150^{ème} palette, en régime stabilisé.

Ces deux expérimentations constituent l'une et l'autre un point de départ. En fonction des résultats obtenus et de l'analyse qui en sera faite, des expérimentations complémentaires pourront être réalisées, sur la base des mêmes conditions. En particulier, les valeurs des temps de cycle $T_{cyc}(i)$ seront initialement fixées arbitrairement, mais d'autres valeurs seront testées si nécessaire.

La figure 18 représente le circuit ou le routage optimal d'une palette du jeu d'essais N°1. Il s'agit d'un routage optimal car il n'y a pas de rebouclage de la palette dû à une impossibilité d'accéder à l'un des postes de travail.

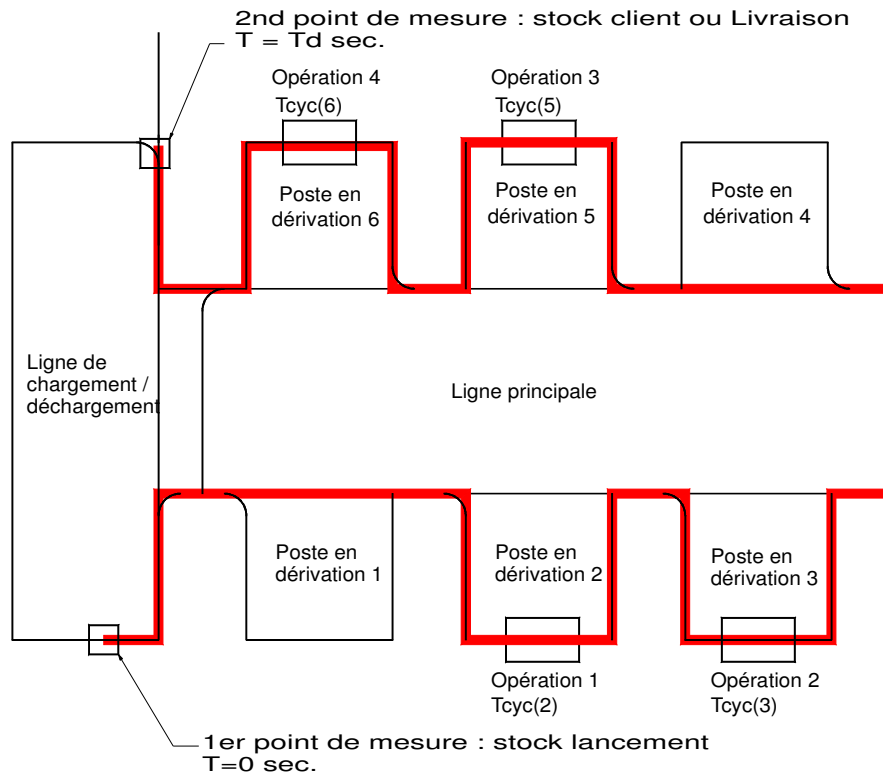


Figure 18 : Graphe de circulation d'une palette

V.1-b) Jeu N°2 (Lancement MULTIPRODUIT)

- Expérimentation N°1 :

Horizon de fabrication FINI, correspondant à un lancement séquentiel de 5 ordres de fabrication. L'ensemble des O.F lancés est appelé "Campagne".

- Expérimentation N°2 :

Horizon de fabrication INFINI, correspondant à une succession de campagnes identiques à celle utilisée pour l'expérimentation N°1. Prise de mesures entre la 100^{ème} et la 200^{ème} palette, en régime stabilisé.

Les figures 19 à 23 représentent les routages des palettes des cinq ordres de fabrication générés arbitrairement pour ce jeu d'essais.

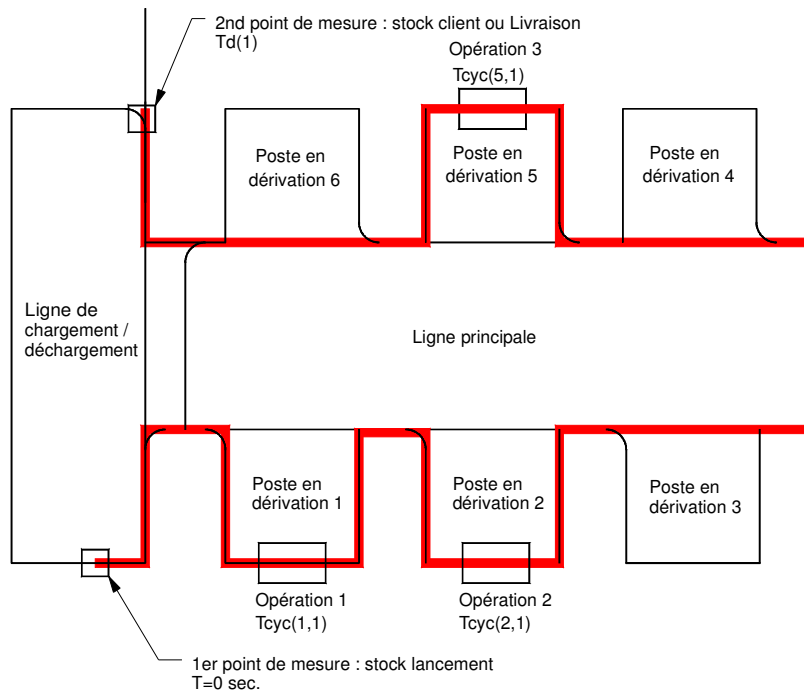


Figure 19 : Graphe de circulation d'une palette de l'O.F. N°1

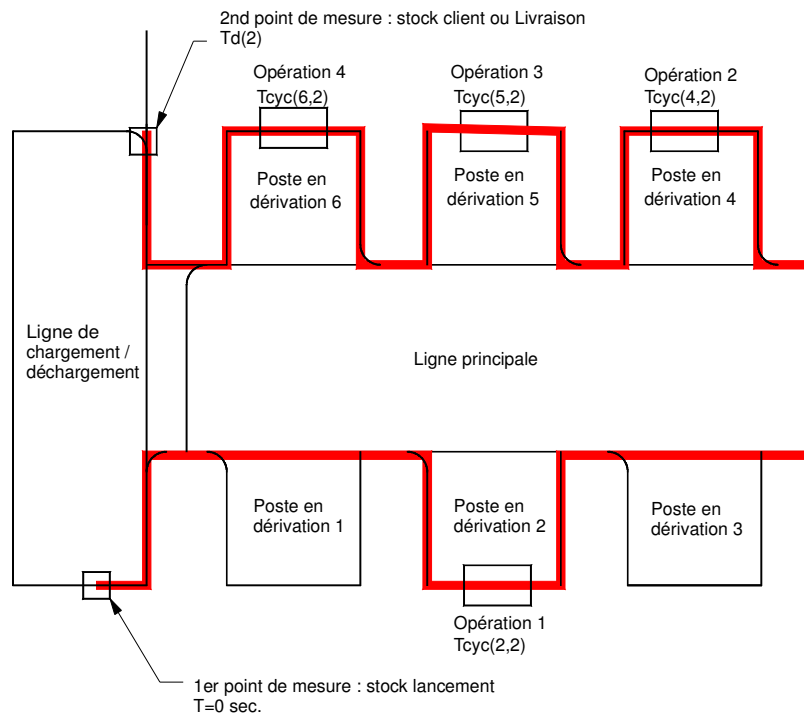


Figure 20 : Graphe de circulation d'une palette de l'O.F. N°2

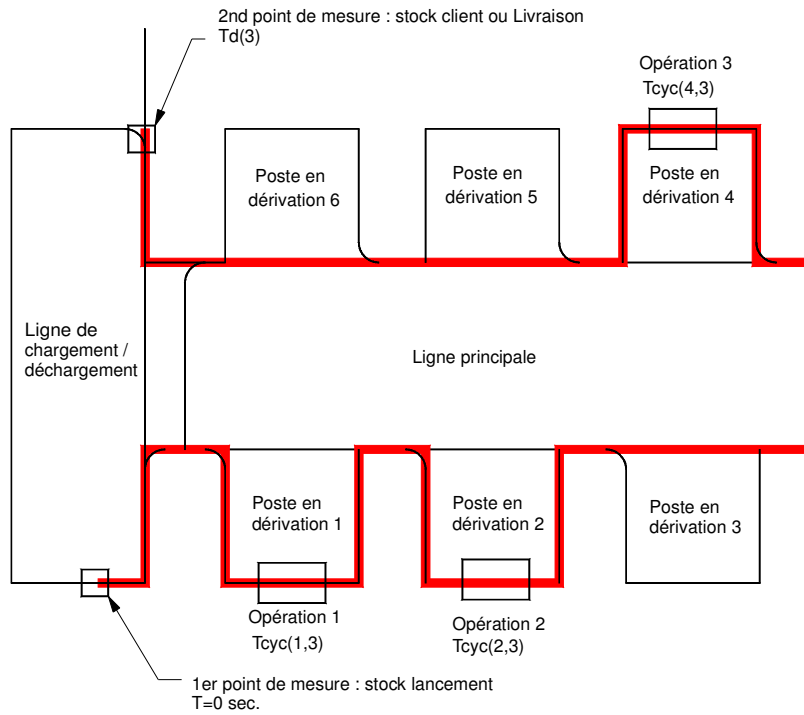


Figure 21 : Graphe de circulation d'une palette de l'O.F. N°3

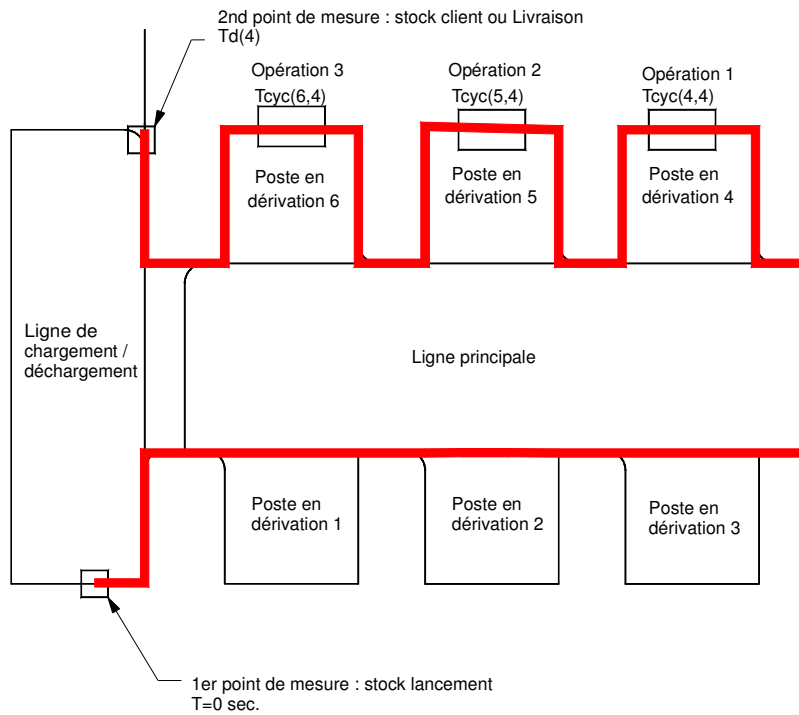


Figure 22 : Graphe de circulation d'une palette de l'O.F. N°4

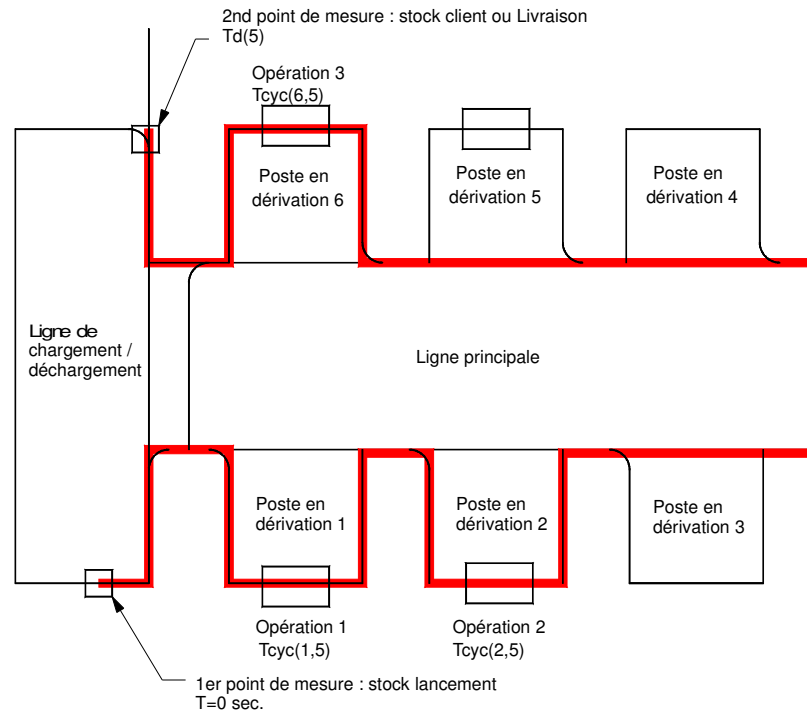


Figure 23 : Graphe de circulation d'une palette de l'O.F. N°5

□ V.2- Variables de pilotage testées à chaque expérimentation

Compte tenu du nombre de variables et du grand nombre de valeurs qu'elles peuvent prendre, on restreint l'étendue du problème à uniquement deux variables :

- **Cadence du lancement demandée = Temps inter-Lancement (TiLa)** en secondes : valeurs entières comprises entre 1 et 20 sec par pas de 1 sec. On se limite au cas où la cadence est constante entre chaque palette et pour tous les ordres de fabrication. On a donc 20 combinaisons possibles de la variable TiLa.

- **Capacité des stocks statiques amont = StAm(i)** comprise entre 1 et 6 palettes au maximum. On se limite au cas où la capacité est identique sur chaque poste en dérivation. On a donc 6 combinaisons possibles de la variable StAm.

A chaque expérimentation, il y a donc 20 x 6 = 120 simulations à réaliser.

Les autres variables seront donc bloquées à une valeur choisie arbitrairement :

- Capacité des stocks statiques aval $StAv(i) = 6$ palettes au maximum.
- Règle de priorité en sortie de poste : Règle 1₁
- Nombre de palettes à lancer $Np = 25 =$ réserve de palettes Rp .

VI- Mesures des variables de sortie (ou indicateurs de performance) réalisées à chaque simulation

□ VI.1- Temps de défilement de chaque palette **Td** (en sec)

Le temps de défilement (**Td**) est le temps qui sépare le lancement d'une palette sur la ligne (heure de lancement = $HLan$) à son déchargement (heure de livraison = $HLiv$).

$$\mathbf{Td = HLiv - Hlan}$$

Pour chaque jeu d'essais, les figures 18 à 23 indiquent le temps de défilement minimal (voir sur chaque figure le temps indiqué au 2nd point de mesure).

Le temps de défilement est dans ce cas mesuré dans des conditions de flux optimales :

- la ligne est vide, aucune autre palette que celle utilisée pour la mesure n'est en circulation,
- les postes de travail sont disponibles sans attente,
- les stocks tampons sont vides.

□ VI.2- Délai de Fabrication d'une campagne **D** (en sec)

Le délai de Fabrication **D** est le temps qui sépare le lancement de la première palette de la campagne (HLanp) au déchargement de la dernière palette de la campagne (HLivd).

$$\mathbf{D} = \mathbf{HLivd} - \mathbf{HLanp}$$

□ VI.3- Nombre de réglages par poste de travail **Nr(i)**

On considère qu'il y a réglage sur un poste de travail quand on change d'article (passage d'une référence de produit à une autre). Le temps de réglage n'est pas comptabilisé car c'est une valeur qui peut fortement varier en fonction des cas. On mesure donc uniquement le nombre de fois où le poste a eu à passer d'un produit à un autre.

□ VI.4- Ratio de fluidité de chaque palette **Rf**

Le ratio de fluidité (appelé aussi ratio de tension du flux) est le rapport entre la somme des temps opératoires et le temps de défilement.

$$\mathbf{Rf} = \frac{\sum_{i=1}^6 \mathbf{Tcyc}(i,k)}{\mathbf{Td}}$$

Avec $\mathbf{Tcyc}(i,k)$ = Temps de cycle du poste \mathbf{P}_i ($i = 1 \dots 6$) pour une palette de l'ordre de fabrication $\mathbf{N}^\circ k$ ($k = 1 \dots 5$).

□ VI.5- Encours de palettes sur la ligne entre le lancement et le déchargement **EC**.

L'encours de palettes se calcule en faisant la différence entre le nombre de palettes lancées (NLan) et le nombre de palettes déchargées (NLiv). L'encours est calculé par exemple à chaque lancement et à chaque déchargement de palette, ou à intervalle de temps régulier.

$$\mathbf{EC = N\text{Lan} - N\text{Liv}}$$

Pour chacune de ces variables de sortie, des statistiques pourront être établies afin de déterminer les valeurs mini, maxi, moyenne, l'écart type et éventuellement l'intervalle de confiance. On pourra aussi représenter la distribution des valeurs dans des histogrammes.

Conclusion

Cette partie met en évidence la complexité du problème traité en terme de multiplicité des combinaisons. Bien que nous disposions d'un modèle de simulation de la ligne qui permet, comme cela a été présenté dans la partie 4, de réaliser des expérimentations virtuelles avec un temps de réponse très faible par rapport au système réel, on constate que le nombre de combinaisons est très élevé.

Après une classification des hypothèses et des variables de pilotage, on structure ces hypothèses en plusieurs expérimentations, chacune d'elles correspondant à une combinaison des hypothèses. A chaque expérimentation, il faut faire varier les valeurs (quantitatives ou qualitatives) des variables de pilotages. Cela multiplie encore le nombre de simulations.

Des simulations exhaustives ne sont donc pas envisageables. On limite donc les simulations à deux cas significatifs, appelés jeux d'essais : lancements Monoproduits et lancements Multiproduits. Pour chacun d'eux, deux expérimentations sont définies : en horizon fini ou en horizon infini.

De la même façon, seules deux variables de pilotage sont utilisées : la cadence du lancement, appelé Temps inter-Lancement (TiLa), et la capacité des stocks statiques Amont (StAm). Nous obtenons 120 combinaisons possibles à partir de ces seules variables.

Toutes les simulations qui seront réalisées ont pour but de recueillir des données. Dans le dernier chapitre, des variables de sortie sont définies à partir des données qui seront recueillies pendant et après la simulation.

- PARTIE 6 -
***Analyse des résultats
des simulations***

Sommaire

Introduction	173
I- Analyse des résultats du jeu d'essais N°1	174
□ I.1- Rappel des conditions expérimentales générales	174
□ I.2- Expérimentation N°1	174
I.2-a) Rappel des conditions propres à cette expérimentation	174
I.2-b) Résultats	176
I.2-c) Analyse des données récoltées et recherche d'une combinaison optimale	177
I.2-d) Relations établies à partir de l'expérimentation N°1	190
I.2-e) Expérimentation avec temps de cycle différents	194
I.2-f) Expérimentation avec temps de cycle équilibrés	195
□ I.3- Expérimentation N°2	196
I.3-a) Rappel des conditions propres à cette expérimentation	196
I.3-b) Résultats	196
II- Analyse des résultats du jeu d'essais N°2	202
□ II.1- Rappel des conditions expérimentales générales	202
□ II.2- Expérimentation N°1	202
II.2-a) Rappel des conditions propres à cette expérimentation	202
II.2-b) Résultats	203
II.2-c) Analyse des données récoltées et comparaison au cas monoproduit	205
II.2-d) Expérimentation avec $N_p > 25$ palettes	211
□ II.3- Expérimentation N°2	214
II.3-a) Rappel des conditions propres à cette expérimentation	214
II.3-b) Résultats	215
II.3-c) Analyse des données récoltées	217
II.3-d) Expérimentation avec $N_p = 30$ palettes	220
Conclusion	223
Bibliographie	224

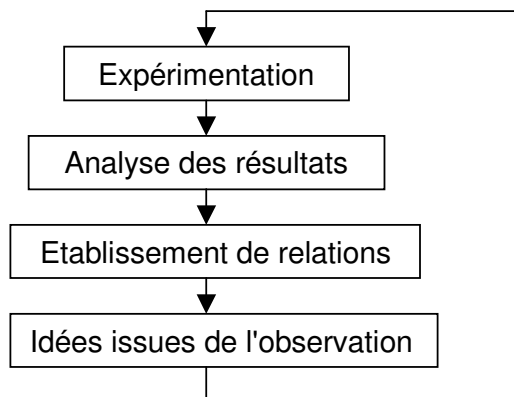
Introduction

Dans cette partie, des expérimentations sur le modèle de simulation vont être réalisées en faisant varier uniquement deux paramètres de pilotage : le **T**emps **i**nter-**L**ancement (TiLa) et la capacité des **S**tocks **A**mont (StAm).

Les objectifs de ces expérimentations sont multiples :

- Recueillir des données pour pouvoir analyser le comportement dynamique du système en fonction des paramètres de pilotage;
- Proposer une méthodologie, à partir de l'analyse des données recueillies, permettant de déterminer les valeurs optimales des paramètres de pilotage pour minimiser une des variables de sortie : le délai de réalisation (D) de la campagne entre le lancement de la première palette et la livraison de la dernière palette. Dans un second temps, une autre variable de sortie pourra être retenue afin de pouvoir trancher entre des solutions distinctes donnant le même délai de réalisation. On utilisera en particulier la valeur de l'encours moyen (EC moyen) comme critère complémentaire.
- Trouver ou vérifier des relations permettant de dimensionner certains paramètres physiques (nombre de palettes, longueur des tronçons servant de stock amont, ...) dans un but de conception d'une ligne transfert asynchrone. Dans ce cas, la ligne est utilisée comme un système générique paramétrable.

L'intérêt de cette démarche réside dans le bouclage entre les expérimentations, l'analyse des résultats, l'établissement de relations, et la réalisation de nouvelles expérimentations à partir d'idées issues de l'observation, comme le montre le schéma ci-dessous.



En fonction de l'analyse des données obtenues, les différentes expérimentations planifiées dans la partie 5 pourront donner lieu à des expérimentations complémentaires non programmées initialement.

I- Analyse des résultats du jeu d'essais N°1

□ I.1- Rappel des conditions expérimentales générales

- Lancement de type "MONOPRODUIT".
- Nombre de palettes en réserve : $R_p = 25$, inférieur ou égal à la quantité à lancer.
- Séquencement des opérations dans un ordre imposé, sans redondance avec :
 $P_i[j]$ = opération sur le poste i ($i=1..6$) pendant un temps de cycle de j secondes.

□ I.2- Expérimentation N°1

I.2-a) Rappel des conditions propres à cette expérimentation

- Horizon de fabrication FINI : la ligne est vide en fin de simulation.
- Lancement de 25 palettes.
- Les temps de cycle sont fixés arbitrairement aux valeurs suivantes :

$$\mathbf{P2[7] - P3[12] - P5[8] - P6[3]}$$

On peut constater que les valeurs des temps de cycle sont fortement déséquilibrées. Si l'on considère que le taux d'équilibrage des temps de cycle d'une ligne de production est :

$$Eb = \frac{\sum_{i=1}^6 T_{cyc}(i)}{w \cdot \text{Max}(T_{cyc}(i))}$$

avec, pour un routage donné :

$T_{cyc}(i)$ = Temps de cycle du poste de travail N° i ($i = 1..6$),

w = nombre de postes de travail utilisés ($w = 1..6$),

$\text{Max}(T_{cyc}(i))$ = Temps de cycle le plus long = Temps du poste goulet = T_g ,

Le taux d'équilibrage est de $(7 + 12 + 8 + 3) / (4 \times 12) = 0,625$ soit 62,5 %. Si tous les postes de travail avaient le même temps de cycle, le taux d'équilibrage serait de 100 %.

Un tel déséquilibre des temps de cycle n'est pas totalement arbitraire. En effet, la vocation première d'une telle ligne de production, avec un transfert asynchrone et des postes de travail en dérivation est de permettre une production flexible (voir la partie 3). Contrairement à des lignes de production dédiées à un seul produit où l'on cherche effectivement à équilibrer le plus possible les temps de cycle des postes de travail, une ligne flexible pourra difficilement satisfaire ce critère d'équilibrage. D'autre part, plus le déséquilibre est important entre les postes de travail, plus le pilotage du flux est complexe. Ces raisons justifient donc le choix des valeurs relatives entre les temps de cycle de cette expérimentation. En fonction des résultats obtenus, d'autres expérimentations, basées sur les mêmes conditions, mais avec des valeurs de temps de cycle différentes, pourront être réalisées.

- Routage des palettes :

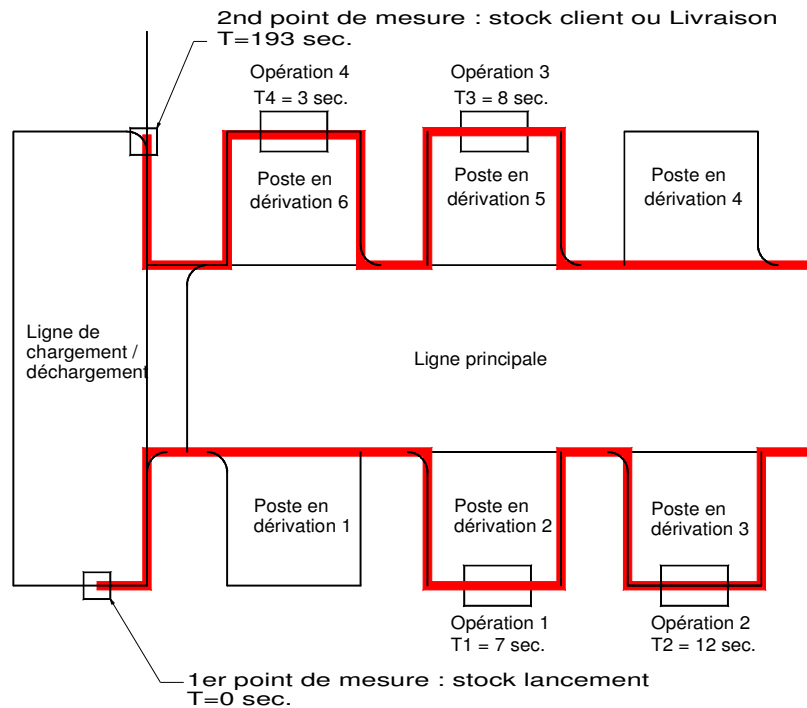


Figure 1 : Routage des palettes du jeu d'essais N°1

- Temps de défilement minimal d'une palette avec ce routage et ces temps de cycle sur les postes de travail : 193 secondes.

I.2-b) Résultats

Les courbes ci-dessous représentent la variation du délai de production (D) en fonction du Temps Inter-Lancement (TiLa), pour chaque valeur de capacité du stock Amont (StAm).

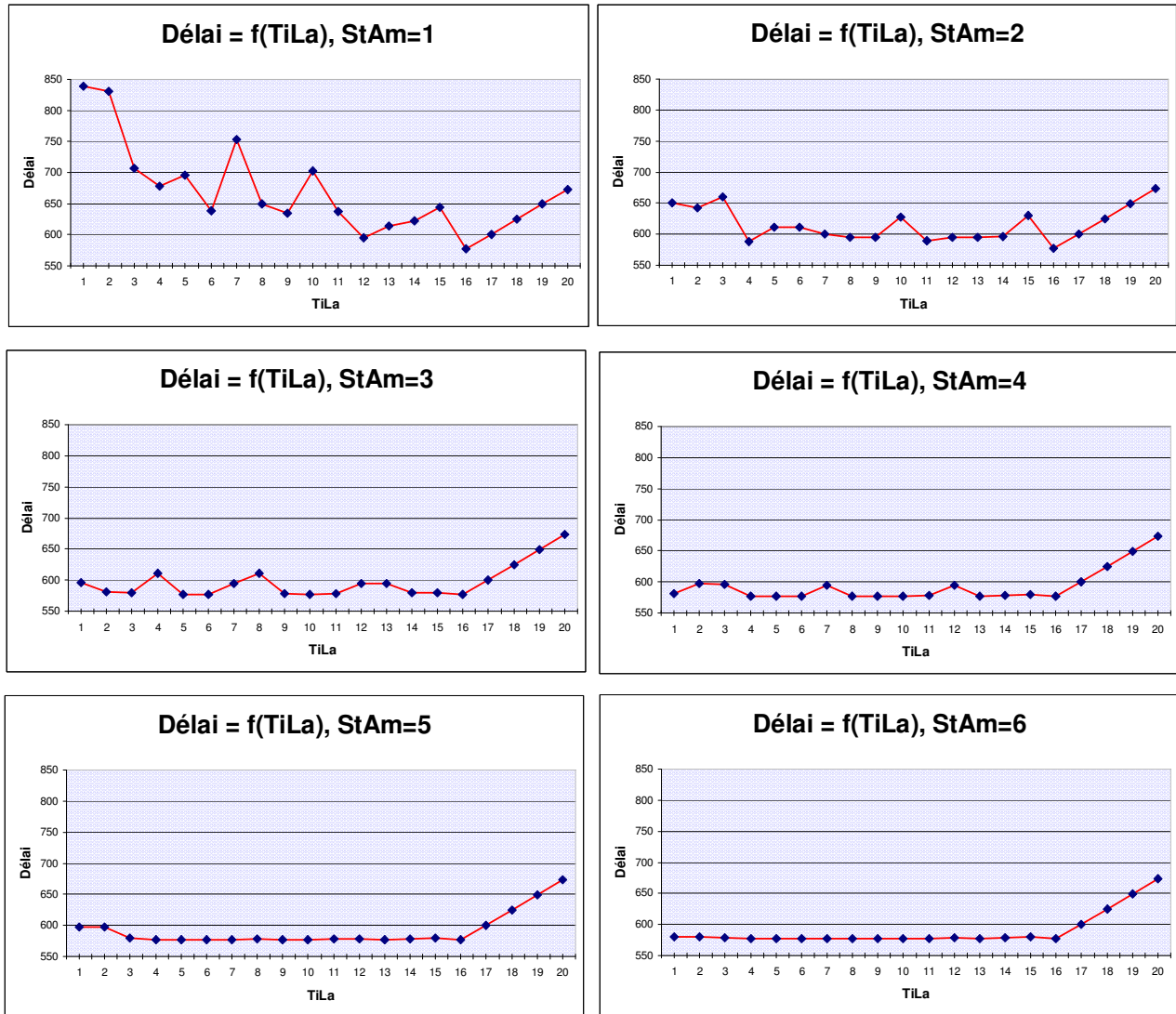


Figure 2 : courbes $D = f(TiLa)$ pour chaque valeur de StAm

Le graphique de la figure 3 est une superposition des courbes précédentes et met en évidence l'influence des stocks amont (StAm) et du Temps inter-Lancement (TiLa) sur le délai (D).

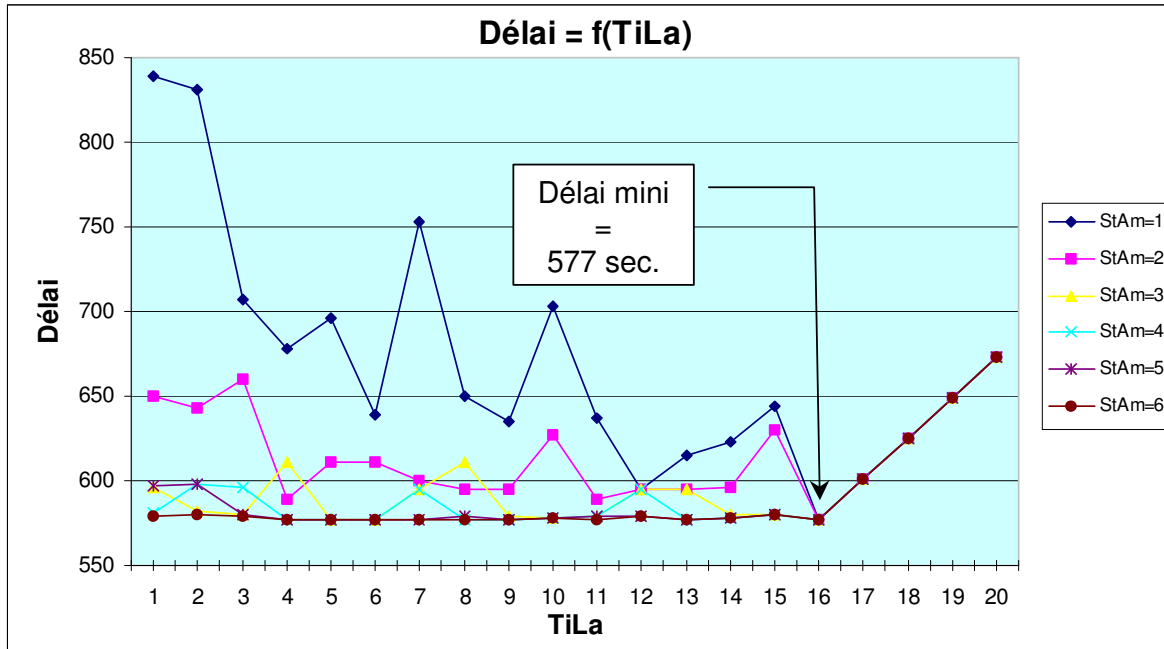


Figure 3 : Graphique de synthèse des résultats

1.2-c) Analyse des données récoltées et recherche d'une combinaison optimale

Les courbes **Délai = f(TiLa)** mettent en évidence le rôle des stocks tampons. Plus la capacité des stocks amont augmente, plus la courbe **Délai = f(TiLa)** s'aplatit pour les valeurs de TiLa inférieures ou égales à 16 secondes, ce qui correspond à un amortissement des variations ou des perturbations du système. Dans cette expérience, où les pannes ne sont pas prises en compte, les stocks tampons permettent donc d'amortir les variations de flux entre le lancement des palettes sur la ligne de chargement/déchargement et les différents postes de travail. On voit très bien que quand les capacités des stocks tampons sont élevées (StAm=6), la cadence de lancement TiLa peut prendre n'importe quelle valeur à l'intérieur d'une plage assez importante (de 1 sec. à 16 sec.) sans que cela produise de variations importantes sur le délai. Par contre, quand la capacité des stocks tampons est limitée à 1 (StAm=1), la courbe présente une allure multimodale, avec plusieurs minimums locaux, mais un seul minimum global.

Dans le cas où la capacité des stocks amont est de 1, le tracé montre le comportement "perturbé" ou "chaotique" du système : une variation d'un pas de la cadence de lancement TiLa peut provoquer de fortes variations du délai de fabrication. Par exemple, pour $TiLa = 6$ sec. où l'on a un minimum local, le délai est de 638 sec, mais pour $TiLa=7$ sec., le délai passe à 752 sec., soit une augmentation de 18 %.

Par ailleurs, on constate que tous les tracés sont confondus et croissants au delà de $TiLa=16$ secondes, valeur qui donne un délai minimal, quelle que soit la valeur du stock tampon. Cette valeur de $TiLa$ correspond donc très certainement à une caractéristique du système.

A partir des résultats et des courbes précédentes, le graphique de la figure 4 met en évidence les 20 combinaisons des valeurs des stocks amont ($StAm$) et du Temps inter-Lancement ($TiLa$) parmi les 120 combinaisons testées qui permettent de minimiser le délai (tous les points correspondent à un délai minimal de 577 sec.).

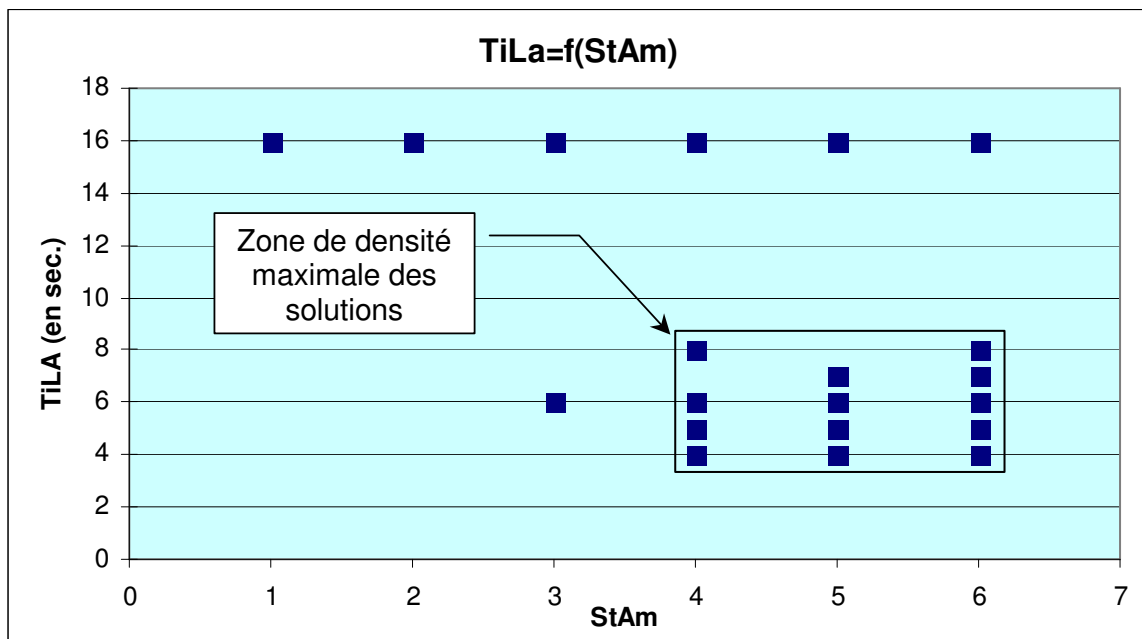


Figure 4 : Combinaisons optimales de $TiLa$ et $StAm$

Le problème qui se pose est donc de déterminer quelles sont les valeurs optimales de cadence de lancement ($TiLa$) et de stocks amont ($StAm$) à choisir parmi les vingt combinaisons possibles qui minimisent le critère (D).

Si l'on s'en tient uniquement à la minimisation du délai (D), nous allons maintenant proposer une méthode de sélection d'une solution, basée sur l'utilisation d'un critère complémentaire et indépendant de toute autre variable de sortie.

On peut identifier sur le graphique de la figure 4 une **zone de densité maximale** (en encadré) qui correspond à un regroupement de solutions optimales dans les deux dimensions (TiLa) et (StAm). Cela exclut donc toutes les solutions correspondant à $TiLa = 16$ sec., car elles sont trop isolées et risquent donc de ne pas garantir une bonne "stabilité" du délai, bien qu'étant regroupées, mais seulement dans une dimension. En effet, les tracés montrent qu'une variation de la valeur de (TiLa) peut dans ce cas entraîner une augmentation du délai. Par contre, dans la zone de densité maximale, plusieurs solutions sont envisageables, et en particulier celles qui garantissent la stabilité de la réponse, même en cas de variation de l'une ou de l'autre des variables d'entrée. Il suffit pour cela de prendre une solution, si elle existe, qui soit elle-même encadrée dans les deux dimensions par d'autres solutions. Dans le cas de cette expérimentation, on peut sélectionner deux solutions qui répondent à ce nouveau critère, que l'on appelle **critère de stabilité** (voir figure 5) :

Solution 1 : StAm = 5, TiLa = 6

Solution 2 : StAm = 5, TiLa = 5

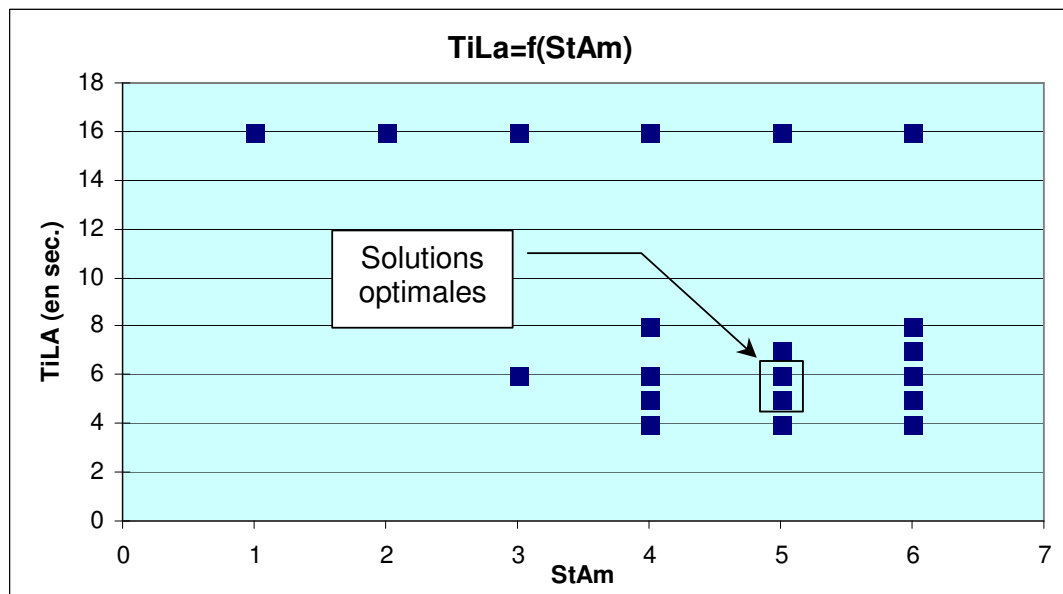


Figure 5 : Solutions optimales d'après le critère de stabilité

On peut donc choisir l'une ou l'autre de ces solutions en garantissant une stabilité du délai, même en cas de variation des paramètres de pilotage. Ces variations sont toutefois limitées à ± 1 sur chacune des échelles.

Ce critère de stabilité n'est pas à négliger, car le système réel est soumis à des perturbations aléatoires, généralement d'amplitude limitée, de ses paramètres de pilotage.

Nous allons maintenant prendre en compte un nouveau critère, l'encours moyen de palettes (**EC**), dont la minimisation est un objectif important en production, au même titre que le délai de production :

L'analyse des courbes d'encours **EC = f(temps)** (voir figures 6 et 7) met en évidence les 3 phases décrites dans les hypothèses propres pour l'horizon de fabrication (voir partie 5, chap. II.2) : montée en régime, régime stabilisé et vidage de la ligne. Dans le cas où $TiLa=16$ sec et quelle que soit la valeur de la capacité des stocks amont (figure 6), le tracé est régulier, symétrique et le régime stabilisé est bien évident (entre environ 180 et 400 sec.). L'encours maximal ne dépasse pas 12 palettes, l'encours moyen sur la période totale de fonctionnement est de 8,86 palettes. Dans le cas où $TiLa$ est inférieur à 16 sec., la variation de l'encours dans le temps est beaucoup moins nette. Par exemple le tracé que l'on obtient pour $TiLa = 6$ sec. et $StAm = 5$ (figure 5) fait aussi apparaître les trois phases caractéristiques d'une production à horizon fini, mais la phase de régime stabilisé est quasiment nulle. Après une montée en régime très rapide (pic à 100 sec.), la ligne se vide progressivement sans jamais avoir été en régime stabilisé. De plus, l'encours maximal est beaucoup plus important que dans le cas précédent : 25 palettes (ce qui correspond à la réserve R_p) au lieu des 12 nécessaires pour obtenir le même délai de fabrication. L'encours moyen sur la période est de 12,5 palettes.

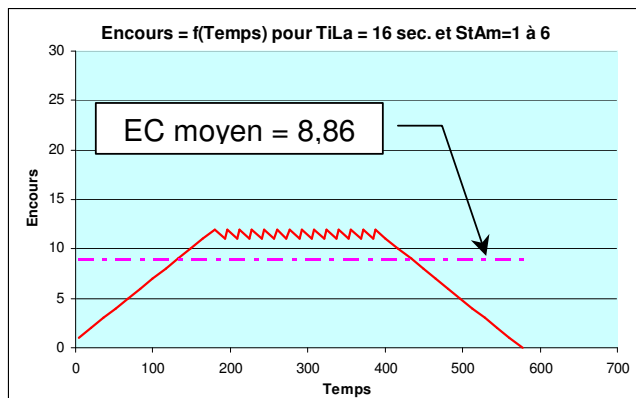


Figure 6 : EC pour $TiLa = 16$ et $StAm = 1$

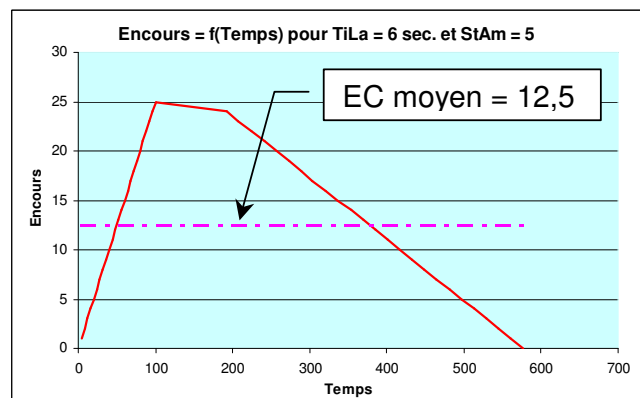


Figure 7 : EC pour $TiLa = 6$ et $StAm = 5$

La conclusion est donc moins évidente qu'il n'y paraît, puisque les valeurs des paramètres ($TiLa$) et ($StAm$) permettant de minimiser l'encours sont différentes de celles trouvées précédemment pour minimiser le délai avec le critère de stabilité (voir figure 8).

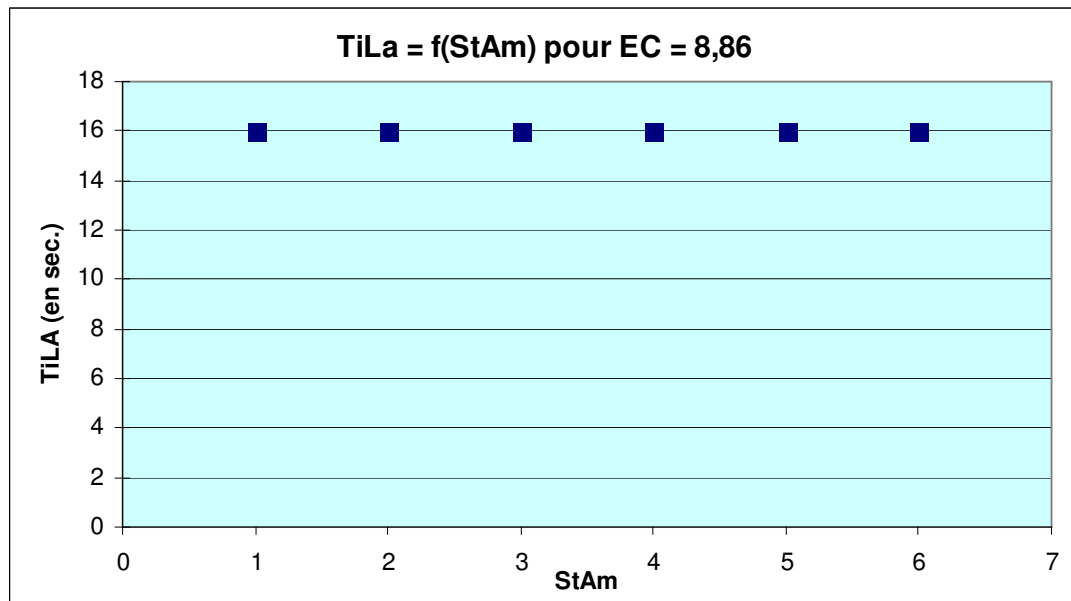


Figure 8 : Solutions optimales de TiLa et StAm pour minimiser l'encours EC

Il faut donc rechercher un compromis entre un délai minimal et un encours minimal, tout en garantissant une stabilité de l'un et de l'autre.

Dans cette analyse, la solution proposée est de tolérer une légère dégradation de l'un et de l'autre de ces critères, afin d'augmenter le nombre de solutions possibles. Une dégradation d'environ 5% du délai minimal (577 sec. < D < 598 sec.) augmente considérablement les solutions possibles (voir figure 9).

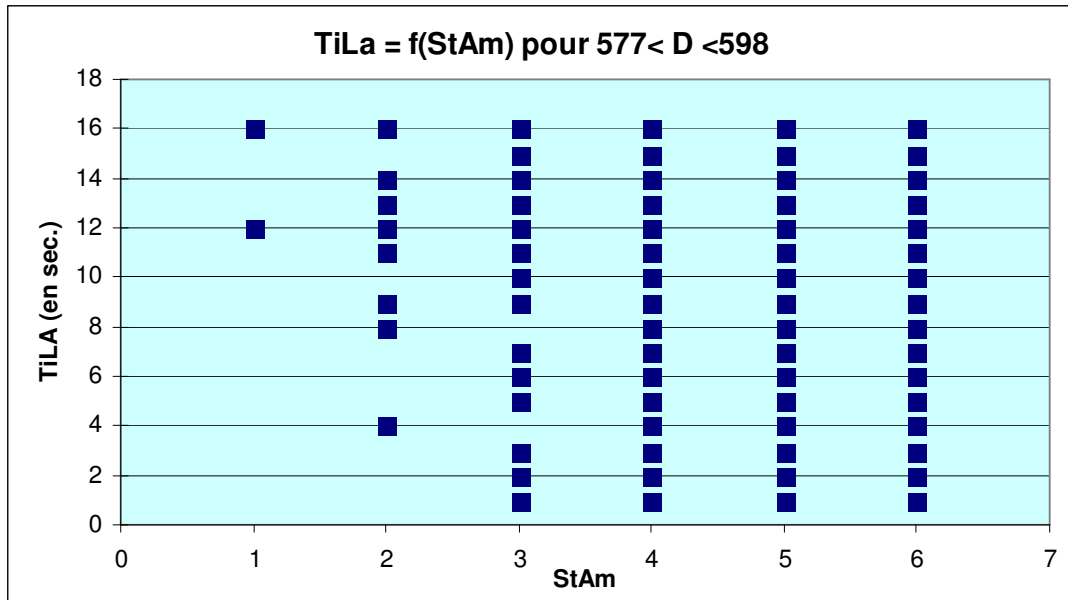


Figure 9 : Ensemble de solutions avec dégradation du délai de 5%

L'étude de l'encours moyen pour chaque solution montre que celui-ci est minimal (8,86 palettes) pour toutes valeurs de TiLa supérieures ou égales à 16 secondes, quelle que soit la valeur de StAm. Une dégradation de l'encours moyen de 10% ($8,86 < EC < 10,18$) permet de prendre en compte des solutions avec des TiLa de 14 et 15 secondes (voir figure 10).

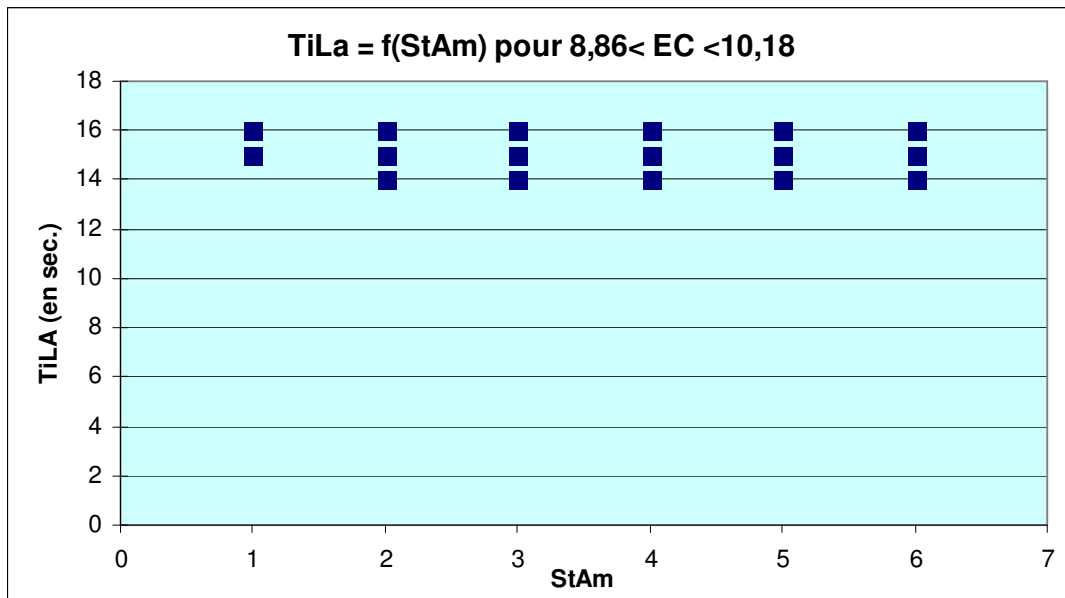


Figure 10 : Ensemble de solutions avec dégradation de l'encours (EC) de 10%

On peut donc, à partir de ces deux critères dégradés, proposer un ensemble de solutions correspondant à l'intersection des deux ensembles, sur lequel un critère de stabilité, identique à celui énoncé dans l'optimisation du délai, est retenu (figure 11).

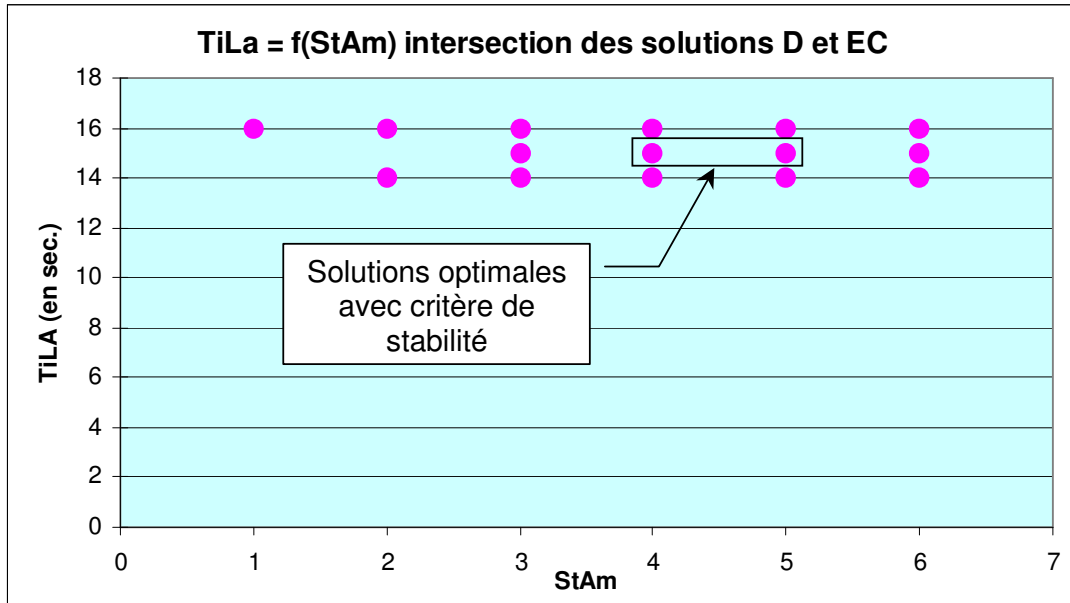


Figure 11 : Solutions retenues avec critère de stabilité

Les solutions retenues sont donc les suivantes :

TiLa = 15 sec. , StAm= 4

TiLa = 15 sec. , StAm= 5

Pour ces solutions, le délai de production est de 580 secondes (mini 577) et l'encours moyen est de 9,7 palettes (mini 8,86). En cas de variation de +/- 1 de l'un ou de l'autre des paramètres de pilotage, le critère de stabilité retenu pour ces 2 solutions garantit une stabilité du délai et de l'encours.

Ces premières analyses et cette méthodologie d'optimisation pourraient permettre de conclure qu'il est préférable de prévoir une capacité de stocks tampons importante pour ne pas avoir à se soucier du choix du paramètre de cadence de lancement (TiLa). Cette conclusion est toutefois à pondérer :

- Nous ne pouvons pas définir de capacité supérieure à 6 sur les stocks tampons, ce qui implique qu'on ne peut pas généraliser leur rôle en disant que l'amortissement du flux (donc la minimisation du délai) est directement proportionnel à la capacité des stocks tampons. En effet, il faudrait vérifier si l'amortissement du flux est linéaire et continu sur un plus grand domaine de

variation de la capacité des stocks. Il est possible qu'au delà d'une certaine valeur de capacité des stocks amont, on constate une dégradation du délai de fabrication.

- On travaille sur un routage et des temps de cycle sur les postes de travail définis de façon totalement arbitraire. Avec d'autres routages ou d'autres temps de cycle, on peut penser que la stabilité du délai n'est pas toujours linéaire par rapport à la capacité des stocks amont.

Nous reviendrons par la suite sur ce problème pour chercher une relation permettant de dimensionner la capacité des stocks amont (voir chapitre I.2-e).

I.2-d) Relations établies à partir de l'expérimentation N°1

Nous allons d'abord proposer une explication à la valeur de $TiLa$ de 16 secondes qui donne le délai (D) minimal, quel que soit la valeur de (StAm).

Comme on l'a vu, les courbes ***Délai = f(TiLa)*** font apparaître une valeur caractéristique de 16 secondes pour laquelle tous les minimums sont confondus, quelle que soit la valeur de la capacité du stock amont (StAm). Une analyse du système d'un point de vue "Flux matière" permet de donner une explication à cette valeur. En effet, si l'on assimile la ligne de production à un processus hydraulique, composé de réservoirs reliés par des tuyaux, le flux de palettes correspond donc à un débit de fluide circulant au travers de ces réservoirs et tuyaux. Le processus est alimenté en liquide avec un débit (dR) qui peut s'exprimer en volume/unité de temps. Sur la ligne, ce débit correspond au flux entrant de palettes, caractérisé par le Temps Inter-Lancement ($TiLa$). De la même façon, ce même processus se vide avec un débit (dV) exprimé dans la même unité que (dR) et qui correspond, sur la ligne, au flux sortant de palettes.

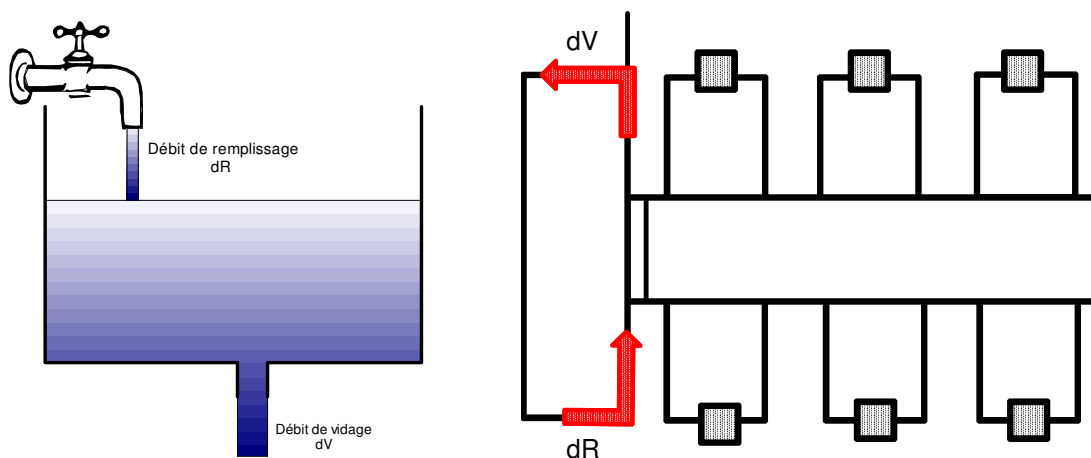


Figure 12 : Analogie débit hydraulique et flux de palettes

Plus précisément, un flux de production va circuler entre plusieurs postes de travail, ce qui va donc correspondre à plusieurs réservoirs en série se vidant les uns dans les autres. Par analogie, le débit de remplissage de chaque réservoir correspond au flux entrant de chaque poste de travail, donc à sa charge (voir partie 1, chap. I.2) et le débit de vidage correspond au flux sortant du poste de travail. Le flux sortant d'un poste de travail ne peut pas être supérieur à sa capacité (voir partie 1, chap. I.2). Si la charge est supérieure ou égale à la capacité, le débit de remplissage d'un réservoir (le flux entrant d'un poste de travail) correspond alors au débit de vidage du réservoir précédent (à la capacité du poste amont). Si il y a une accumulation de liquide dans un réservoir (des pièces devant un poste de travail), cela signifie que le débit de vidage est limité par les dimensions de l'orifice de vidage (le flux sortant est limité par la capacité du poste de travail). S'il n'y a pas d'accumulation dans un réservoir, les dimensions de l'orifice de vidage (la capacité du poste de travail) sont telles que le débit de vidage correspond au débit de remplissage (la capacité du poste de travail est supérieure ou égale à sa charge).

Cette analogie avec l'hydraulique met en évidence la notion de contrainte (ou "goulet d'étranglement") : le réservoir (le poste de travail) qui a le débit de vidage (la capacité) le plus faible et incapable d'absorber la charge demandée.

L'une des règles de la Théorie des Contraintes établit que le flux sortant maximal d'un système ne peut être supérieur à la capacité de la ressource goulet, quel que soit le flux entrant (voir partie 1, chap. IV.5).

L'utilisation de cette règle dans le cas de cette expérience permet d'identifier le poste de travail "goulet" : c'est le poste N°3, car son temps opératoire est le plus long (12 sec.).

La différence de 4 sec. qu'il y a entre ce temps opératoire et le TiLa optimal de 16 sec. provient du temps nécessaire à une palette pour aller du stock tampon amont au poste d'indexage (donc 4 secondes de temps mort entre chaque opération).

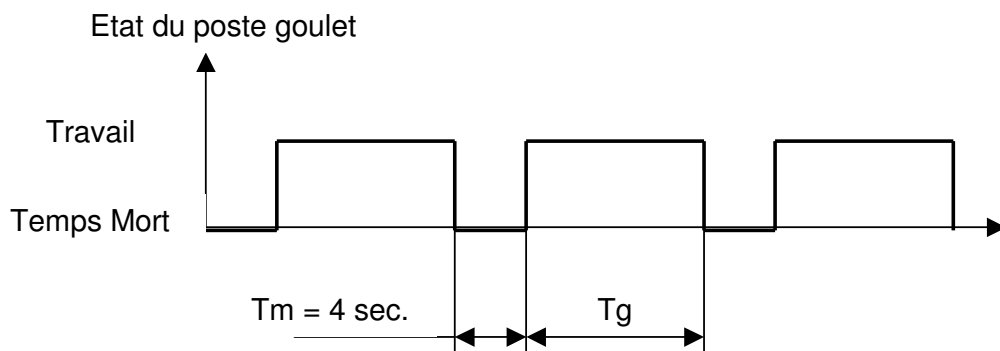


Figure 13 : Les 2 états du poste goulet

Cette règle permet de mettre en équation le délai de production dans un tel cas [FONTANILI 98]:

Soit :

D = Délai de production (en sec.)

Tg = Temps goulet = Temps opératoire le plus long

Tm = Temps mort = 4 sec.

TiLa = Temps inter-Lancement (en sec.)

Td = Temps de défilement de la première palette (en sec.) entre le point de lancement et le point de livraison

N = Nombre de palettes du lot de production

Si $TiLa \geq (Tg + Tm)$ alors

$$D = Td + [(Tg + Tm) \times (N-1)]$$

Sinon

$$D = \dots \text{???? impossible à mettre en équation facilement}$$

Il est à noter que dans le cas où $TiLa \geq (Tg + Tm)$, les stocks tampons ne seront jamais en accumulation car le flux entrant est inférieur ou égal à la capacité du poste goulet. Une augmentation du flux entrant ne peut pas provoquer d'augmentation du flux sortant, mais par contre certains stocks amont vont se remplir ce qui va provoquer une augmentation de l'encours. Les mesures obtenues pour des valeurs de TiLa inférieures à 16 secondes démontrent cela clairement, puisque toutes les meilleures combinaisons obtenues pour ces valeurs le sont avec des capacités de stocks tampons supérieures à 1 et avec des encours moyens d'environ 12,5 palettes.

Application à l'expérience N°1 :

Td = 193 sec.

N= 25 palettes

Tg = 12 sec.

Tm = 4 sec.

TiLa = 12 + 4 = 16 sec.

$$D = 193 + (12 + 4) \times (25 - 1) = 577 \text{ sec.}$$

ce qui correspond bien au délai obtenu en pratique avec un TiLa de 16 sec.

L'intérêt de cette équation ne réside pas dans le choix des paramètres de pilotage, puisque comme on l'a démontré précédemment, il ne semble pas être judicieux d'opter pour un TiLa de 16 secondes pour garantir la stabilité du délai et de l'encours.

Par contre, cette équation permet de déterminer très rapidement et très simplement le délai minimal théorique que l'on est en mesure d'obtenir pour la production d'un lot de produits identiques. Ce délai minimal théorique peut donc ensuite servir de référence pour une recherche des paramètres optimum de pilotage avec un critère de stabilité.

Nous proposons maintenant une mise en équation du nombre optimal de palettes à utiliser (N_p) pour assurer une production sans attente (voir partie 5, chap. IV.5) tout en minimisant l'encours.

Sachant que l'expérimentation N°1 a été réalisée avec un nombre de palettes égal à la quantité de palettes en réserve ($N_p = R_p$), nous pouvons faire la remarque suivante :

- Dans le cas où $TiLa = 16$ sec. et $StAm=1$, la courbe de l'encours = $f(\text{temps})$ montre que la quantité de palettes nécessaire est de 12 (variation entre 11 et 12 de l'encours pendant la phase de régime nominal). L'encours est mesuré entre le point de lancement et le point de livraison. On ne tient donc pas compte, dans l'encours, des palettes présentes sur la ligne de déchargement, entre le point de livraison et le point de déchargement. Pour cette valeur de $TiLa$ (et pour toutes les valeurs de $StAm$), le nombre de palettes peut être déterminé par la relation suivante :

Soit :

T_d = temps de défilement (en sec.)

$T_g = TiLa$ = Temps goulet (en sec.)

T_m = Temps mort sur poste de travail (= 4 sec.)

$N_{p \text{ mini}}$ = réserve minimale de palettes

$$N_{p \text{ mini}} = T_d / (T_g + T_m)$$

Application numérique :

$T_d = 193$ sec.

$T_g + T_m = 12 + 4 = 16$ sec.

$$N_{p \text{ mini}} = 193 / (12 + 4) = 12 \text{ palettes}$$

Si l'on tient compte du temps de déchargement de chaque palette, il faut, dans ce cas, rajouter au temps de défilement le temps nécessaire à l'acheminement de la palette du

point de livraison au poste de déchargement et au déchargement proprement dit, soit environ 37 sec. On calcule donc le nombre total de palettes y compris celles circulant sur la ligne de déchargement en rajoutant 37 secondes au temps de défilement (T_d).

Application numérique :

$$N_{p\text{mini}} = (193 + 37) / (12 + 4) = 15 \text{ palettes}$$

1.2-e) Expérimentation avec temps de cycle différents

Nous avons constaté précédemment que des capacités de stocks amont élevées (à 6) ont la propriété d'amortir les variations de délai, pour des valeurs de Temps inter-Lancement ($TiLa$) comprises entre 1 et ($T_g + T_m$). Nous allons chercher à vérifier cette propriété et éventuellement la généraliser.

Pour l'expérimentation N°1, les temps opératoires fixés arbitrairement sur les postes de travail, et en particulier sur le poste goulet, sont d'une durée relativement grande par rapport aux valeurs possibles de Temps inter-Lancement ($TiLa$) pour permettre une accumulation des palettes dans les stocks amont.

Nous allons donc procéder à une série d'expérimentations identiques à la précédente, en changeant simplement les valeurs des temps opératoires, le routage restant le même.

- Série 1 : **P2[1] – P3[4] – P5[3] – P6[2]** ($T_d = 173$ secondes)

Les temps opératoires sont différents, mais le taux d'équilibrage est identique à l'expérimentation N°1 (voir 1.2-a) :

$$E_b = (1 + 4 + 3 + 2) / (4 \times 4) = 0,625$$

A partir des relations établies dans le chapitre précédent, on peut déterminer :

- le Temps inter-Lancement optimal : $TiLa = T_g + T_m = 4 + 4 = 8$ secondes.
- le délai théorique : $D = 173 + (25 - 1) \times (4 + 4) = 365$ secondes.
- le nombre optimal de palettes à utiliser : $N_p = (T_d + 37) / (4 + 4) = 27$ palettes.

On sait maintenant qu'il est inutile de faire des simulations pour des valeurs de $TiLa$ supérieures à $T_g + T_m$. Nous réalisons donc des simulations en faisant varier $TiLa$ entre 1 et 10 secondes (au lieu de 20 dans l'expérimentation précédente).

Les courbes de la figure 14 représentent la variation du délai de production (D) en fonction du Temps inter-Lancement ($TiLa$), pour chaque valeur de capacité du stock Amont ($StAm$).

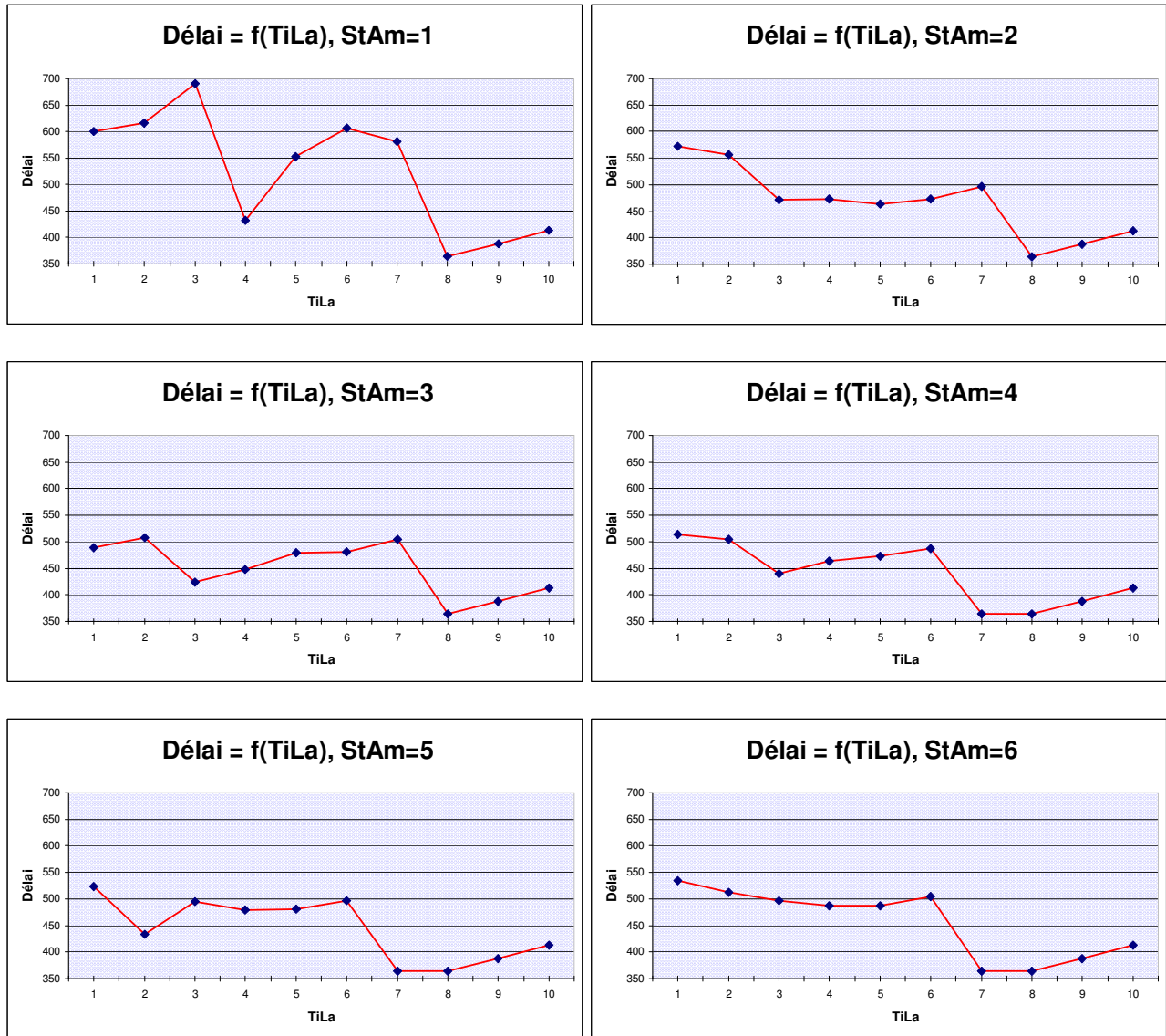


Figure 14 : Courbes $D = f(TiLa)$ pour chaque valeur de $StAm$

Le graphique de la figure 15 est une superposition des courbes précédentes et met en évidence l'influence des stocks amont ($StAm$) et du Temps inter-Lancement ($TiLa$) sur le délai (D).

On peut constater que dans cette expérimentation, contrairement à la précédente, l'augmentation de la capacité du stock amont n'entraîne pas les mêmes effets. Il y a un amortissement de la courbe $D = f(TiLa)$, mais il ne se situe pas à la valeur minimale du délai.

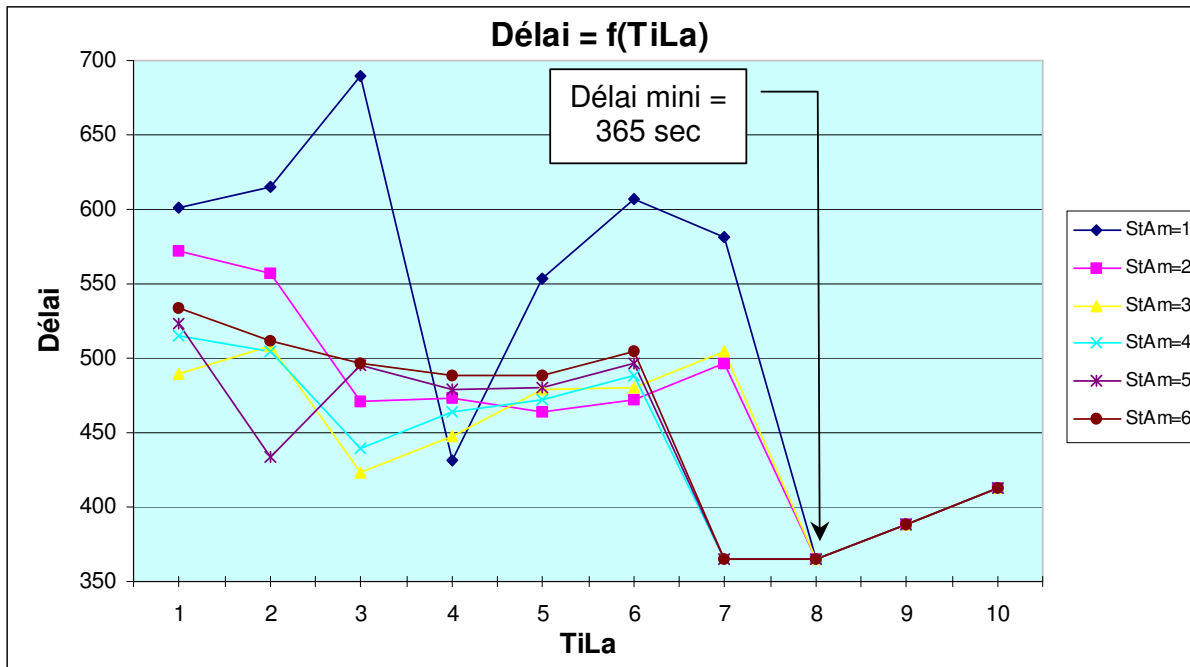


Figure 15 : Graphique de synthèse des résultats

L'observation et l'analyse des résultats des deux expérimentations précédentes nous amènent à chercher une relation faisant intervenir les temps opératoires $T_{cyc}(i)$, peut être en particulier le temps du poste goulet ($T_g = \text{Max}(T_{cyc}(i))$), et la capacité des stocks amont ($StAm$), pour garantir un amortissement de la courbe du délai (D) à la valeur minimale.

Pour répondre à ce problème, il faut analyser de façon plus détaillée le fonctionnement d'un poste en dérivation (figure 16).

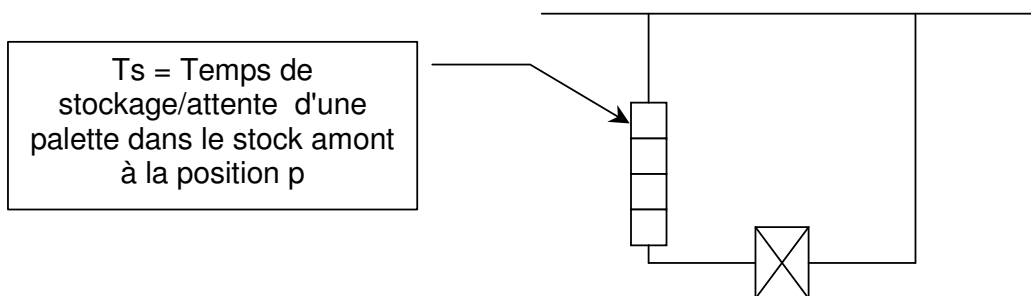


Figure 16 : Temps d'attente dans le stock amont

Soit :

T_s = Temps de stockage/attente d'une palette dans le stock amont

p = position d'arrivée de la palette dans le stock amont ($p = 1 \dots StAm$)

$T_{cyc}(i)$ = Temps de cycle du poste i ($i = 1..6$)

T_m = Temps mort entre le travail de deux palettes successives

On peut donc écrire la relation suivante :

$$T_s = p \cdot (T_{cyc}(i) + T_m)$$

En particulier, pour le poste goulet et si $p = StAm$, le temps d'attente/stockage maximal d'une palette dans le stock amont est :

$$\text{Max}(T_s) = StAm \cdot (T_g + T_m) \quad (1)$$

avec T_g = Temps de cycle du poste goulet.

Pour assurer un fonctionnement optimal du poste goulet, il est indispensable que celui-ci soit alimenté sans aucune rupture, en dehors des temps morts entre deux palettes. Or, dans le cas où la capacité maximale du stock amont est atteinte, une palette qui devait rentrer sur le poste de travail devra en fait reboucler sur la ligne principale (figure 17).

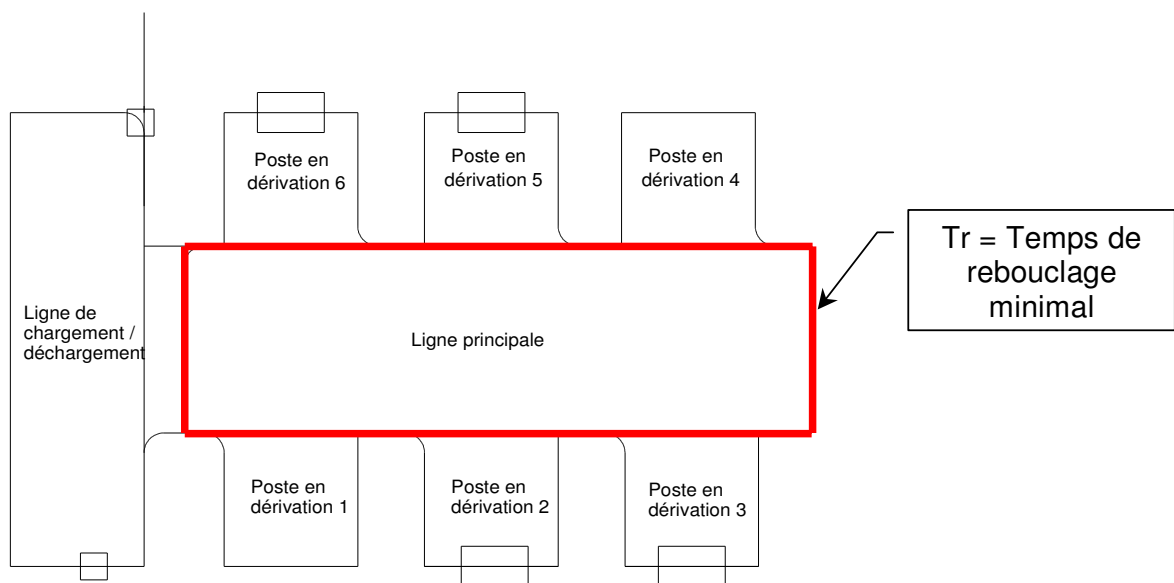


Figure 17 : Temps de rebouclage minimal d'une palette

Soit T_r = Temps de rebouclage minimal d'une palette sur la ligne = 94 secondes.

Le Temps de rebouclage (T_r) est minimal si la palette est prioritaire quand elle passe devant la sortie de chaque poste en dérivation (voir partie 5, chap. IV.4). En fonctionnement à régime nominal, quand plusieurs palettes circulent sur l'ensemble des postes de travail et sur la ligne principale, le temps de rebouclage minimal est rarement atteint à cause de la règle de priorité en sortie de poste en dérivation qui implique un arrêt, donc une attente, de la palette circulant sur la ligne principale.

A partir de ces observations, on peut distinguer les deux cas suivants :

Si $\text{Max}(T_s) < T_r$

Alors risque d'attente du poste

Sinon ($\text{Max}(T_s) \geq T_r$)

Pas d'attente du poste : Fonctionnement optimal

On peut donc, à partir de ce constat et de l'équation (1) déduire une relation permettant de dimensionner le stock amont du poste goulet afin d'obtenir un fonctionnement optimal :

$$\text{StAm}(i) \geq \frac{T_r}{(T_{\text{cyc}}(i) + T_m)}$$

avec i = numéro du poste goulet

Si l'on applique cette relation à la première expérimentation, pour laquelle $i = 3$ et $T_{\text{cyc}}(3) = T_g = 12$ secondes, on obtient :

$$\text{StAm}(3) \geq 94 / (12 + 4) = 5,87$$

Ce résultat explique donc pourquoi avec un stock amont fixé à 6 sur tous les postes de travail, nous avons obtenu un délai de fabrication minimal, avec un excellent amortissement quelle que soit la valeur du Temps inter-Lancement (T_{La}).

Pour la deuxième expérimentation, on a $i = 3$ et $T_{\text{cyc}}(3) = T_g = 4$ secondes, on obtient :

$$\text{StAm}(3) \geq 94 / (4 + 4) = 11,75$$

On voit donc qu'avec une capacité du stock amont ne pouvant pas être supérieure à 6, il est impossible d'obtenir une stabilité du délai à son niveau minimal quelle que soit la valeur du Temps inter-Lancement (T_{La}).

On peut transformer la relation précédente afin de savoir, pour une capacité du stock amont fixée, en dessous de quelle valeur le temps de cycle du poste goulet entraîne une dégradation du délai :

$$T_{cyc}(i) \geq \frac{Tr}{StAm(i)} - T_m \quad (2)$$

avec $i = N^\circ$ du poste goulet

Dans le cas qui correspond à celui des expérimentations précédentes, on obtient donc :

$$T_{cyc}(i) \geq (94 / 6) - 4 = 11,66 \text{ secondes}$$

Dans tous les cas de production où le temps de cycle du poste goulet sera supérieur ou égal à 11,66 secondes et où le stock amont sera égal à 6, on obtiendra un délai minimal de production qui sera stable pour toutes valeurs du Temps inter-Lancement (TiLa) inférieure ou égale à $T_g + T_m$.

Afin de vérifier ces relations, nous réalisons une nouvelle expérimentation en prenant le même routage mais en fixant le temps de cycle du poste goulet à une valeur juste inférieure à la valeur déterminée par la relation (2) : $T_{cyc}(3) = T_g = 10$ secondes.

- Série 2 : **P2[1] – P3[10] – P5[3] – P6[2]** ($T_d = 179$ secondes)

Le taux d'équilibrage est plus faible que pour les expérimentations précédentes :

$$Eb = (1 + 10 + 3 + 2) / (4 \times 10) = 0,40$$

A partir des relations établies dans le chapitre précédent, on peut déterminer :

- le Temps inter-Lancement optimal : $TiLa = T_g + T_m = 10 + 4 = 14$ secondes.
- le délai théorique : $D = 179 + (25 - 1) \times (10 + 4) = 515$ secondes.
- le nombre optimal de palettes à utiliser : $N_p = (T_d + 37) / (10 + 4) = 16$ palettes.

Le graphique de la figure 18 présente les résultats de cette expérimentation.

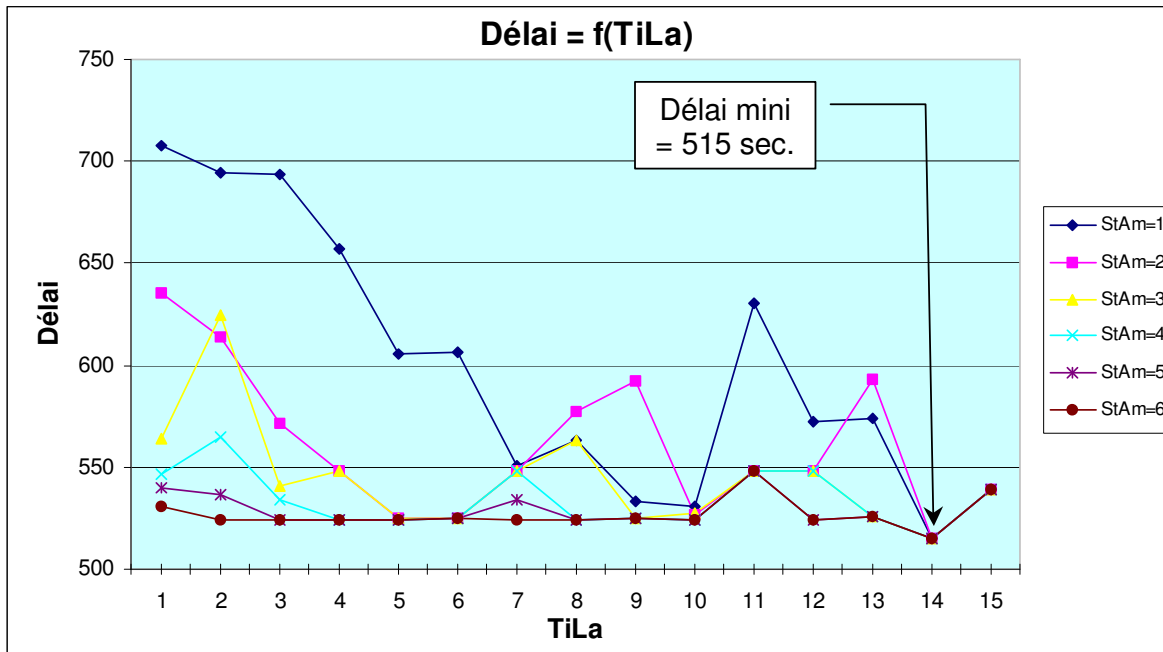


Figure 18 : Courbes $D = f(TiLa)$

On constate dans cette série un amortissement presque parfait de la courbe (StAm = 6) sur la valeur minimale du délai (515 sec.), ce qui permet de vérifier la relation (2).

La conclusion de ces observations et relations est la suivante :

- Pour une ligne de production existante, déjà dimensionnée au niveau des stocks amont, si le temps de cycle du poste goulet respecte la relation (2), on peut obtenir un délai de fabrication minimal pour toutes valeurs de TiLa inférieures ou égales à $(Tg + Tm)$. Le choix d'une solution particulière parmi toutes celles possibles peut se faire en utilisant le critère de stabilité présenté dans le chapitre I.2-c. Si par contre la relation (2) n'est pas respectée, l'obtention d'un délai minimal n'est garantie que pour certaines valeurs de TiLa, à priori proche de $(Tg + Tm)$. De la même façon, la stabilité de la solution est incertaine. C'est donc dans ce cas précis où la relation (2) n'est pas vérifiée que l'optimisation est plus délicate à résoudre.
- Pour une ligne à concevoir, l'utilisation des relations précédentes permet de dimensionner les stocks amont, éventuellement la longueur de la boucle principale (pour fixer le temps de rebouclage Tr) et le nombre de palettes à utiliser.

I.2-f) Expérimentation avec temps de cycle équilibrés

On propose maintenant d'étudier l'influence d'un équilibrage parfait, en théorie, des temps de cycle des différentes opérations. On conserve toujours le même routage et on prend comme temps de cycle une valeur inférieure au seuil critique calculé par la relation (2) du chapitre précédent, par exemple 4 secondes pour comparer avec l'expérimentation précédente :

- série 3 : **P2[4] – P3[4] – P5[4] – P6[4]** (Td = 179 secondes)

Le taux d'équilibrage est donc de 100%.

Le graphique de la figure 19 présente les résultats.

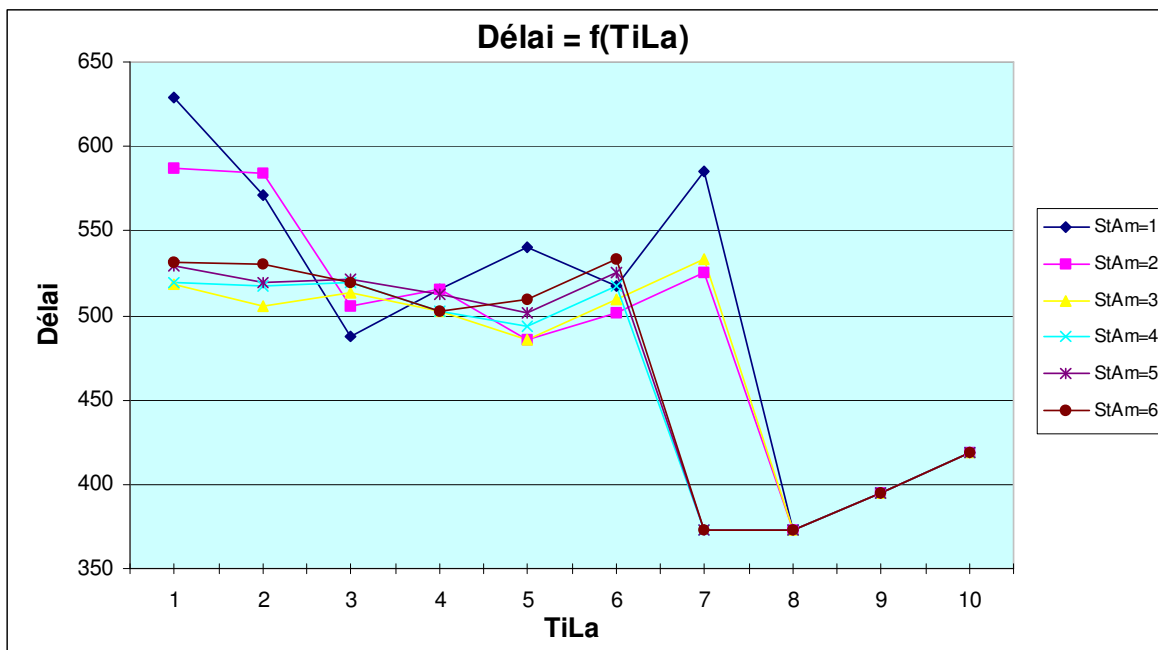


Figure 19 : Courbes $D = f(TiLa)$

En comparant ce graphique à celui de la figure 15, on remarque une allure générale très proche entre les différentes courbes. On peut donc, sur ce cas, conclure que l'équilibrage des temps de cycle de la ligne ne simplifie pas particulièrement le choix des paramètres de pilotage.

□ I.3- Expérimentation N°2

I.3-a) Rappel des conditions propres à cette expérimentation

- Pas de phénomènes aléatoires.
- Horizon de fabrication INFINI : la ligne est toujours en charge.
- Mesure du délai (D_{100}) écoulé entre la 50^{ème} et la 150^{ème} palette, en régime nominal.
- Les temps de cycle sont fixés arbitrairement aux valeurs suivantes :

P2[7] – P3[12] – P5[8] – P6[3]

I.3-b) Résultats

Le graphique de la figure 20 représente la variation du délai de production de 100 palettes en régime nominal (D_{100}) en fonction du Temps inter-Lancement (TiLa), pour chaque valeur de capacité du stock amont (StAm).

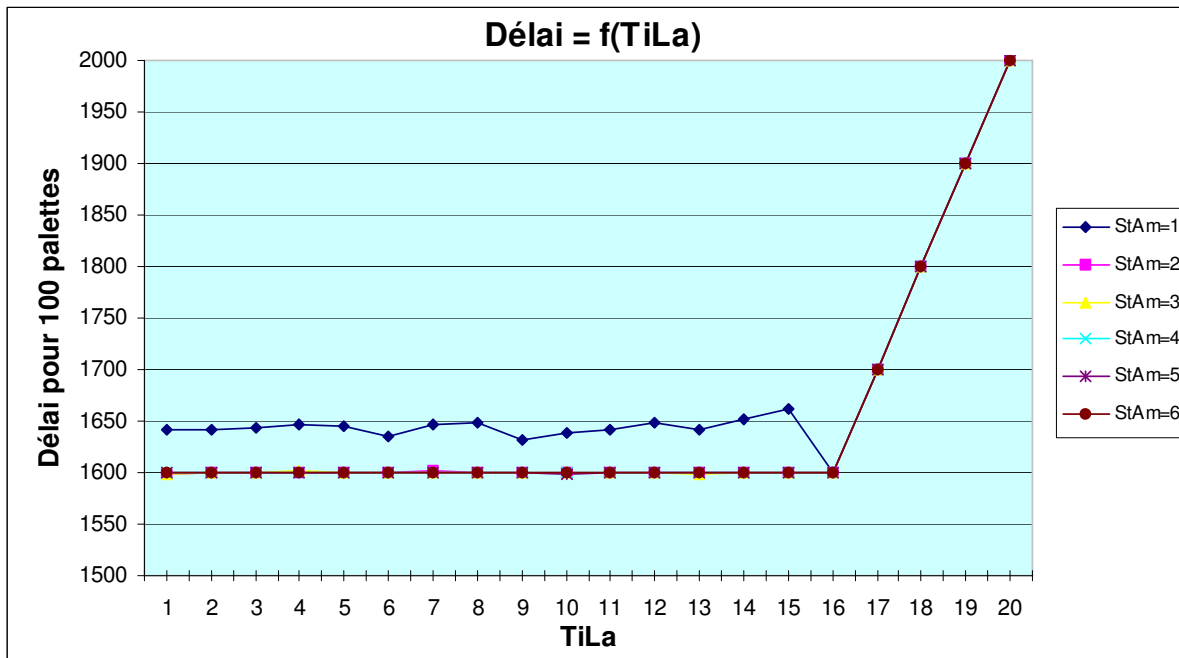


Figure 20 : Courbes $D_{100} = f(TiLa)$

On constate sur le tracé de StAm=1 une très faible variation du délai pour des valeurs de TiLa comprises entre 1 et 15 sec. Ce tracé est distinct des autres, qui sont eux pratiquement confondus pour Tila compris entre 1 et 16 sec., quelle que soit la valeur de

(StAm). Le délai moyen (D_{100}) obtenu pour tous ces tracés est de 1600 secondes. Les conclusions de cette expérimentation font apparaître que, sauf pour une valeur de (StAm) de 1, le délai de fabrication est constant, quelle que soit la valeur de TiLa choisie entre 1 et 16 sec. et pour des capacités des stocks amont (StAm) comprises entre 2 et 6. Ce constat est intéressant puisqu'il permet de s'affranchir d'un choix précis des variables TiLa et StAm en horizon de fabrication infini.

Cette stabilité des résultats peut s'expliquer par le fait que le système s'autorégule : en effet, étant donné que le nombre de palettes est fini ($N_p = 25$), le lancement d'une palette est conditionné par la présence et la disponibilité d'une palette revenue au poste de déchargement après être passée sur les postes nécessaires à la réalisation de la gamme suivie. La cadence de lancement réelle des palettes peut donc être très différente de celle qui est demandée, puisqu'elle va correspondre à la cadence de livraison des palettes qui sont déchargées. Cette cadence de livraison est mesurée par le Temps inter-Livraison moyen (TiLi) qui est d'environ 16 secondes, et qui correspond au Temps de cycle du poste goulet ($T_g = 12$ sec.) plus le temps mort (T_m) de 4 secondes déjà décrit précédemment.

Le schéma de la figure 21 illustre le flux de palettes suivant ce principe d'autorégulation :

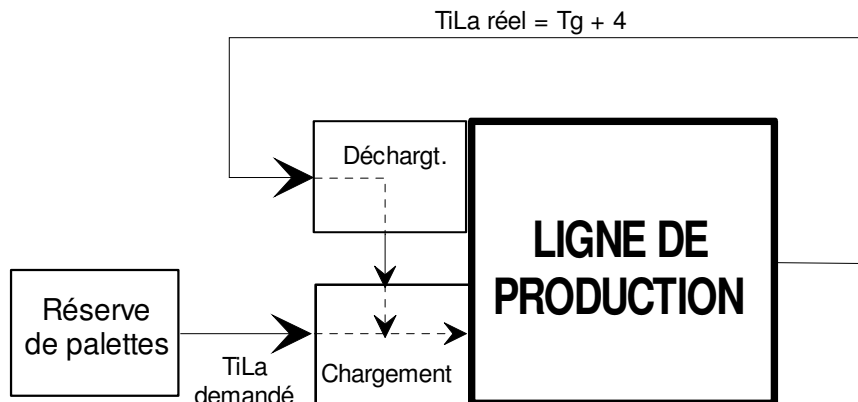


Figure 21 : Principe d'autorégulation du flux

Cette autorégulation est assez similaire à un système fonctionnant en **KANBAN** [SHINGO 86], le nombre de palettes en réserve jouant le rôle des étiquettes dans le système Kanban. La différence notable avec le Kanban est que l'ordre de fabrication, sous la forme d'un retour de palette, est introduit directement en amont de la ligne par le déchargement d'une palette, alors que dans le système Kanban, chaque poste de travail émet un ordre vers le poste de travail immédiatement en amont à chaque fabrication (voir partie 1, chap. IV.2). Le Kanban régule le niveau des encours entre deux postes de travail en fonction de la consommation réelle du poste aval, alors que le fonctionnement obtenu sur la ligne régule le flux d'entrée en fonction du flux de sortie réel. Cette autorégulation d'une ligne de production à transfert asynchrone s'apparente au principe du **CONWIP** "CONstant Work In

Process", présenté dans [SPEARMAN 90]. L'expérimentation réalisée dans le cas d'un horizon de production infini permet de vérifier ce principe. Il reste à déterminer, comme pour le Kanban, le nombre d'étiquettes, donc de palettes (N_p), à mettre en circulation.

Le graphique de la figure 22 montre les valeurs de l'encours moyen (EC) pour chaque capacité du stock amont (StAm). Lors de cette expérimentation, le nombre de palettes (N_p) est de 25 (correspond à la réserve de palettes R_p). On voit sur le graphique que l'encours est compris entre 21 et 22 palettes pour toutes les valeurs de (StAm) et pour toutes les valeurs de (TiLa) entre 1 et 10. Puis l'encours moyen chute à 11,3 palettes pour $TiLa = 16$ secondes. L'encours continue à baisser pour des valeurs supérieures de (TiLa), mais si l'on observe parallèlement la courbe $D_{100} = f(TiLa)$ de la figure 20, on voit que le délai est minimal pour toutes valeurs de TiLa inférieures ou égales à 16 secondes.

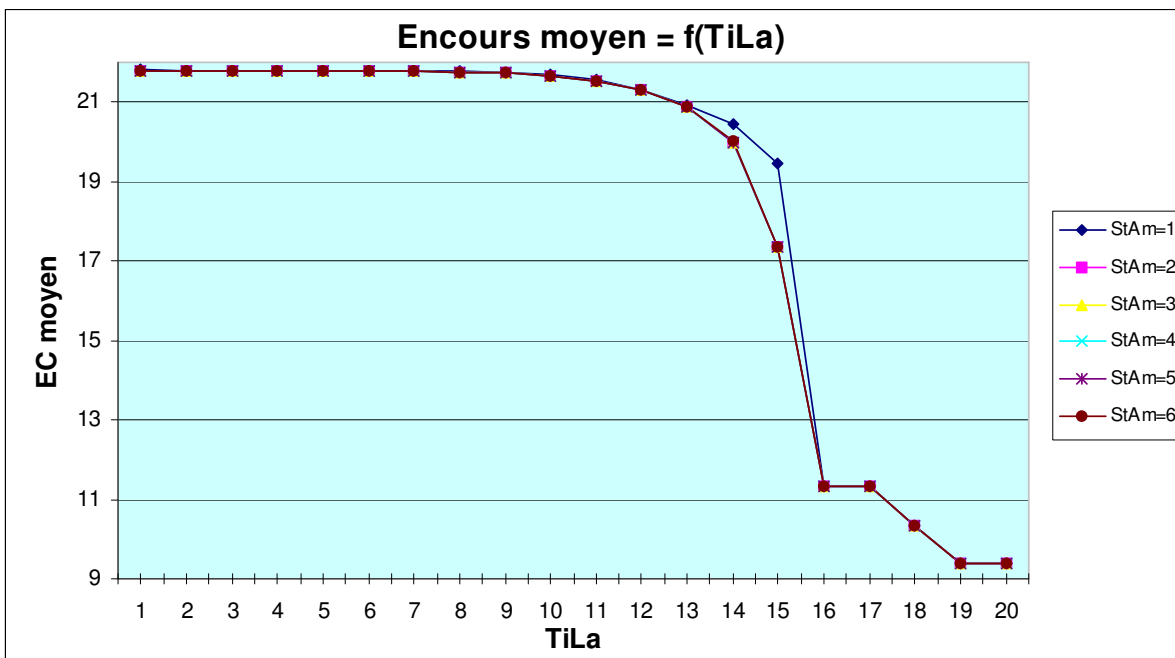


Figure 22 : EC moyen = f(TiLa)

L'analyse précédente fait donc ressortir qu'une autre variable est à prendre en considération dans un cas de production à horizon infini : Le nombre de palettes à utiliser (N_p).

Si l'on revient sur l'expérience initiale réalisée avec 25 palettes, le graphique de la figure 23 montre la variation du ratio de fluidité (Rf) qui est le rapport entre la somme des temps de cycle et le temps de défilement (voir partie 5, chap. VI.4). Etant donné que cette expérimentation correspond à une production monoproduit, le ratio de fluidité est uniquement fonction du temps de défilement de chaque palette. Le graphique montre une forte instabilité du ratio d'une palette à l'autre, ce qui signifie que certaines palettes rebouclent une ou plusieurs fois sur la ligne principale, pendant que d'autres suivent leur routage sans rebouclage.

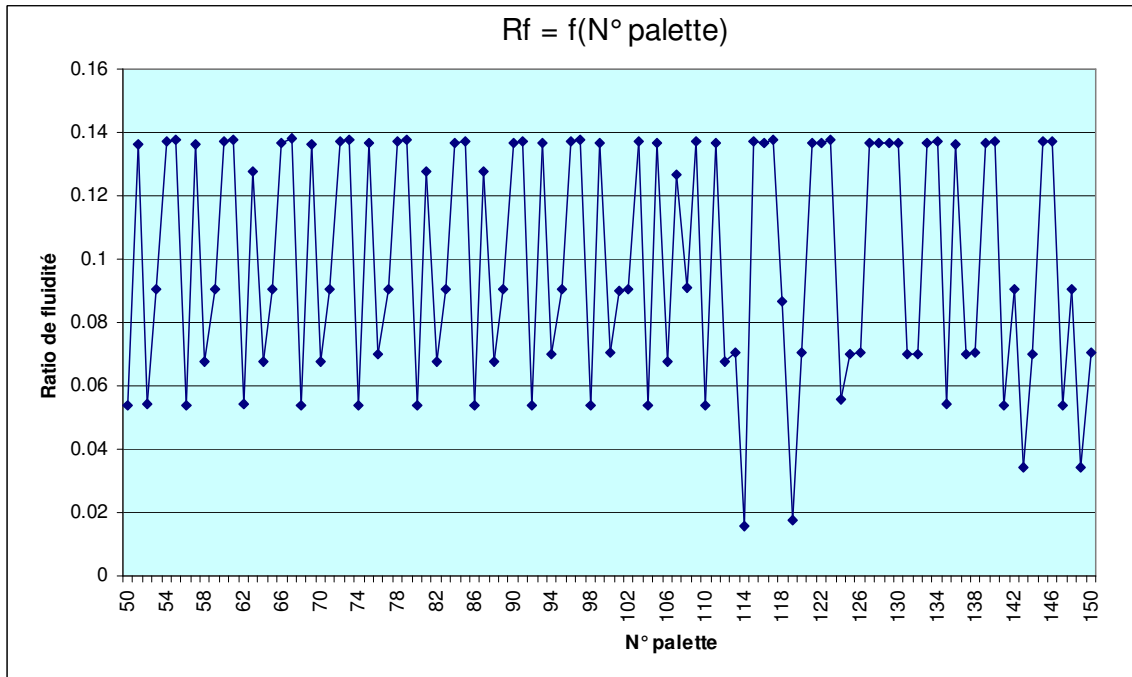


Figure 23 : Variation du Ratio de fluidité ($N_p = 25$ palettes)

En prenant pour (N_p) une valeur de 15, qui est issue de la relation établie dans l'expérimentation N°1 (voir chap. I.2-d), le délai (D_{100}) minimal est le même que pour 25 palettes (soit 1600 secondes), seule la courbe pour $StAm = 1$ diffère en étant moins stable (voir figure 24).

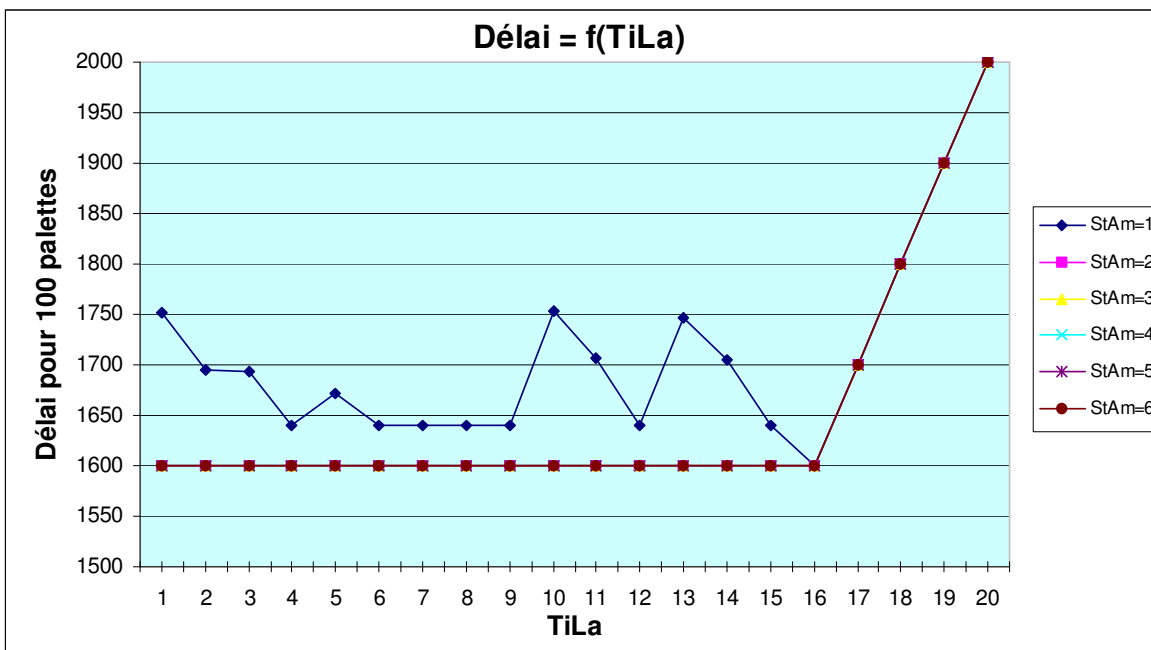


Figure 24 : $D_{100} = f(TiLa)$ pour $N_p = 15$ palettes

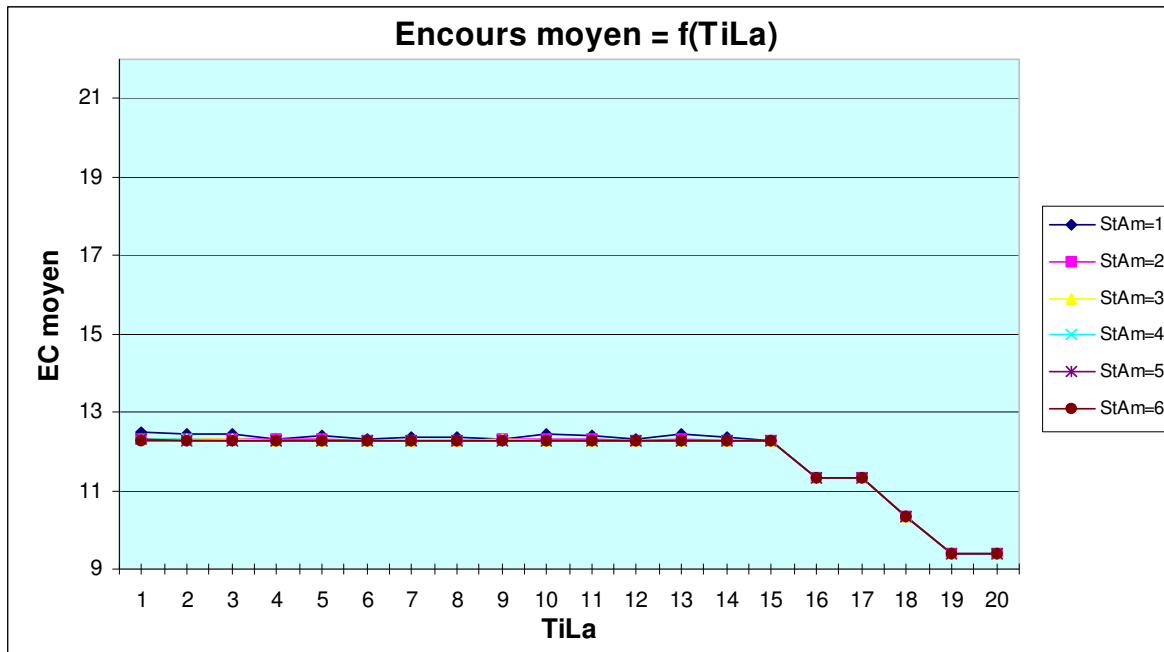


Figure 25 : EC moyen = f(TiLa)

On voit très nettement sur le graphique de la figure 25 que l'encours moyen (EC) est compris entre 12 et 13 palettes pour toutes les valeurs de (StAm) et pour toutes les valeurs de (TiLa) inférieures ou égales à 16.

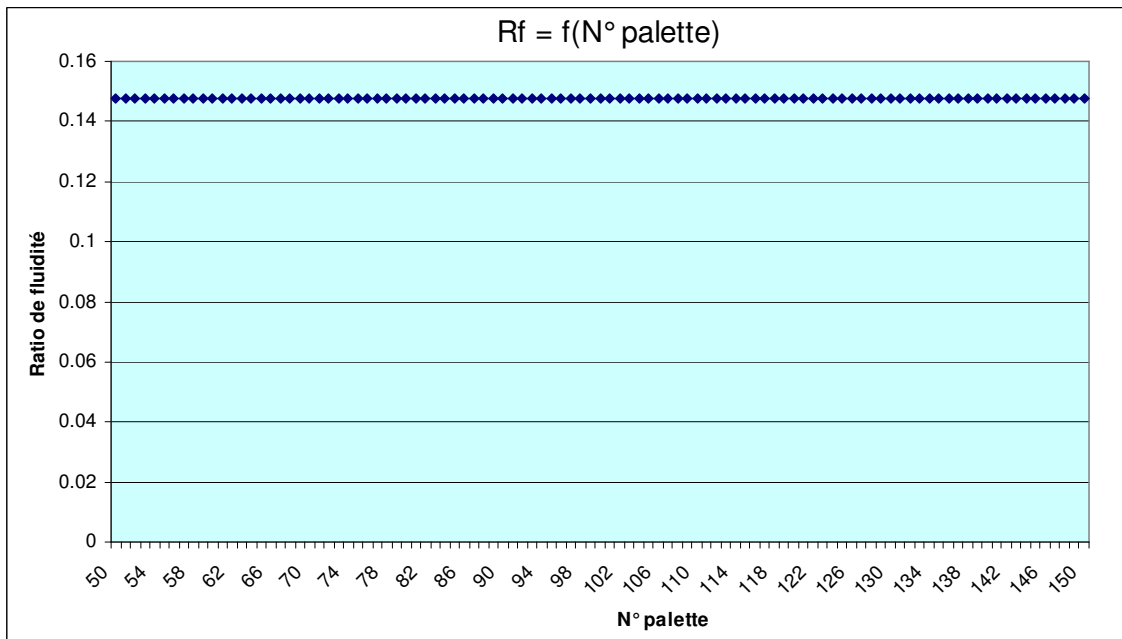


Figure 26 : Rf = f(N° palette)

En conclusion, on obtient donc exactement le même délai (D_{100}) que dans l'expérimentation précédente, mais en utilisant 15 palettes au lieu de 25. L'encours moyen EC est donc plus faible et le tracé du ratio de fluidité $R_f = f(N^\circ \text{ palette})$ de la figure 26 montre une régularité parfaite du flux.

La détermination du nombre de palettes à utiliser est donc un paramètre important. Un nombre trop faible de palettes provoque une augmentation du délai car on ne sature pas le poste goulet. Un nombre trop important de palettes perturbe le flux et entraîne une augmentation de l'encours moyen.

II- Analyse des résultats du jeu d'essais N°2

□ II.1- Rappel des conditions expérimentales générales

- Lancement de type "MULTIPRODUIT".
- Nombre de palettes en réserve : $R_p = 25$, inférieur ou égal à la quantité totale de palettes à lancer.
- Séquencement des opérations dans un ordre imposé, sans redondance avec :
 $P_{i[j]}$ = opération sur le poste i ($i=1\dots6$) pendant un temps de cycle $T_{cyc}(i,k)$ de j secondes pour l'ordre de fabrication k ($k = 1..n$).

□ II.2- Expérimentation N°1

II.2-a) Rappel des conditions propres à cette expérimentation

- Horizon de fabrication FINI : la ligne est vide en fin de simulation.
- Lancement d'une campagne de 5 ordres de fabrication successifs comportant un total de 44 palettes.
- La matrice des flux de la figure 27 précise les temps de cycle et les tailles de lot (valeurs fixées arbitrairement) de chaque ordre de fabrication, avec :

$T_{lot}(k)$ = Taille de lot de l'ordre de fabrication N° k ($k = 1..5$),

$T_{cyc}(i,k)$ = Temps de cycle du poste i ($i = 1..6$) pour une palette de l'ordre de fabrication N° k ($k = 1..5$),

$T_d(k)$ = Temps de défilement optimal d'une palette de l'O.F N° k ($k = 1..5$),

$R_f(k)$ = Ratio de fluidité optimal d'une palette de l'O.F N° k ($k = 1..5$)

	Tlot (k)		Tcyc(1.k)	Tcyc(2.k)	Tcyc(3.k)	Tcyc(4.k)	Tcyc(5.k)	Tcyc(6.k)	Td(k)	Rf(k)
	Taille Lot		Poste 1	Poste 2	Poste 3	Poste 4	Poste 5	Poste 6	(sec.)	optimal
OF1	10	Gamme A	4	2			3		172	0.052
OF2	5	Gamme B		1		4	3	2	173	0.058
OF3	8	Gamme C	3	4		5			175	0.069
OF4	15	Gamme D				3	4	5	175	0.069
OF5	6	Gamme E	1	2				3	169	0.036

Figure 27 : Matrice des flux du jeu d'essais N°2

Le temps de défilement optimal $Td(k)$ et le ratio de fluidité optimal $Rf(k)$ sont obtenus en lançant une seule palette sur la ligne.

Le séquençement des O.F au niveau du lancement est représenté sur la figure 28.

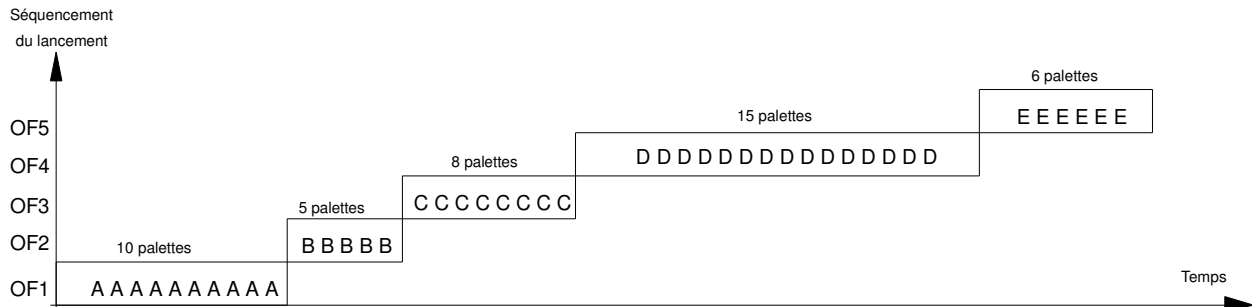


Figure 28 : Séquençement du lancement des O.F

Remarque : La figure 28 permet de soulever le problème du choix de la séquence optimale de lancement. On rappelle que pour ce jeu d'essais et pour les expérimentations réalisées dans cette thèse, on fait l'hypothèse que cette séquence est imposée : A puis B puis C puis D puis E.

Cela étant, le choix de la séquence de lancement semble être une piste à explorer pour de futures recherches.

II.2-b) Résultats

Les courbes de la figure 29 représentent la variation du délai de production de la campagne (D) en fonction du Temps inter-Lancement ($TiLa$), dans le cas où l'on utilise toutes les palettes en réserve ($Np = Rp = 25$).

En effet, on constate que pour ce jeu d'essais, le nombre total de palettes de la campagne (44) est supérieur au nombre de palettes en réserve (25). Toutes les palettes en réserve sont donc utilisées.

Nous allons tout d'abord réaliser une première expérimentation dans les conditions énoncées précédemment, sans tenir compte des relations établies pour une production de type monoproduit. Nous tenterons ensuite d'appliquer certaines de ces relations pour voir si elles sont généralisables à des productions de type multiproduit.

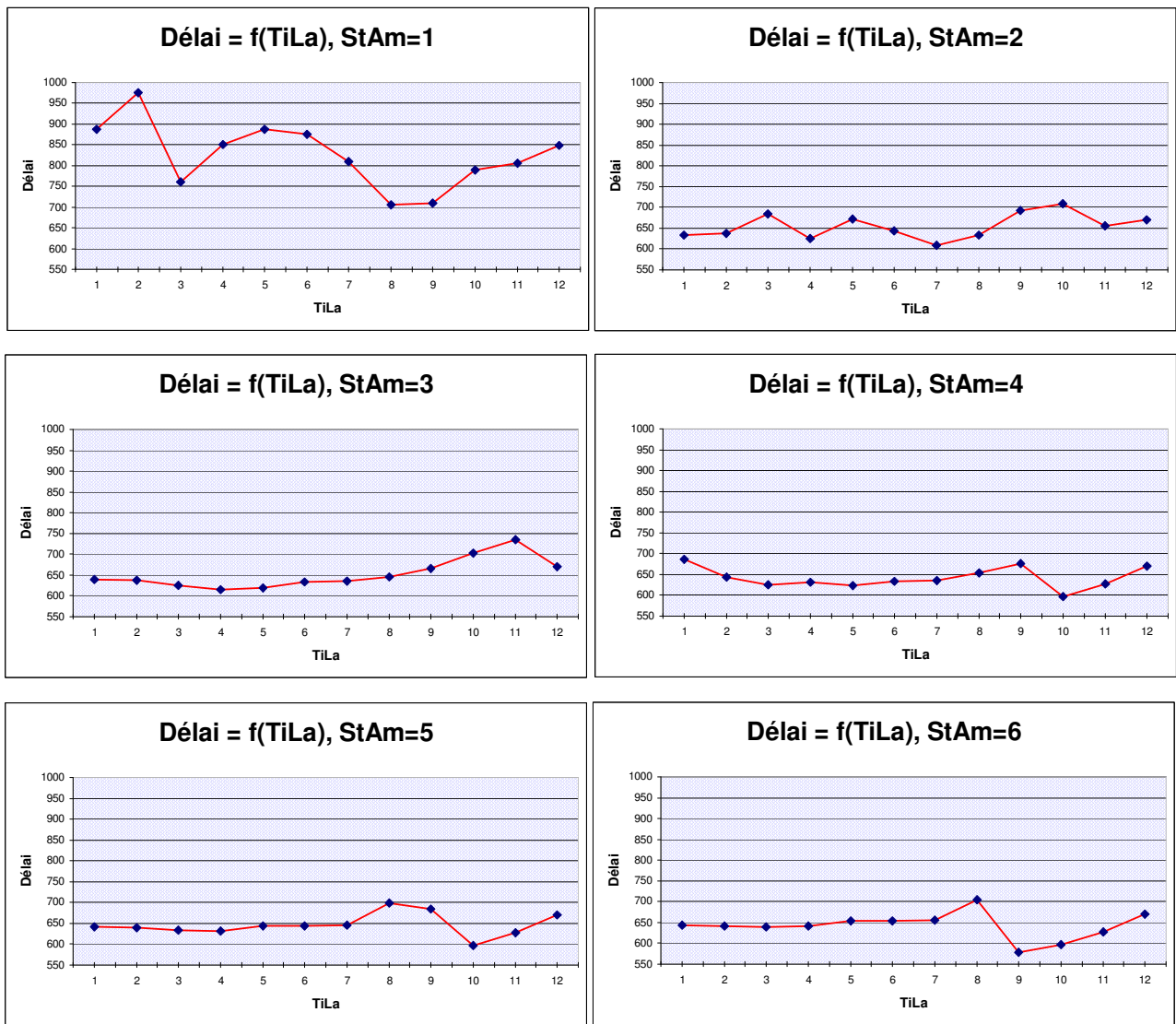


Figure 29 : Courbes $D = f(TiLa)$ pour chaque valeur de StAm

Le délai minimal (579 sec.) est obtenu pour une seule combinaison de (StAm) et de (TiLa) :
Solution unique : StAm = 6, TiLa = 9

Le graphique de la figure 30 est une superposition des courbes précédentes et met en évidence l'influence des stocks amont (StAm) et du Temps inter-Lancement (TiLa) sur le délai (D).

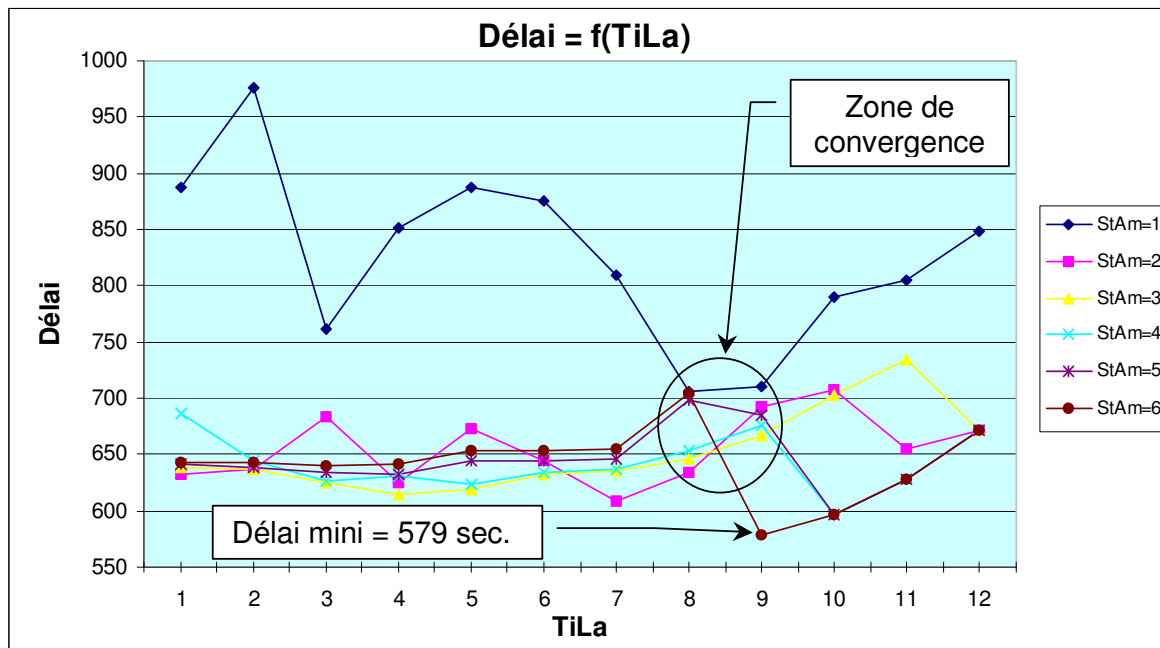


Figure 30 : Graphique de synthèse des résultats

II.2-c) Analyse des données récoltées et comparaison au cas monoproduit

Le premier constat fait en analysant le graphique de la figure 30 est que toutes les courbes ne convergent pas vers un même délai minimum et unique, contrairement à ce que l'on obtenait dans le cas d'une production monoproduit (voir par exemple le graphique de la figure 3). On peut seulement distinguer une zone de convergence où il semble y avoir un resserrement des valeurs de l'ensemble des courbes ($610 \text{ sec.} \leq D \leq 734 \text{ sec.}$, pour $TiLa$ compris entre 8 et 9). Toutefois, cette zone n'englobe pas le délai mini ($D = 579 \text{ sec.}$, pour $TiLa = 9$ et $StAm = 6$), mais elle s'y trouve à proximité. Ce premier constat laisse donc apparaître une difficulté plus importante que dans une production monoproduit pour déterminer le délai de production minimal.

On peut constater quand même certaines similitudes avec le jeu d'essais précédent. En particulier, le rôle stabilisateur des stocks amont quand la valeur de ceux-ci est supérieure à 1 et pour des valeurs de ($TiLa$) situées à gauche de la zone de convergence.

Nous allons maintenant tenter d'appliquer au cas multiproduit le calcul du délai qui a été utilisé dans le jeu N°1.

La figure 31 rappelle le principe de ce calcul :

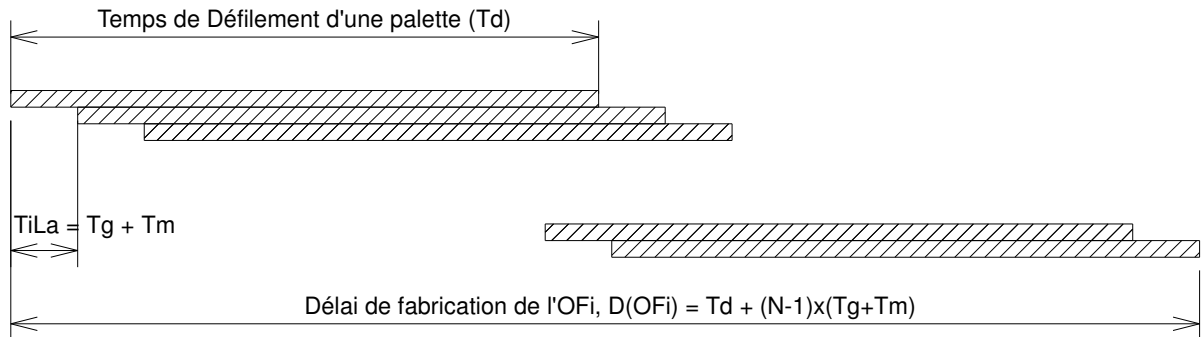


Figure 31 : Principe du calcul du délai d'un O.F

Transposé au cas multiproduit, il faut donc déterminer un Temps inter-Lancement (TiLa) propre à chaque O.F de la campagne, puisque la gamme correspondant à chaque O.F ne comporte pas obligatoirement le même temps goulet (Tg).

Le tableau de la figure 32 illustre l'application de ce calcul au cas multiproduit.

	Temps goulet Tg(k)	TiLa(k) (Tm=4sec.)	Poste goulet Tcyc maxi	Délai th. OFk	Décalage OFk à OFk+ 1
OF1	4	8	Poste 1	244	80
OF2	4	8	Poste 4	205	40
OF3	5	9	Poste 4	238	72
OF4	5	9	Poste 6	301	135
OF5	3	7	Poste 6	204	

Figure 32 : TiLa et Délai théorique de chaque O.F

On peut voir dans ce tableau que la valeur du Temps inter-Lancement peut être différente d'un O.F à l'autre. Ceci fait apparaître une première différence entre le calcul du délai et l'expérimentation : dans le cas de l'expérimentation, on fixe le Temps inter-Lancement (TiLa) à la même valeur pour tous les O.F de la campagne.

Le diagramme de la figure 33 montre le résultat issu de ce calcul du délai.

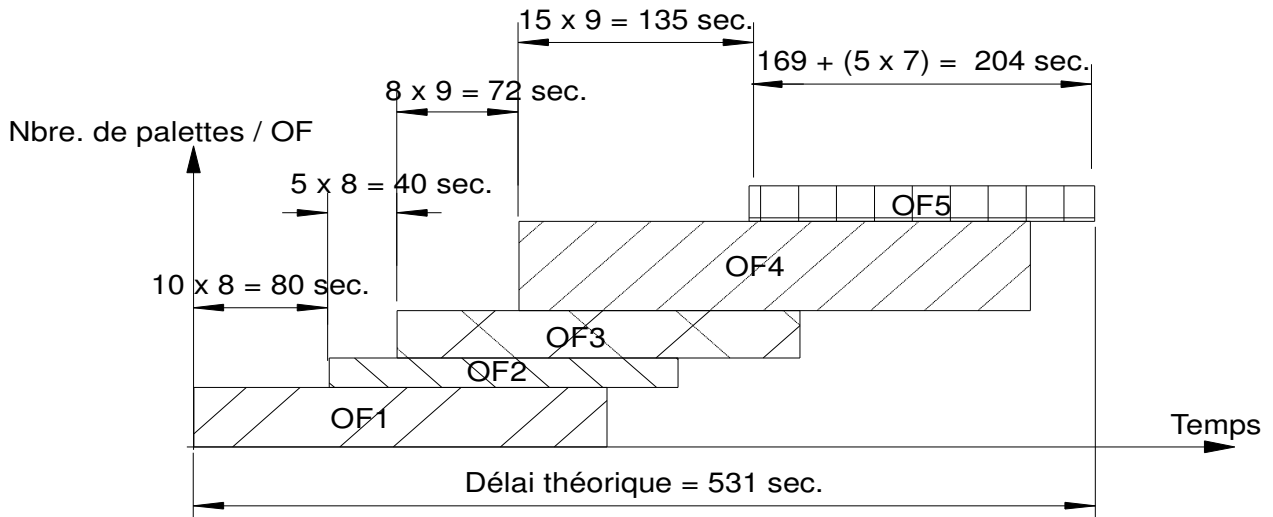


Figure 33 : Ordonnancement théorique des O.F et calcul du délai

A partir de ce cas particulier, la relation générale permettant de calculer le délai estimé (D_{th}) d'une campagne de production multiproduit peut se mettre sous la forme :

$$D_{th} = [Td(\text{Max}[k]) + (Tlot(\text{Max}[k]) - 1) \cdot (Tg(\text{Max}[k]) + Tm)] + \sum_{k=1}^{\text{Max}[k]-1} Tlot(k) \cdot (Tg(k) + Tm)$$

avec $\text{Max}[k]$ = indice de l'O.F le plus élevé.

Appliquée à l'expérimentation, on obtient donc (pour $\text{Max}[k] = 5$), en utilisant les données de la figure 27 :

$$\begin{aligned} D_{th} &= [169 + (6 - 1) \cdot (3 + 4)] \\ &\quad + [10 \cdot (4 + 4)] \\ &\quad + [5 \cdot (4 + 4)] \\ &\quad + [8 \cdot (5 + 4)] \\ &\quad + [15 \cdot (5 + 4)] \\ &= 531 \text{ secondes} \end{aligned}$$

On voit que le délai calculé par cette relation est inférieur à celui obtenu en pratique. Cela peut s'expliquer par le choix des valeurs de Temps inter-Lancement (TiLa), qui sont différentes pour chaque O.F en théorie, alors qu'en pratique, (TiLa) a la même valeur pour tous les O.F. Si l'on applique cette relation en mettant la même valeur de (TiLa) à chaque O.F, on obtient malgré tout un résultat différent de l'expérimentation.

Si l'on prend par exemple la valeur de (TiLa) qui a donné le délai minimal lors de l'expérimentation (TiLa = 9 sec., D = 579 sec.), on obtient :

$$\begin{aligned}
 D_{th} &= [169 + (6 - 1) \cdot (9)] \\
 &+ [10 \cdot (9)] \\
 &+ [5 \cdot (9)] \\
 &+ [8 \cdot (9)] \\
 &+ [15 \cdot (9)] \\
 &= 556 \text{ secondes, au lieu de } 579 \text{ secondes en pratique}
 \end{aligned}$$

A priori, le calcul du délai minimal est donc plus délicat, ou moins précis, dans un cas de production multiproduit que dans un cas de production monoproduit. Ceci peut s'expliquer par le fait que lors d'une campagne de production multiproduit, plusieurs routages d'ordres de fabrication sont réalisés simultanément, ce qui implique des chevauchements qui sont difficiles à modéliser.

La relation précédente présente le délai comme le cumul de plusieurs OF ayant chacun un temps goulet.

On peut donc essayer de déterminer une autre relation permettant de calculer le délai théorique minimal, en tenant compte plus particulièrement du poste goulet pour l'ensemble de la campagne. Le poste goulet pour l'ensemble de la campagne est défini comme le poste de travail qui a la charge la plus élevée. La charge d'un poste Ch(i) est déterminée par la relation suivante :

$$Ch(i) = \sum_{k=1}^{k=Max(k)} \sum_{i=1}^{i=Max(i)} [T_{cyc}(i,k) + T_m] \cdot T_{lot}(k)$$

avec k = indice de l'ordre de fabrication

i = indice du poste de travail

Le poste goulet est donc défini pour i correspondant à Max(Ch(i)).

Le tableau de la figure 34 présente le résultat du calcul de la charge Ch(i) pour chaque poste de travail de la campagne expérimentale.

	Tlot (k)		Tcyc(1,k)	Tcyc(2,k)	Tcyc(3,k)	Tcyc(4,k)	Tcyc(5,k)	Tcyc(6,k)
	Taille Lot		Poste 1	Poste 2	Poste 3	Poste 4	Poste 5	Poste 6
OF1	10	Gamme A	4	2			3	
OF2	5	Gamme B		1		4	3	2
OF3	8	Gamme C	3	4		5		
OF4	15	Gamme D				3	4	5
OF5	6	Gamme E	1	2				3
	44							
Somme Tcyc(i,k)			8	9	0	12	10	10
Ch(i)			166	185	0	217	225	207

Figure 34 : Calcul de la charge de chaque poste

Le poste goulet est donc le poste N° 5, qui a une charge de 225 secondes d'occupation..

On peut remarquer que le poste goulet de l'ensemble de la campagne ne correspond à aucun des postes goulets considérés pour chaque OF dans la relation précédente.

A partir du calcul des charges, on peut établir que le délai de fabrication de la campagne ne peut être inférieur à $\text{Max}(\text{Ch}(i))$. A cela, il faut rajouter le temps de défilement le plus court parmi les gammes des ordres de fabrication passant par le poste goulet. Le délai minimal (Dth) peut donc s'exprimer par la relation suivante :

$$D_{th} = \text{Max}(\text{Ch}(i)) + \min(\text{Td}(k))$$

avec $\min(\text{Td}(k))$ déterminé pour i correspondant à $\text{Max}(\text{Ch}(i))$ et pour $\text{Tcyc}(i,k) \neq 0$.

Appliquée à l'expérimentation, on obtient donc :

$$\min(\text{Td}(k)) = 172 \text{ secondes pour l'OF1}$$

$$D_{th} = 225 + 172 = 397 \text{ secondes.}$$

Cette relation donne donc un délai encore plus faible que celui obtenu par la relation précédente (397 secondes au lieu de 531 secondes).

Cela s'explique par le fait que cette nouvelle relation ne tient compte que du poste goulet, sans prendre en considération le fonctionnement des autres postes.

En conclusion, même si les relations de calcul de délai semblent être, l'une et l'autre, assez imprécises, elles ont au moins l'intérêt de servir d'objectif à une optimisation des différents paramètres de pilotage. On peut en effet tenter de se rapprocher le plus possible de ces valeurs. Cependant, contrairement à une production monoproduit où il est possible d'atteindre en pratique le délai calculé, et parfois avec plusieurs solutions, cela ne semble pas être le cas pour une production multiproduit, en tout cas pas avec les contraintes imposées aux différents paramètres de pilotage lors de ces expérimentations.

On peut conclure cette expérimentation en étudiant les courbes d'encours moyen en fonction du Temps inter-Lancement (voir figure 35) :

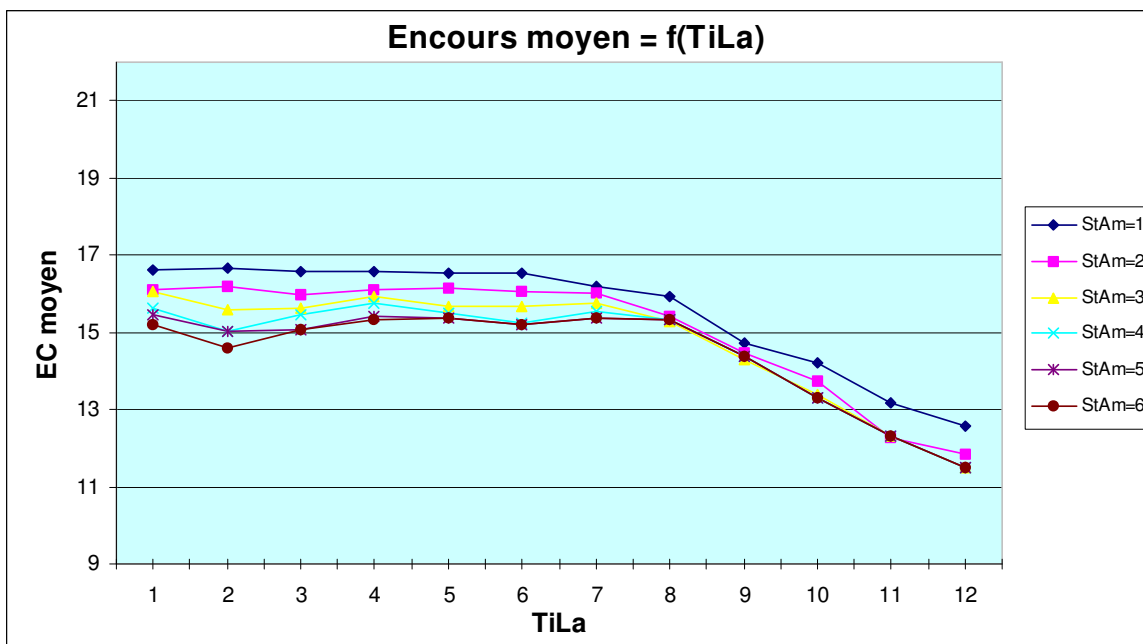


Figure 35 : Courbes EC moyen = f(TiLa)

Celles-ci ne font pas apparaître de grosses variations en fonction des deux paramètres (TiLa) et (StAm). On constate que l'encours moyen est régulier pour toutes les valeurs de (TiLa) comprises entre 1 et 8 secondes, puis chute au delà. Si l'on rapproche cette analyse à celle du délai, on remarque que la valeur minimale de celui-ci est justement obtenue pour TiLa = 9 secondes et que la zone de convergence des différentes courbes se trouvait entre 8 et 9 secondes de Temps inter-Lancement.

II.2-d) Expérimentation avec $N_p > 25$ palettes

Afin de valider les relations précédentes et de voir l'influence de l'encours sur le délai de fabrication, nous allons réaliser une nouvelle expérimentation, en changeant simplement la quantité de palettes en réserve (R_p). En réalité, celle-ci est limitée à 25 palettes, mais nous pouvons bien sûr simuler une production en utilisant une quantité supérieure.

Pour vérifier si la réserve de palettes est un facteur influent, nous réalisons donc maintenant une expérimentation avec 44 palettes en réserve, soit le nombre total de palettes à lancer pour cette campagne. Le graphique de la figure 36 présente les courbes de délai en fonction du Temps inter-Lancement avec $N_p = 44$ palettes. La comparaison de ce graphique avec celui de la figure 30 où le nombre de palettes à lancer était limité à 25 met en évidence un gain de performance en ce qui concerne le délai de fabrication de la campagne.

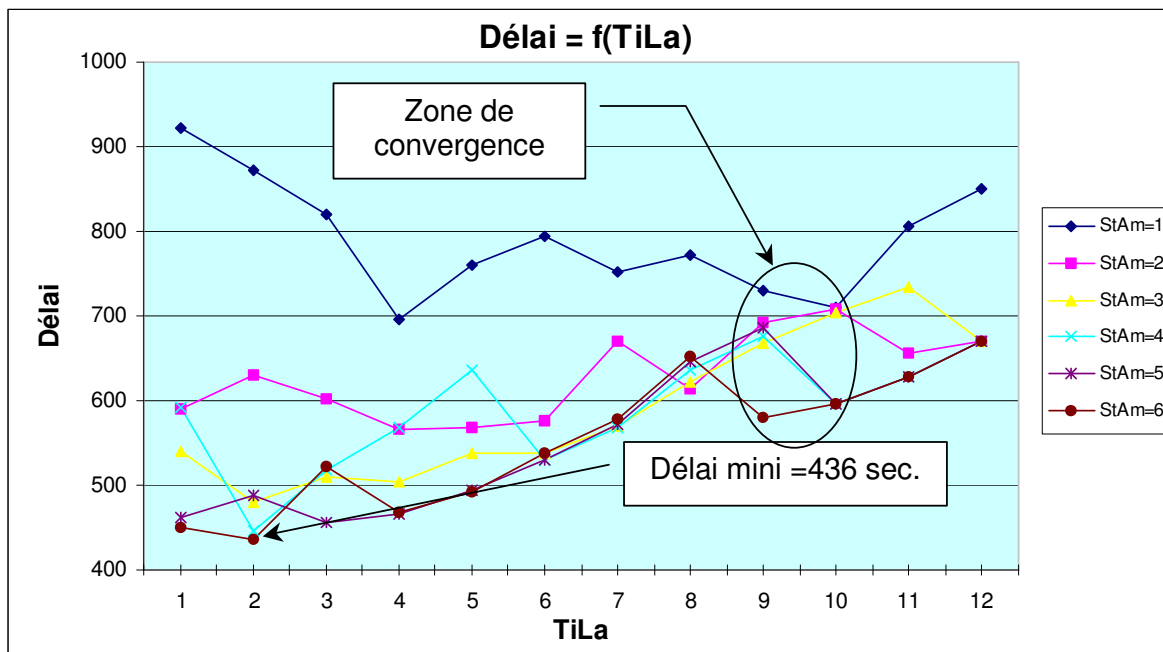


Figure 36 : Courbes $D = f(TiLa)$ pour $N_p = 44$ palettes

Le délai minimal obtenu est de 436 secondes contre 576 secondes pour $N_p = 25$ palettes. Cela signifie certainement que le nombre de palettes à utiliser doit être supérieur à 25 pour minimiser le délai. On remarque aussi une nette différence pour la valeur de (TiLa) qui minimise le délai : $TiLa = 2$ secondes pour $N_p = 44$ palettes, alors que nous avons $TiLa = 9$ secondes pour $N_p = 25$ palettes.

On peut aussi délimiter une zone de convergence où toutes les courbes se rapprochent, mais celle-ci est pratiquement située dans les mêmes limites que pour $N_p=25$ palettes :

$$9 \text{ sec.} \leq TiLa \leq 10 \text{ sec. pour } 579 \text{ sec.} \leq \text{Délai} \leq 730 \text{ sec.}$$

Par contre, l'analyse des courbes EC moyen = f(TiLa) de la figure 37 montre une variation importante de la valeur de l'encours en fonction de (TiLa) : de 11,5 palettes pour TiLa = 12 secondes à 22 palettes pour TiLa = 1 seconde. Pour TiLa = 2 secondes qui correspond au délai minimal, l'encours moyen est à la valeur maximale de 22 palettes.

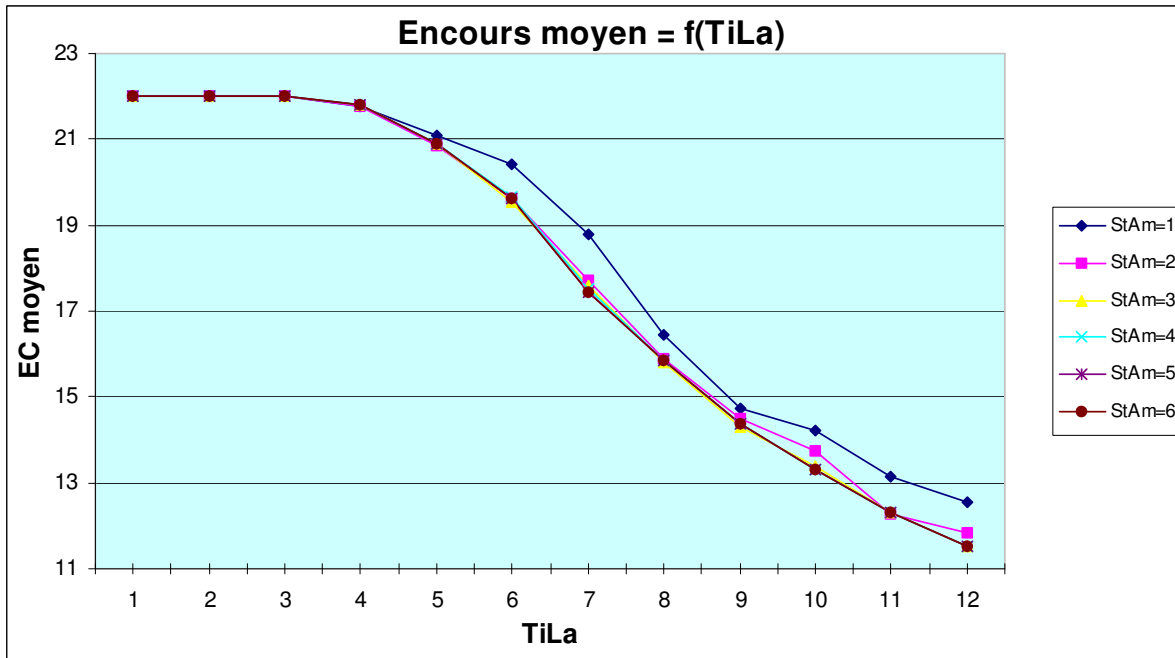


Figure 37 : Courbes EC moyen = f(TiLa)

On peut essayer de calculer la quantité optimale de palettes à utiliser en se basant sur la relation déterminée pour une production monoproduit.

Nous avons montré que le nombre optimal de palettes à utiliser (Np) permettant de minimiser à la fois le délai et l'encours moyen était :

$$Np = (Td + 37) / (Tg + Tm)$$

En appliquant cette formule à une campagne multiproduit, nous ne pouvons déterminer le nombre de palettes que relativement à chaque ordre de fabrication. On obtient :

$$OF1 : Np(1) = (172 + 37) / (4 + 4) = 27 \text{ palettes}$$

$$OF2 : Np(2) = (173 + 37) / (4 + 4) = 27 \text{ palettes}$$

$$OF3 : Np(3) = (175 + 37) / (5 + 4) = 24 \text{ palettes}$$

$$OF4 : Np(4) = (175 + 37) / (5 + 4) = 24 \text{ palettes}$$

$$OF5 : Np(5) = (169 + 37) / (3 + 4) = 30 \text{ palettes}$$

On peut donc, par exemple, lancer une expérimentation avec le nombre de palettes maximal calculé pour chaque O.F, soit 30 palettes pour l'OF 5. La figure 38 montre les courbes $D = f(TiLa)$ et la figure 39 les courbes $EC = f(TiLa)$.

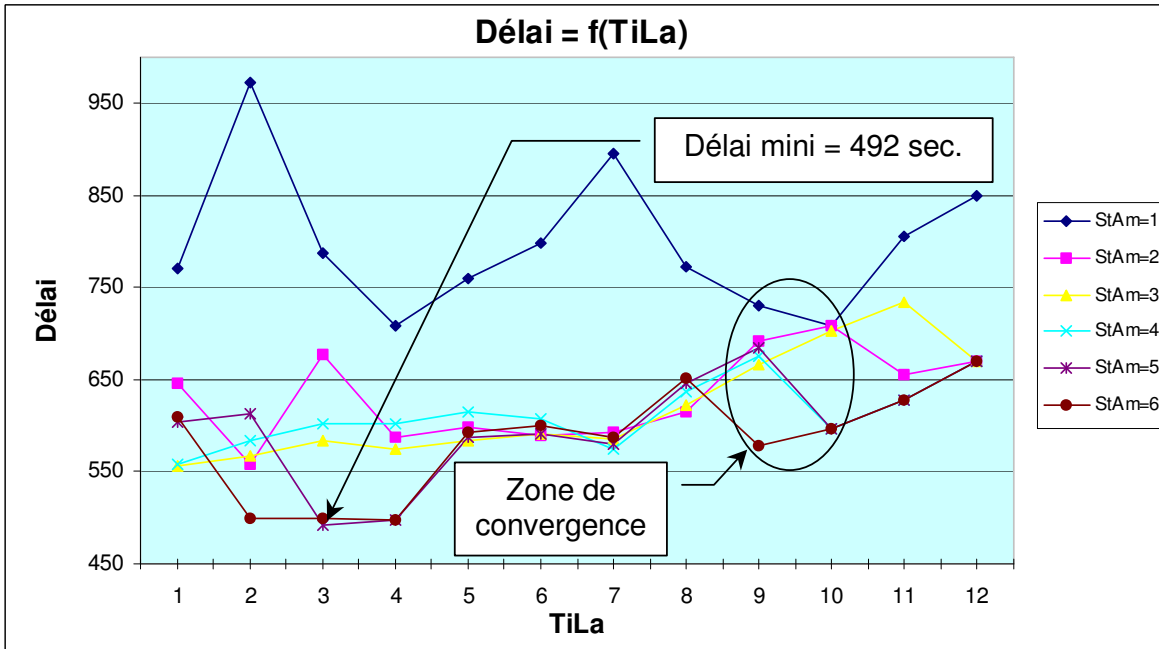


Figure 38 : Courbes $D = f(TiLa)$ pour $Np = 30$ palettes

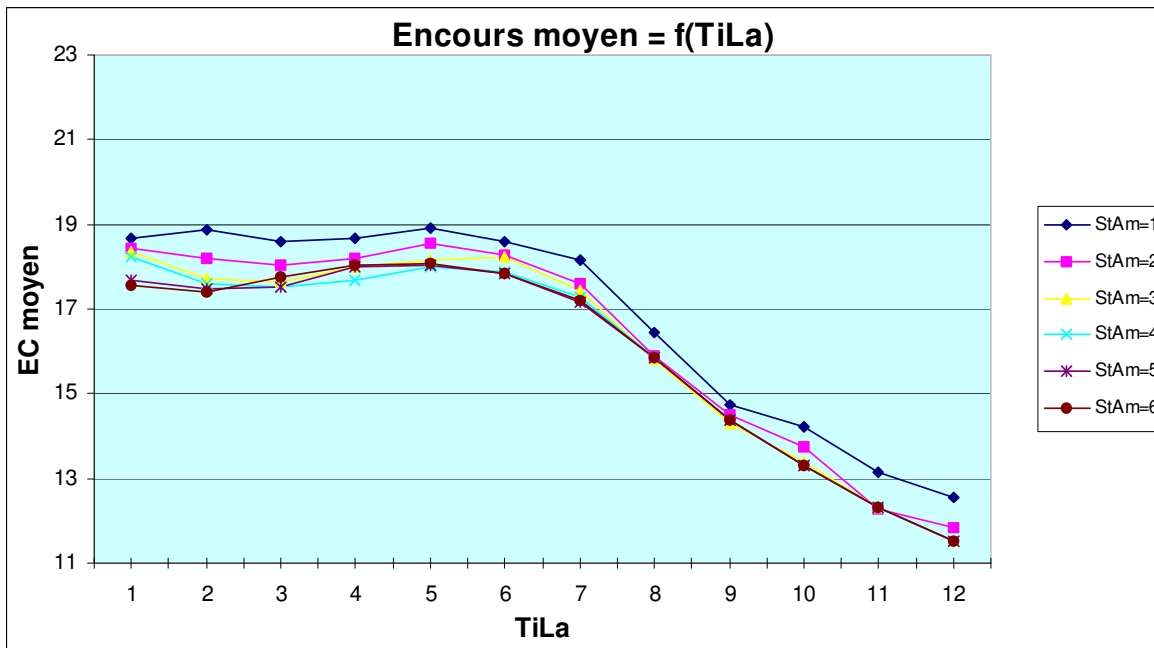


Figure 39 : Courbes $EC \text{ moyen} = f(TiLa)$

Le délai minimal obtenu avec $N_p = 30$ palettes est de 492 secondes, pour $StAm = 5$ et $TiLa = 3$ sec.

Le délai est donc supérieur à celui obtenu avec $N_p = 44$ palettes, mais l'encours moyen est plus faible (EC moyen = 18 palettes au lieu de 22 palettes).

Comme pour les 2 autres expérimentations avec $N_p = 25$ et $N_p = 44$ palettes, on retrouve sensiblement au même endroit une zone de convergence des courbes :

$$9 \text{ sec.} \leq TiLa \leq 10 \text{ sec. pour } 579 \text{ sec.} \leq \text{Délai} \leq 730 \text{ sec.}$$

A ce stade expérimental, on peut donc conclure que le cas d'une production multiproduit est, comme on s'en doutait, plus difficile à résoudre que le cas d'une production monoproduit car aucune relation théorique n'a pu être vérifiée précisément par l'expérimentation. D'autre part, les différentes courbes $D = f(TiLa)$ ne convergent pas vers un optimum unique permettant de choisir la valeur du Temps inter-Lancement ($TiLa$). Nous allons maintenant passer à des expérimentations en horizon infini pour confirmer ou pas ces premières remarques.

□ II.3- Expérimentation N°2

II.3-a) Rappel des conditions propres à cette expérimentation

- Horizon de fabrication INFINI : la ligne est toujours en charge.
- Lancement en séquence d'un nombre infini de campagnes identiques comportant chacune 5 ordres de fabrication successifs d'un total de 44 palettes. Chaque campagne est identique à celle définie pour l'expérimentation en horizon fini (voir matrice des flux de la figure 27, chap. II.2-a).

La figure 40 illustre le principe du séquençage des campagnes en horizon infini.

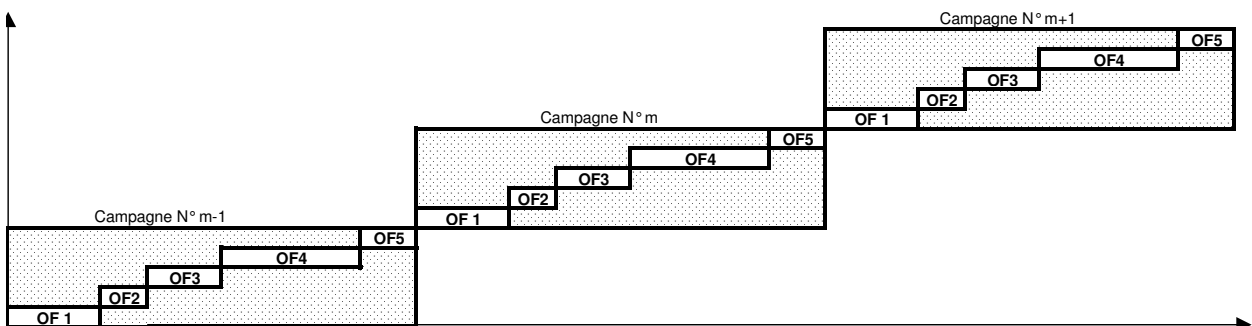


Figure 40 : Séquençage du lancement des campagnes

- Prise de mesures sur les palettes de la 5^{ème} campagne, en régime nominal. Pour la campagne N° 5, mesure du délai écoulé entre la livraison de la 1^{ère} et de la 44^{ème} palette. Une autre possibilité consiste à mesurer le délai entre le lancement de la 1^{ère} palette et la livraison de la 44^{ème} palette de la campagne. Dans le premier cas, la mesure se faisant uniquement en sortie de ligne, on l'utilise pour calculer le flux sortant (en articles / seconde) ou le temps de cycle de la campagne en régime nominal (en secondes / article), ce que les anglo-saxons appellent le "Takt Time" [MEYERS 99]. Dans le second cas, la mesure correspond au délai total de production de la campagne, entre le lancement et la livraison, et intègre directement le temps de réponse de la ligne. La figure 41 illustre ces deux possibilités de mesure. Lors de la première expérimentation, nous effectuerons simultanément les deux mesures afin de comparer les résultats.

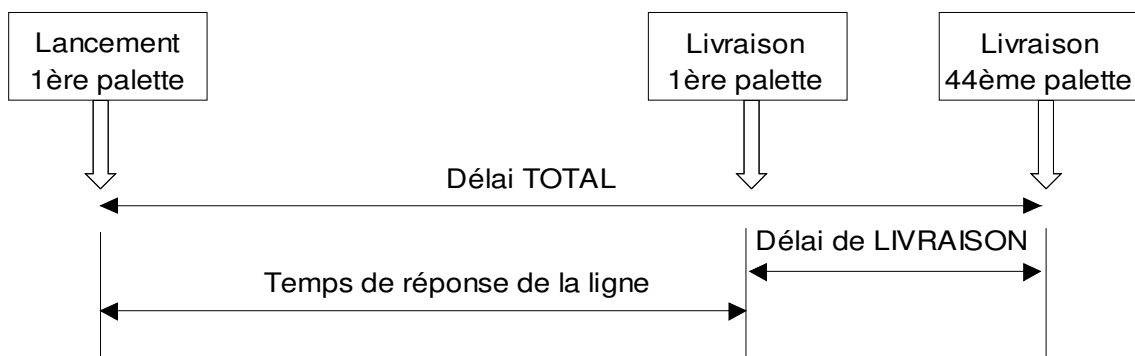


Figure 41 : Possibilité de prise de mesures

Notation :

- Délai entre la livraison de la 1^{ère} et la livraison de la 44^{ème} palette : D_{Liv}
- Délai entre le lancement de la 1^{ère} et la livraison de la 44^{ème} palette : D_{Lan}

II.3-b) Résultats

La première expérimentation est lancée avec $N_p = R_p = 25$ palettes.

La figure 42 présente les différentes courbes obtenues pour chaque valeur de (StAm). Sur chaque courbe, on a le tracé de D_{Liv} (tracé inférieur) et le tracé de D_{Lan} (tracé supérieur). On peut constater que les deux tracés sont sensiblement parallèles dans tous les cas.

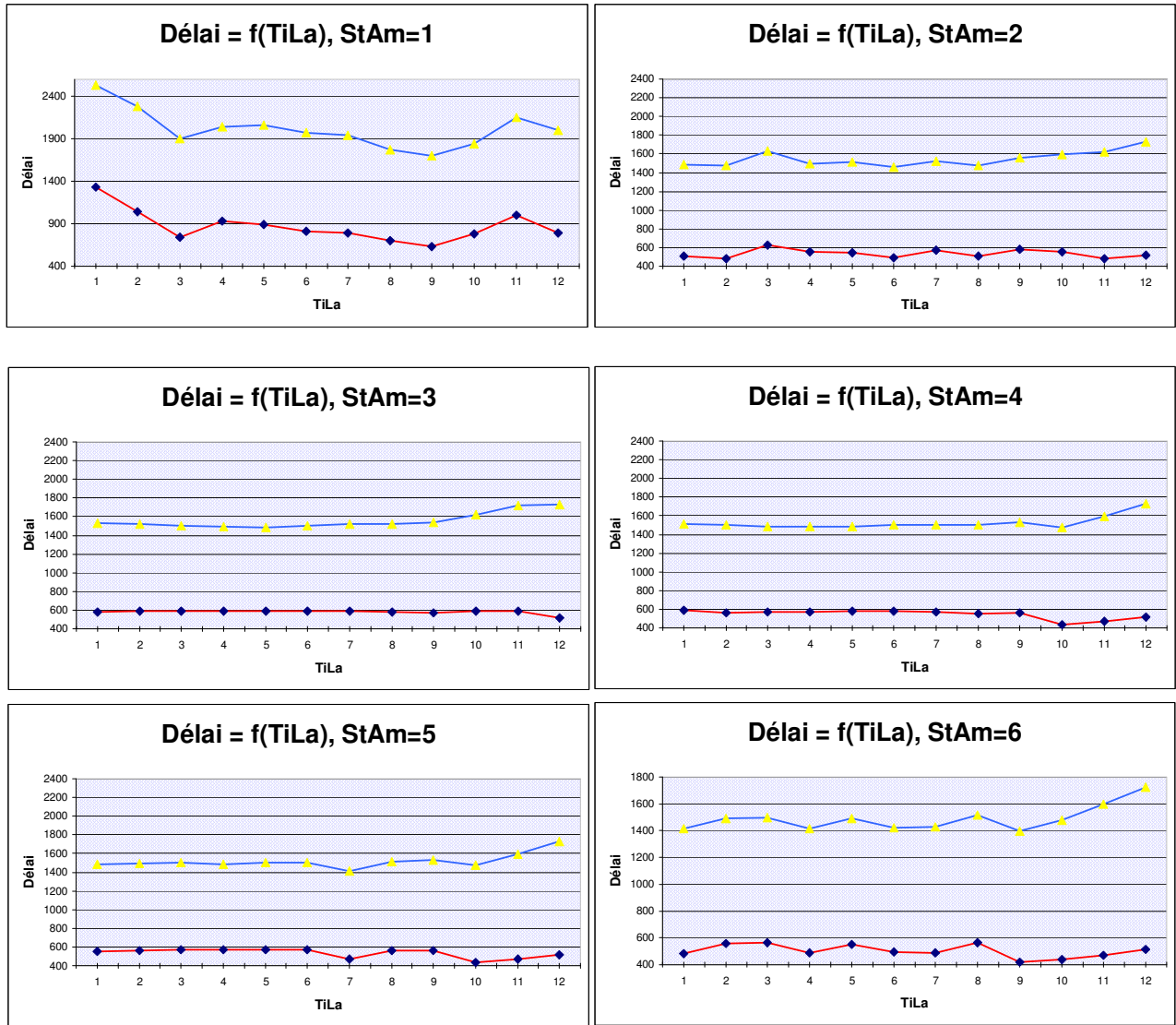


Figure 42 : Courbes D_{Lan} et $D_{Liv} = f(TiLa)$

	StAm = 1	StAm = 2	StAm = 3	StAm = 4	StAm = 5	StAm = 6
Ecart moyen	1148	1011	978	980	985	987
Ecart mini	1063	944	900	907	921	928
Ecart Maxi	1244	1214	1214	1214	1214	1214

Le tableau ci-dessus met en évidence les écarts statistiques entre D_{Lan} et D_{Liv} pour chaque valeur de (StAm). On remarque que cet écart varie très peu entre chaque courbe.

Ces différents constats nous amènent à dire que les deux mesures semblent être très dépendantes l'une de l'autre en proportionnalité.
Pour la suite de ce travail, on choisit arbitrairement de ne prendre que la mesure de D_{Liv} .

Le graphique de la figure 43 est une superposition des courbes $D_{Liv} = f(TiLa)$ pour chaque valeur de (StAm).

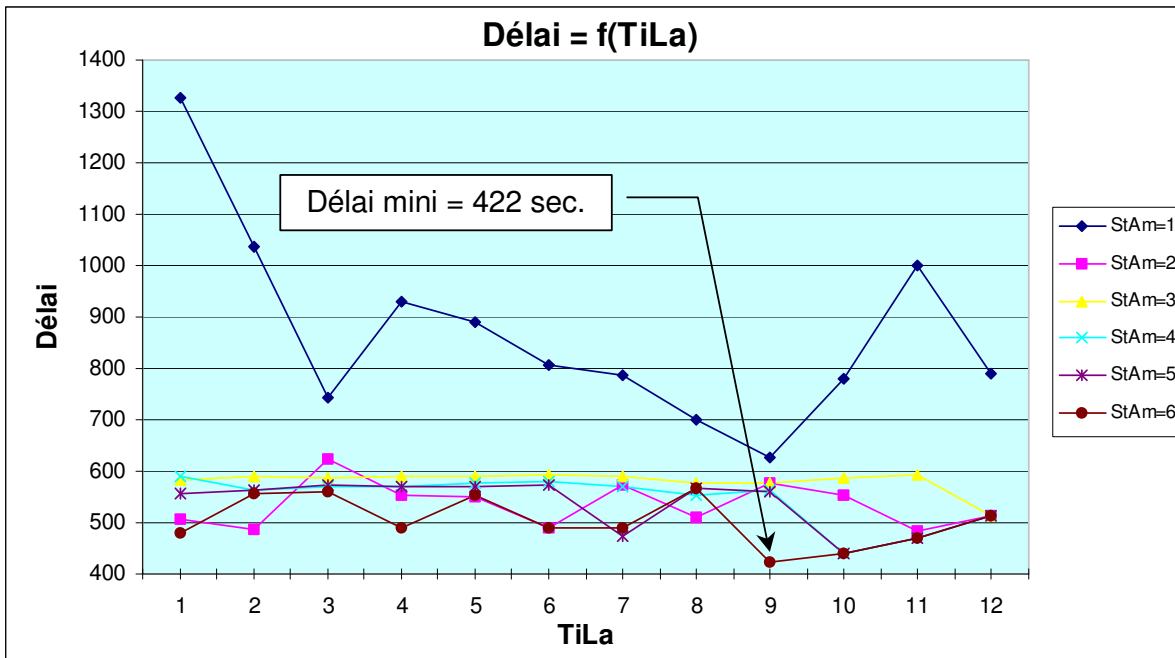


Figure 43 : Graphique de synthèse des résultats

II.3-c) Analyse des données récoltées

L'analyse du graphique de la figure 43 montre une nette différence d'amplitude de variation entre la courbe (StAm = 1) et les autres courbes. On remarque une similitude avec l'expérimentation à horizon fini en ce qui concerne la focalisation du réseau de courbes vers une valeur de (TiLa) particulière (ici, $TiLa = 9$ sec.).

La courbe (StAm = 1) étant très différente des autres au niveau de son amplitude, l'échelle choisie sur la figure 43 ne fait pas suffisamment ressortir les écarts entre les autres courbes. La figure 44 reprend les mêmes courbes mais avec une échelle de délai plus petite (de 400 à 700 secondes). La courbe (StAm = 1) n'apparaît que pour sa valeur mini (628 secondes).

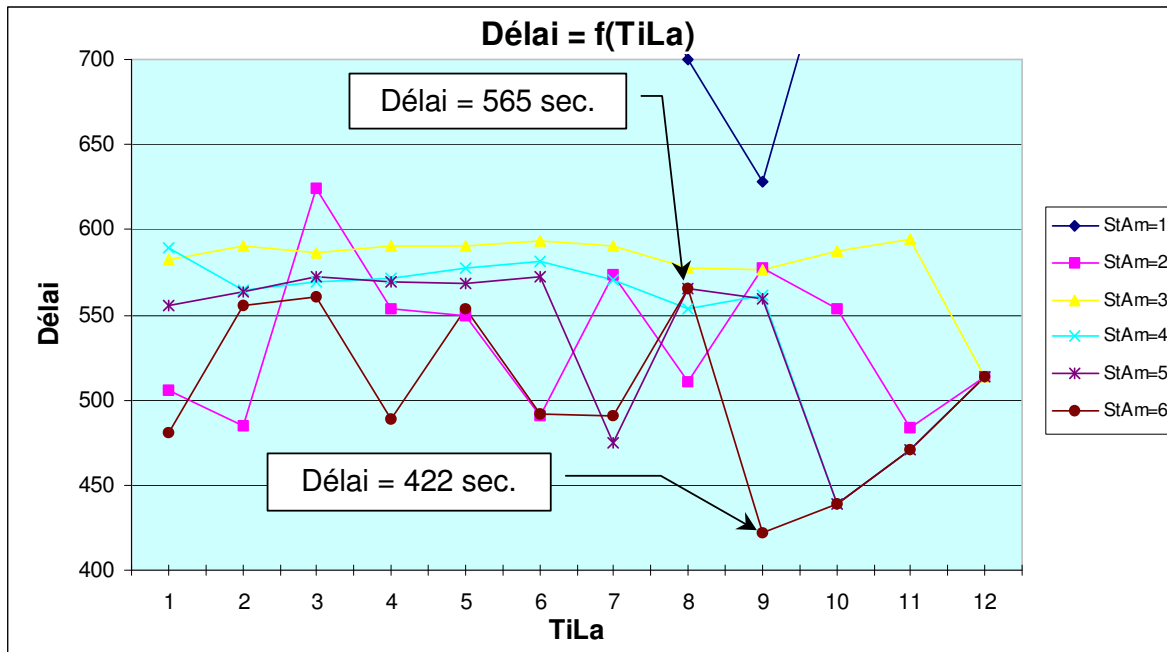


Figure 44 : Courbes $D_{Liv} = f(TiLa)$ pour $400 \text{ sec.} \leq D_{Liv} \leq 700 \text{ sec.}$

On peut voir sur ce graphique que toutes les courbes, à l'exception de (StAm = 1), convergent sur la même valeur de délai pour TiLa = 12 secondes.

On constate aussi un risque d'instabilité assez important si l'on opte pour la solution (StAm = 6 et TiLa = 9 sec.) qui donne le délai minimal (422 sec.). En effet, si (TiLa) est réduit à 8 secondes au lieu de 9 secondes, le délai passe à 565 secondes, soit une augmentation de 34 %. Dans cette situation, si l'on risque d'avoir une variation de la variable de pilotage (TiLa), la solution qui donne le délai minimal n'est pas nécessairement la meilleure. Il est alors plus judicieux de choisir une solution qui donne un délai plus élevé, mais qui garantisse une bonne stabilité de ce délai. On peut par exemple prendre, pour la même valeur de (StAm), TiLa = 10 secondes (D = 438 sec.). En cas de variation de +/- 1 de (TiLa), le délai ne varie pas de plus de 11 % entre les deux solutions extrêmes (de 422 sec. à 470 sec.).

La variation de l'encours moyen apparaît sur la figure 45.

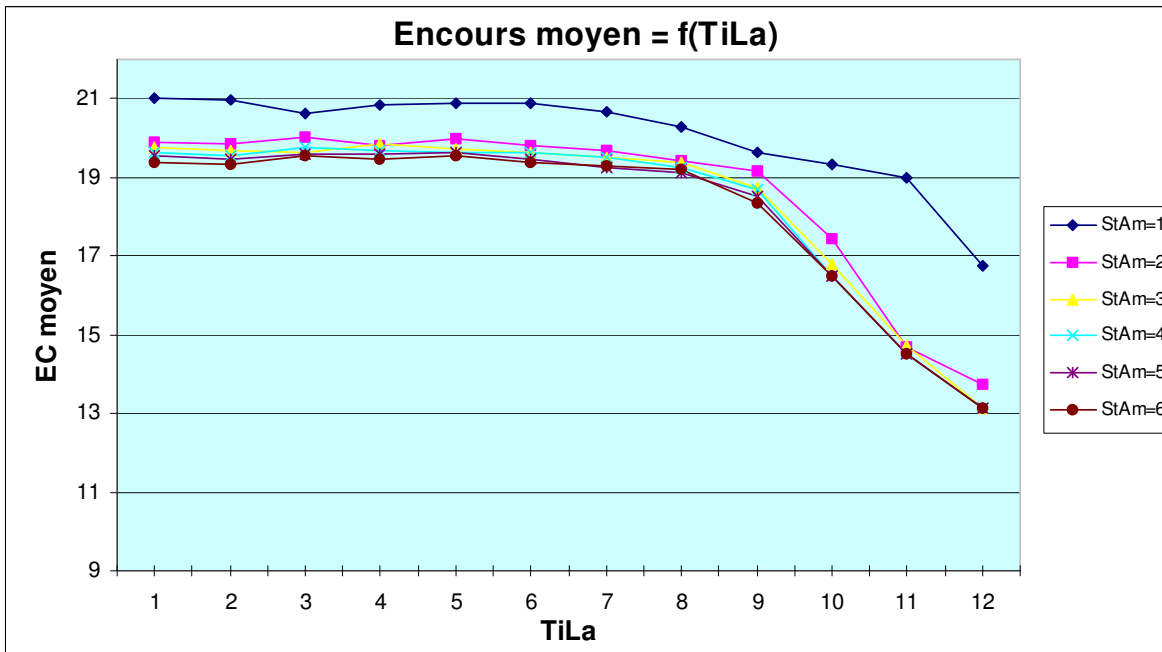


Figure 45 : $EC\ moyen = f(\text{TiLa})$ pour $Np = 25$ palettes

Si l'on retient la solution ($\text{TiLa} = 10$ sec., $\text{StAm} = 6$), l'encours moyen est de 16,48 palettes (contre 18,32 pour $\text{TiLa} = 9$ sec., ce qui confirme l'intérêt de choisir la solution "stable").

II.3-d) Expérimentation avec $N_p = 30$ palettes

Comme pour l'expérimentation à horizon fini, on lance une expérimentation avec le nombre de palettes déterminé par la relation du chapitre I.2-d et reprise dans le II.2-c. On avait alors trouvé 30 palettes au maximum. La figure 46 présente les résultats pour $N_p = 30$ palettes.

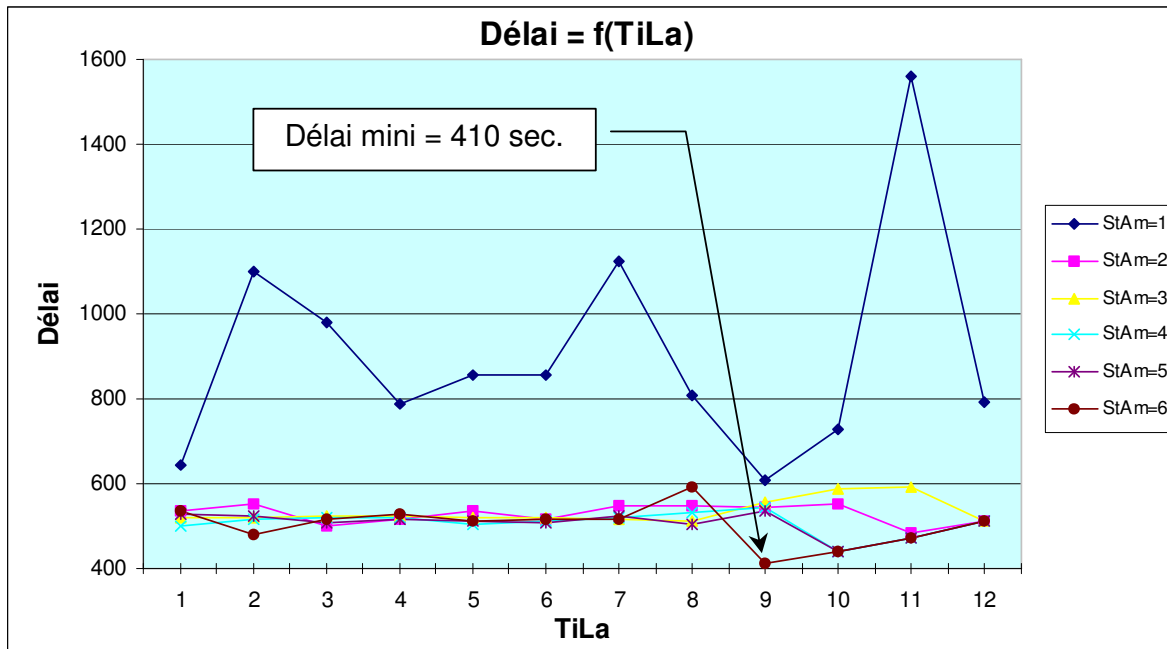


Figure 46 : Courbes $D = f(TiLa)$ pour $N_p = 30$ palettes

C'est essentiellement la courbe pour ($StAm = 1$) qui présente quelques différences de forme par rapport à l'expérimentation précédente. Malgré cela, le délai minimal est toujours situé à $TiLa = 9$ secondes pour $StAm = 1$ comme pour $StAm = 6$. Comme pour l'expérimentation avec $N_p = 25$ palettes, on peut réduire l'échelle pour mieux visualiser les variations des autres courbes. La figure 47 présente les mêmes courbes avec une échelle de délai allant de 400 secondes à 600 secondes. On peut établir exactement la même règle que pour $N_p = 25$ au sujet de la stabilité du délai. Dans ce cas encore, il est préférable d'opter plutôt pour la solution ($StAm = 6, TiLa = 10$ sec.) qui garantit un minimum de variation du délai, même en cas de variation de ± 1 seconde du paramètre ($TiLa$).

L'encours moyen obtenu pour la solution ($StAm = 6, TiLa = 10$) est pratiquement le même (EC moyen = 16, 47 palettes) que pour $N_p = 25$ palettes (voir figure 48).

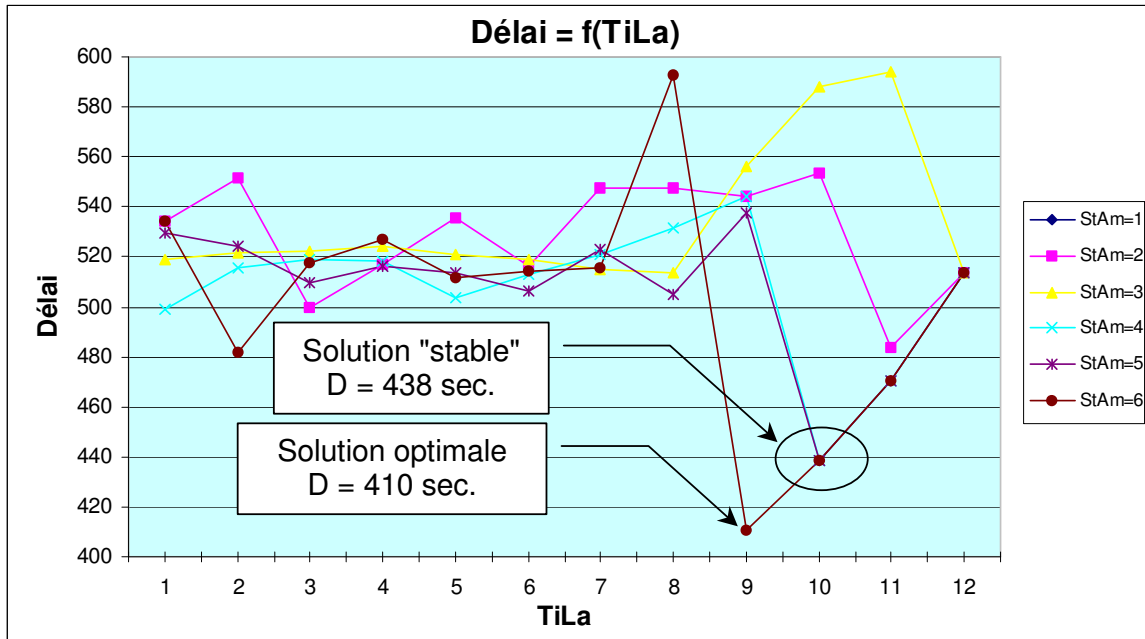


Figure 47 : Courbes $D_{Liv} = f(TiLa)$ pour $400 \text{ sec} \leq D_{Liv} \leq 600 \text{ sec}$

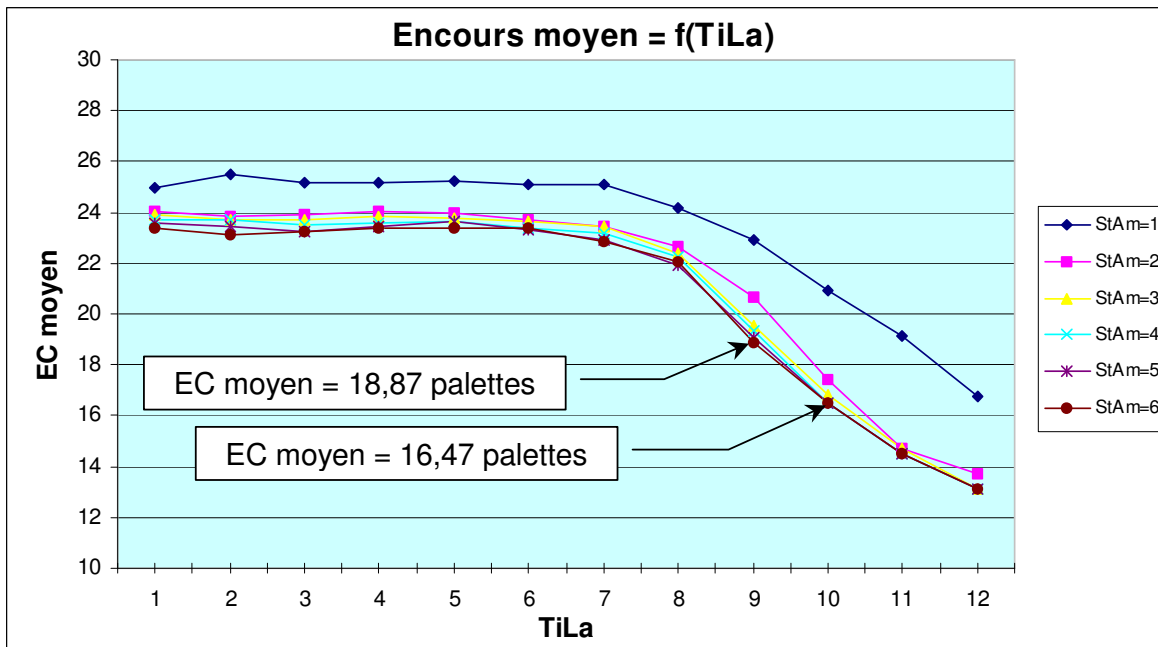


Figure 48 : $EC \text{ moyen} = f(TiLa)$ pour $Np = 30 \text{ palettes}$

Cette expérimentation n'apporte pas de gain substantiel au niveau du délai et l'encours moyen reste à peu près identique. Les solutions optimales et les solutions "stables" sont les mêmes pour $N_p = 25$ et pour $N_p = 30$ palettes.

Conclusion

Pour une production de type monoproduit, les relations que nous avons établies permettent de calculer le délai minimal et nous avons vérifié expérimentalement que l'on pouvait en pratique obtenir ce délai, avec même plusieurs solutions de choix des paramètres de pilotage. On a fixé des critères complémentaires, à la fois de stabilité de la solution retenue et de minimisation de l'encours moyen, afin de choisir parmi toutes les solutions disponibles, la mieux adaptée.

Pour une production de type multiproduit, nous avons pu établir deux relations permettant de calculer le délai minimal de production de la campagne. Les expérimentations réalisées n'ont pas permis d'atteindre le délai calculé. Ces relations, et en particulier la deuxième, permettent toutefois de fixer un objectif à atteindre en sachant que celui-ci constitue un minimum en dessous duquel il est impossible de descendre. En horizon fini, les expérimentations ont permis de mettre en évidence l'importance du choix du nombre de palettes à utiliser (N_p). Avec un nombre de palettes trop faible, on ne peut pas se rapprocher du délai minimal, alors qu'en augmentant (N_p), on arrive à réduire de façon conséquente le délai, mais en contrepartie, l'encours moyen (EC moyen) augmente. L'utilisation empirique de la relation permettant de déterminer le nombre de palettes à utiliser pour une production monoproduit fournit un compromis au niveau du délai tout en maintenant l'encours à un niveau acceptable. En horizon infini, nous avons pu constater sur les deux expérimentations réalisées, en changeant uniquement le nombre de palettes à utiliser ($N_p = 25$ puis $N_p = 30$ palettes), que dans un cas comme dans l'autre, il semble être judicieux d'opter pour une solution "stable" plutôt que pour la solution optimale.

A partir de ces expérimentations, et notamment celles correspondant au cas multiproduit, on a pu mettre en évidence les limites d'une méthodologie d'optimisation basée sur l'utilisation de relations théoriques ou la réalisation d'expérimentations, en limitant fortement le nombre de combinaisons. Par exemple, pour le cas multiproduit, le délai théorique obtenu par calcul n'a pas été atteint en pratique, et aucune relation n'a été établie permettant de trouver une solution théorique donnant un délai minimal.

On peut donc rechercher une méthodologie, permettant de répondre à la question : "Existe t-il une combinaison des paramètres de pilotage qui tende vers le délai minimum théorique tout en minimisant l'encours moyen ?"

La méthodologie employée jusque là ne permet pas de proposer de solutions théoriques. Il faut donc se tourner vers des méthodes d'optimisation algorithmiques qui vont être présentées dans la partie 7.

Bibliographie de la partie 6 : "Analyse des résultats des simulations"

[SHINGO 86] : "Maîtrise de la production et méthode Kanban"

Editions d'Organisation, 1986

S. Shingo

[SPEARMAN 90] : "ConWip : A pull alternative to Kanban"

International Journal of Production Research, Vol. 28-5, p. 879-894

M.L. Spearman

[MEYERS 1999] : "Motion and Time Study for Lean Production"

Editions Prentice Hall, 1999

F. E. Meyers

[FONTANILI 98] : "A ring-shaped mechanical assembly line optimized by a genetic algorithm"

Congrès IDMME'98, Compiègne

F. Fontanili, A. Vincent, T. Soriano, R. Ponsonnet

- PARTIE 7 -
Optimisation de la production

Sommaire

Introduction	227
I- Modélisation pour l'optimisation.....	229
□ I.1- Optimisation sur le système lui-même.....	229
□ I.2- Optimisation sur un modèle du système	230
I.2-a) Formalisation (ou formulation) mathématique ou analytique	231
I.2-b) Utilisation d'un modèle de simulation de flux	231
II- Mise en évidence de la complexité du problème	233
□ II.1- Objectif de l'optimisation	233
□ II.2- Calcul de la combinatoire des solutions possibles.....	234
□ II.3- Démarche d'optimisation suivie	236
III- Typologie des méthodes d'optimisation.....	237
□ III.1- Méthodes exactes ou de Recherche Opérationnelle	237
□ III.2 - Méthodes approchées ou heuristiques.....	239
III.2-a) Formulation du problème d'optimisation (appliquée à la ligne expérimentale).....	239
III.2-b) Principe du voisinage d'une solution.....	239
III.2-c) Méthodes de recherche locale définies par un voisinage.....	241
III.2-d) Méthodes de recherche globale définies par un voisinage.....	243
III.2-e) Méthode d'exploration aléatoire	247
III.2.f) Algorithmes génétiques	248
Conclusion	258
Bibliographie	259

Introduction

Dans la partie 6, nous avons pu mettre en évidence, au travers des expérimentations réalisées, les difficultés rencontrées pour optimiser les paramètres de pilotage d'une ligne de production. Ces difficultés sont au moins de deux natures :

- Trouver une solution permettant de satisfaire des critères de production pris séparément : minimiser le délai de production, réduire les encours, minimiser le nombre de palettes à utiliser, etc...
- Trouver une solution "stable", c'est à dire qui tolère des petites variations des paramètres de pilotage sans pour autant trop dégrader les critères d'optimisation.

Dans cette partie, l'objectif principal est de recenser les méthodes d'optimisation algorithmiques susceptibles d'améliorer la recherche d'une solution optimale. En effet, comme on l'a vu lors des expérimentations pour des productions de type multiproduit, il semble être difficile de trouver par une relation quelconque une solution théorique qui satisfasse les critères. Nous allons donc plutôt nous diriger vers une approche algorithmique couplée à la simulation, afin d'optimiser le pilotage de la ligne de production.

Dans le chapitre 1, on rappelle l'intérêt d'utiliser le modèle de simulation pour optimiser un système de production.

Alors que dans la partie 6 nous avons limité les paramètres et surtout le nombre de valeurs qu'ils peuvent prendre, dans cette partie beaucoup de contraintes seront levées. Après avoir défini l'objectif d'une optimisation, le chapitre 2 met donc en évidence la complexité du problème dès lors que l'on autorise chaque paramètre à prendre toutes les valeurs de son étendue.

Le chapitre 3 est un état de l'art des principales méthodes d'optimisation utilisables en production. On trouve :

- Les méthodes à base de règles établies par expérience ou par déduction logique. C'est ce qui a été fait dans la partie 6, avec les difficultés que l'on a montrées.
- Les méthodes exactes, dont l'objectif est de déterminer l'optimum exact. Ces méthodes sont bien adaptées à des cas où il est facile de modéliser le

problème sous la forme d'une ou plusieurs équations, ce qui n'est pas le cas de notre ligne de production. Nous ne nous étendrons donc pas sur ces méthodes.

- Les méthodes approchées ou heuristiques, qui au contraire des précédentes, sont bien adaptées à un couplage avec un modèle de simulation de flux et dans les cas où le problème est complexe avec une très forte combinatoire. C'est donc essentiellement ces méthodes qui seront détaillées dans le chapitre 3, avec une insistance particulière pour les algorithmes génétiques, que nous utiliserons dans la partie 8 pour optimiser la ligne de production. Comme leur nom l'indique, ces méthodes ne garantissent pas l'obtention de l'optimum global, mais uniquement de trouver UN optimum intéressant parmi d'autres.

I- Modélisation pour l'optimisation

En suivant une approche systémique [LEMOIGNE 93], un système peut-être modélisé par une boîte noire sur laquelle agissent des variables d'entrées. Ce que l'on veut observer du système est modélisé par des variables de sortie (figure 1).

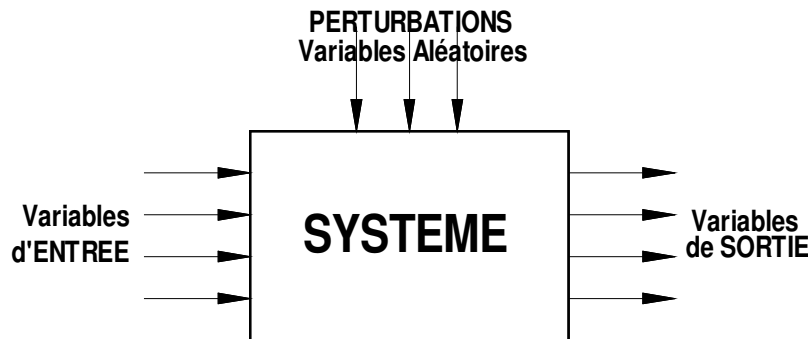


Figure 1 : Modèle systémique

Pour optimiser un système, deux solutions sont possibles , travailler sur le système lui-même, ou utiliser un modèle du système.

□ I.1- Optimisation sur le système lui-même

Dans ce cas, on ne travaille pas sur un modèle, mais sur le système lui-même, ceci à condition que l'on soit capable d'agir facilement sur les variables d'entrée et de mesurer les variables de sortie. Cette solution consiste la plupart du temps à tester plusieurs combinaisons de variables et de valeurs de variables d'entrée et d'en déduire une combinaison optimale. Dans ce cas, plusieurs inconvénients peuvent apparaître :

On immobilise le système pour effectuer l'optimisation, ce qui n'est pas sans causer des problèmes budgétaires si l'optimisation doit durer pendant un temps assez long par rapport au temps de cycle du système.

Si le temps de réponse (*) du système est important, il est indispensable de réduire le nombre de combinaisons à tester. Des méthodes ont été développées au début du siècle (FISHER 1925) et plus récemment (TAGUSHI 1960) afin d'optimiser des systèmes réels : Plans d'expériences et Méthode Tagushi [PILLET 92] [DEMONSANT 96]. Ces méthodes

(*) Le Temps de Réponse du système correspond à la durée qui s'écoule entre la validation des variables d'entrée et la mesure de leur effet sur les variables de sortie.

ont pour but principal d'organiser le déroulement des expériences et de réussir à obtenir une combinaison optimale des variables d'entrée en un minimum d'essais (plans d'expériences fractionnaires). Le principal attrait de la méthode Tagushi est de proposer des expériences dans lesquelles plusieurs variables d'entrée peuvent changer de valeur simultanément. En effet, dans un plan d'expérience complet, la méthode consiste à bloquer toutes les variables d'entrée sauf une, et de mesurer l'évolution des variables de sortie. Le principal inconvénient est donc de ne pas tenir compte des éventuelles interactions qu'il peut y avoir entre certaines variables.

Avec la méthode Tagushi, non seulement le nombre d'essais est réduit, mais on tient aussi compte des éventuelles interactions entre plusieurs variables. Malgré tout, cette méthode se limite à une dizaine de variables d'entrées (appelées Facteurs) ne prenant pas plus de deux ou trois valeurs différentes (appelées Niveaux) avec une évolution linéaire entre chaque niveau. Cette méthode, encore peu connue en France il y a cinq ans, est maintenant largement utilisée pour optimiser le réglage des procédés de fabrication sur des machines automatisées : réglage de presses d'injection pour pièces en matière plastique, réglage de robots de peinture, réglage de machines à mouler les carreaux de plâtre, etc....

Pour conclure sur l'optimisation réalisée sur le système réel, on peut dire qu'elle est utilisée quand on ne dispose pas d'un modèle du système réel, soit parce que la modélisation est trop complexe, soit parce qu'elle est trop imprécise.

□ 1.2- Optimisation sur un modèle du système

Dans le cas où le système réel n'existe pas encore car il est en phase de conception, on travaille sur un modèle du système à optimiser. Même dans des cas où le système réel existe, cela peut s'avérer indispensable si l'on désire éviter les inconvénients de la méthode précédente. Le problème qui se pose est donc de réussir à modéliser le système à étudier de telle sorte que le modèle soit le plus fidèle possible du comportement que l'on veut reproduire tout en garantissant des temps de réponse bien plus courts que le système réel. Ce modèle peut être formalisé de deux façons :

1.2-a) Formalisation (ou formulation) mathématique ou analytique

Dans le cas général, il s'agit "simplement" de mettre le problème sous la forme d'une ou plusieurs équations. Malheureusement, cette formalisation n'est pas vraiment bien adaptée à la modélisation de systèmes réels complexes. En effet, si l'objectif du modèle

mathématique est d'optimiser le système qu'il représente, cela dépend essentiellement de l'existence de dérivées.

Dans le premier chapitre de son ouvrage, [GOLDBERG 94] lance un pavé dans la mare des théoriciens de l'optimisation. En effet, il écrit à leur sujet:

"Ils ont trop facilement accepté l'héritage des grands mathématiciens du 18^{ème} et 19^{ème} siècle, qui décrivent un monde propre, fait de fonctions quadratiques, de contraintes idéales, et de dérivabilité omniprésente. Le monde réel à explorer est envahi de discontinuités, d'espaces de recherche bruités et multimodaux, ce qui le rend bien moins adapté au calcul. Il n'est donc pas surprenant que les méthodes soumises aux contraintes de continuité et de dérivabilité ne soient adaptées qu'à une classe de problèmes très limitée. ...".

Dans le cas des systèmes de production, certaines méthodes de modélisation mathématique ont pourtant été développées et appliquées avec succès. Les chaînes de Markov, les réseaux de files d'attente et la programmation dynamique sont aujourd'hui les modèles mathématiques les plus utilisés en production. Par exemple, [SURI 97] reconnaît dans son article que pour modéliser des systèmes de manutention automatisée, les réseaux de files d'attente manquent de précision, mais note que ceux-ci offrent un bon compromis entre exactitude et vitesse de calcul tout en prenant en compte les aspects stochastiques et dynamiques du système. L'utilisation de ces outils mathématiques pour la modélisation des systèmes de production est décrite en détail et avec de nombreux exemples dans les ouvrages de [GERSHWIN 94] et [BUZACOTT 93].

1.2-b) Utilisation d'un modèle de simulation de flux

Dans la partie 4, nous avons déjà utilisé la simulation de flux afin de disposer d'un modèle de la ligne de production pour réaliser des expérimentations.

Dans l'industrie, la simulation de flux connaît un vif intérêt depuis que cet outil informatique est à la portée des non informaticiens grâce à l'utilisation d'interfaces graphiques conviviales et de techniques orientées objet. Outre cet aspect, un modèle de simulation de flux ne requiert pas la programmation d'équations mathématiques complexes et peut représenter très finement le système à étudier. En effet, les progiciels proposent aujourd'hui une visualisation dynamique des flux, ce qui permet de voir ce qui se passe dans l'atelier de façon très réaliste.

Pour justifier le choix de la simulation de flux par rapport à l'utilisation d'un modèle analytique, [ALFARO 97] explique que le modèle analytique est trop réducteur et ne permet pas d'obtenir une représentation réaliste du système par rapport à un modèle de simulation de flux. La simulation de flux permet au concepteur du modèle de définir le niveau de précision qu'il désire avoir par rapport au problème qui le préoccupe.

[QUERE 97] a choisi de modéliser un système de production par la simulation de flux car elle permet d'intégrer la complexité du système dans son ensemble. La simulation de flux répond donc particulièrement bien au manque de formalisme des systèmes de production. Elle permet une théorisation à base d'observations, d'essais et d'erreurs (sans conséquences) et présente pour le chercheur une grande valeur heuristique [UNIVERSALIS 95]. C'est de cette façon que nous l'avons utilisée dans la partie 6.

Pour conclure sur ces deux possibilités de modélisation du système réel, [GERSHWIN 94] semble être le plus pondéré des deux écoles : il pense que le mieux est d'utiliser ces deux outils : modélisation analytique pour dégrossir la conception du système et modélisation pour la simulation de flux pour vérifier et affiner les résultats. Il admet quand même qu'il y a un gouffre entre la théorie et la pratique : "... *Many theoreticians and some practitioners blame this on the ignorance and lack of sophistication of the practitioners, but, in my opinion, this is unfair. Some of the reasons for the gap is due to the absence of science in the theoretician's work.*"

On peut donc dire que lorsque le système n'est pas trop complexe, sans pouvoir facilement quantifier ce niveau de complexité, et que l'on cherche uniquement à dégrossir son optimisation, un modèle mathématique semble être l'outil adapté. Par contre, si le système est complexe ou si l'on désire un degré de précision très élevé par rapport au système réel, le modèle de simulation de flux est l'outil idéal.

En ce qui concerne ce travail de recherche, on dispose d'un système expérimental réel (voir Partie 3), sur lequel on peut donc effectuer des expériences, mais le temps de réponse peut s'avérer trop important pour pouvoir tenter une optimisation directement sur ce système. L'optimisation sera donc réalisée à partir d'un modèle. Pour les raisons évoquées plus haut, une modélisation mathématique de ce système, bien qu'elle présente un intérêt certain en terme de temps de réponse, est trop complexe et ne permet pas de vérifier visuellement l'adéquation entre le fonctionnement du système réel et son modèle, comme le permet un modèle de simulation. D'autre part, l'objectif fixé dans ce travail est d'optimiser le système réel et de ne pas se contenter d'optimiser son modèle. Il faut donc s'assurer de la **robustesse** de l'optimisation, c'est à dire son aptitude à améliorer des variables de sortie, même en présence de perturbations aléatoires. Ce sont justement ces perturbations aléatoires qu'il n'est pas aisé de modéliser analytiquement sur un système de production.

On choisit donc d'utiliser la simulation de flux pour optimiser le système expérimental. La démarche sera la même que dans la partie 6, sauf qu'ici le modèle de simulation sera directement couplé à un algorithme d'optimisation.

La conduite de la modélisation et de la simulation de flux est présentée en détail dans la partie 4 , à la fois de façon générale, mais aussi en particulier pour le support

expérimental que constitue la ligne flexible. Nous considérons donc à ce stade que nous disposons d'un modèle de simulation de la ligne flexible.

II- Mise en évidence de la complexité du problème

□ II.1- Objectif de l'optimisation

L'objectif de l'optimisation est de déterminer les extremums d'une fonction, appelée dans la littérature **FONCTION OBJECTIF** ou **FONCTION d'ADAPTATION** ou **FONCTION d'ADEQUATION**.

Le terme de "**Fonction Objectif**" sera retenu pour la suite de ce travail. Cette fonction correspond à une relation algébrique entre une ou plusieurs variables de sortie du système étudié que l'on appelle "critères" (voir figure 1).

La littérature dans le domaine de l'optimisation ne s'étend pas beaucoup sur la définition de cette fonction objectif. On trouve surtout des contributions et des explications sur différentes méthodes d'optimisation, mais peu d'auteurs se hasardent dans le choix et la pertinence de la fonction objectif. En effet, comme on l'a déjà constaté partiellement dans la partie 6 (compromis à trouver entre minimisation du délai et minimisation de l'encours) et comme cela sera présenté dans la suite de ce travail, on se heurte à la subjectivité quand on doit choisir une fonction objectif, ce qui est d'ailleurs assez paradoxal....

Pour introduire cette notion de fonction objectif, on peut considérer qu'il existe 2 classes de fonction :

- Quand une seule variable de sortie est prise en compte dans la fonction objectif, on dit qu'il s'agit d'une **fonction objectif simple critère ou monocritère**. C'est le cas le plus répandu dans la littérature et il est vrai qu'il peut satisfaire bon nombre d'applications. Par exemple, dans un atelier de production, l'optimisation peut avoir pour fonction objectif la minimisation des temps de changement d'outil. Le seul critère pris en compte est le temps de changement d'outil, mais plusieurs variables d'entrée peuvent influencer sur ce critère. Dans la partie 6, nous avons réalisé des expérimentations dans le but de minimiser, en priorité, le délai de production. Les variables d'entrée étaient le Temps inter-Lancement (TiLa) et la capacité des stocks amont (StAm).

- Quand plusieurs variables de sortie doivent être prises en compte dans la fonction objectif, on dit qu'il s'agit d'une **fonction objectif multicritère**. Si l'on reprend le même exemple que dans la classe précédente, la fonction objectif doit non seulement minimiser le temps de changement d'outil, mais en plus doit maximiser le taux d'utilisation des machines et minimiser les tailles de lot. Il y a donc 3 critères à prendre en compte simultanément : temps de réglage, taux d'utilisation des machines et tailles de lot. Dans la partie 6, après avoir trouvé les solutions qui minimisent le délai, nous avons retenu uniquement celles qui minimisent aussi l'encours moyen. Nous avons pu constater que dans certains cas, le choix d'une solution qui répond à plusieurs critères liés au système doit être complété par un critère de stabilité de la fonction objectif.

Il est clair que cette 2ème classe de fonction objectif présente un intérêt certain pour les problèmes d'optimisation de production, qui sont, comme le rappelle [MOLET 97, page 18], intrinsèquement multicritères, avec en plus l'inconvénient majeur d'être souvent contradictoires. Il s'agit donc de trouver un compromis entre plusieurs critères dont la difficulté est de classer quantitativement ces critères et de leur affecter un poids.

Dans la suite de ce travail, on a tenté de rechercher et de comparer plusieurs fonctions objectifs, de chacune des 2 classes précédentes. Les résultats de ce travail figurent dans la partie 8.

□ II.2- Calcul de la combinatoire des solutions possibles

Le calcul de la combinatoire des solutions consiste à déterminer le nombre total de combinaisons entre les différentes valeurs des variables d'entrée. Ce calcul s'appuie sur des notions d'analyse combinatoire basée sur le Principe du Dénombrement, sur les permutations et les combinaisons [GOLDBERG 94, page 340]. Cela suppose que les problèmes à traiter comportent un ensemble fini, dénombrable mais de cardinal élevé, de solutions. Toutes les méthodes d'optimisation combinatoire ont pour but de résoudre ce type de problème.

Si l'on s'appuie sur le système expérimental étudié, on peut dénombrer (voir Partie 5, chapitre IV sur typologie des variables de pilotage) au moins 4 types de variables de pilotage :

- Le Temps inter-Lancement ou la cadence de lancement (TiLa) en secondes,
- La capacité des stocks amont (StAm) en nombre de palettes,
- La capacité des stocks aval (StAv) en nombre de palettes,
- Les règles de priorité en sortie de poste (Prio).

Sachant que pour chaque type, plusieurs variables homogènes (*) peuvent être utilisées, il est possible de dénombrer les solutions possibles.

Exemple 1 : (correspond aux expérimentations de la partie 6)

On considère uniquement 2 types de variables d'entrée :

*- Le Temps inter-Lancement (**TiLa**) est constant entre chaque palette. Il correspond donc à une seule variable discrète et homogène de pas 1, dont l'étendue est de [1..20] secondes.*

*- La capacité des stocks amont (**StAm**) est la même sur tous les postes de travail. Elle correspond donc à une seule variable discrète et homogène de pas 1, dont l'étendue est [1..6] palettes.*

Le nombre de solutions possibles est de $20 \times 6 = 120$.

Exemple 2 : (correspond à ce qui sera fait dans la partie 8)

On considère 2 types de variables d'entrée :

*- Le Temps inter-Lancement (**TiLa**) est variable d'une palette à l'autre. Il correspond donc à P variables discrètes et homogènes, avec $P = \text{Nombre de palettes} - 1$. Si l'on doit lancer une série de 50 palettes, il y a donc 49 variables **TiLa**, identifiées par un indice : **TiLa(1,2)**, **TiLa(2,3)**, ..., **TiLa(P, P+1)**. L'étendue de chaque variable **TiLa** est de [1..20] secondes.*

*- La capacité des stocks amont (**StAm**) est variable d'un poste à l'autre. Elle correspond donc à N variables discrètes et homogènes, avec $N = \text{Nombre de postes}$. Si l'on dispose de 4 postes de travail, il y a donc 4 variables **StAm**, identifiées par un indice : **StAm(1)**, **StAm(2)**, ..., **StAm(N)**. L'étendue de chaque variable est de [1..6] palettes.*

Le nombre de solutions possibles est :

*- Pour **TiLa** : $[(\text{Max}(\text{TiLa}) - \text{Min}(\text{TiLa})) + 1]^{(P)}$, avec $\text{Max}(\text{TiLa}) = \text{valeur maximale de TiLa}$, soit 20 secondes et $\text{Min}(\text{TiLa}) = \text{valeur minimale de TiLa}$, soit 1 seconde. On obtient donc :*

$[(20 - 1) + 1]^{(50-1)} = 20^{49}$ combinaisons possibles de Temps inter Lancement entre chaque palette.

(*) Un ensemble de variables homogènes est un ensemble de variables de même type, possédant la même unité et la même étendue de variation, avec le même nombre de valeurs s'il s'agit de variables discrètes.

- Pour $StAm$: $[(Max(StAm)-Min(StAm)) + 1]^N$, avec $Max(StAm)$ = valeur maximale de $StAm$, soit 6 palettes et $Min(StAm)$ = valeur minimale de $StAm$, soit 1 palette. On obtient donc :

$$[(6 - 1) + 1]^4 = 6^4 \text{ combinaisons possibles de Capacité des Stocks Amont.}$$

Le nombre total de solutions en combinant les Temps inter-Lancement (TiLa) et les capacités des stocks amont (StAm) est donc de 2049×6^4 .

Ce deuxième exemple met en évidence la complexité du problème d'optimisation. La difficulté de résolution d'un tel problème fait dire aux spécialistes de l'optimisation qu'il est NP-Complet [PRINS 97, page 31] [CHARON 96, page 117]. La particularité d'un tel problème est qu'il ne peut pas être résolu facilement (et surtout rapidement) par une méthode d'optimisation exacte (voir chapitre III.1). On utilisera plutôt pour la résolution de ce genre de problème des **méthodes approchées** ou **heuristiques**.

□ II.3- Démarche d'optimisation suivie

L'objectif final de cette thèse est de trouver un algorithme d'optimisation capable de trouver automatiquement et rapidement une solution d'un problème fortement complexe. La principale difficulté est d'évaluer la performance et la pertinence d'un tel algorithme, puisque si le système est complexe, on ne connaît pas forcément l'optimum global.

La démarche suivie dans ce travail de recherche consiste, comme on l'a fait dans la partie 6, à contraindre le système en limitant le nombre de types de variables et le nombre de variables, les autres variables étant assimilées alors à des contraintes sur lesquelles on ne peut pas agir. En effet, comme on vient de le montrer sur les exemples précédents, même si l'on prend en compte uniquement un ou deux types de variables, le problème peut rapidement devenir très complexe. On s'attachera donc à travailler sur des cas de complexité progressive, afin de s'assurer de la performance de l'algorithme. Avant de passer à la résolution de cas plus complexes, on validera l'algorithme d'optimisation sur un cas plus simple dont on connaît l'optimum global.

III- Typologie des méthodes d'optimisation

On peut classer les méthodes d'optimisation appliquées à la production en trois grandes catégories :

- Les méthodes d'optimisation à base de règles établies par expérience ou par déduction logique. C'est ce qui a été fait dans la partie 6. Nous avons vu les limites de ce type de méthode dès que le problème se complexifie, par exemple passage d'une production monoproduit à une production multiproduit.
- Les méthodes d'optimisation exactes, dont l'objectif est de déterminer l'optimum exact d'une fonction objectif.
- Les méthodes d'optimisation approchées ou heuristiques, qui sont particulièrement adaptées à la résolution des problèmes fortement combinatoires et NP-Complets.

□ III.1- Méthodes exactes

Ces méthodes sont basées soit sur une résolution algorithmique ou analytique, soit sur une énumération exhaustive de toutes les solutions possibles. Elles s'appliquent donc aux problèmes qui peuvent être résolus de façon optimale et rapidement.

On trouve essentiellement deux types de méthodes :

- **Les méthodes analytiques** [GOLDBERG 94, page 4] sont basées sur l'existence de dérivées, donc sur l'existence d'équations ou de systèmes d'équations, linéaires ou non linéaires. Le principe général consiste à rechercher un extremum hypothétique en déterminant les points de pente nulle dans toutes les directions. La méthode du Gradient ou du Quasi-Newton consiste à se déplacer dans une direction dépendant du gradient de la fonction objectif. Les inconvénients de ces méthodes sont nombreux : elles s'appliquent localement, les extremums qu'elles atteignent sont optimaux au voisinage du point de départ et l'existence de dérivées n'est pas systématique. Ces méthodes sont donc peu efficaces pour résoudre des problèmes d'optimisation en production.
- **Les méthodes algorithmiques** [CHARON 96] : La programmation linéaire [SAKAROVITCH 84, page 117] permet de déterminer l'optimum d'une fonction linéaire tout en respectant des contraintes linéaires. L'algorithme du Simplexe a été développé en 1947 par G.B. Dantzig pour résoudre ce type de problème. Par la suite, cet algorithme a été perfectionné pour pouvoir être adapté à sa programmation informatique. On trouve ainsi une forme matricielle de cet algorithme qui permet de décrire le déroulement des opérations en termes algébriques et moyennant la

résolution de systèmes d'équations linéaires. De nombreuses recherches ont été menées et sont encore menées dans l'application de la programmation linéaire à des cas de production. On peut citer par exemple l'ouvrage de [THIEL 90] qui présente des exemples de traitement de problèmes de production en utilisant la Recherche Opérationnelle. Vient ensuite la Théorie des Graphes [SAKAROVITCH 84, page 15], [PRINS 97], et les méthodes arborescentes (ou "*Branch and Bounds Methods*") qui permettent de traiter des problèmes complexes de façon exacte à partir d'une énumération intelligente de l'espace de solutions. Elles peuvent très bien s'appliquer à des problèmes de taille moyenne, mais pour des problèmes de grande taille, leur durée d'exécution peut devenir prohibitive. Dans son ouvrage, [PRINS 97] présente ces méthodes de façon très claire, pédagogique et concrète, sous la forme d'algorithmes de graphes.

Des recherches plus récentes ont permis de développer la **Programmation Dynamique** et la **Programmation Linéaire en Nombres Entiers**, afin de trouver des solutions exactes à des problèmes réputés difficiles. Mais la résolution de problèmes par ces méthodes n'est pas toujours facile, notamment à cause de la difficulté à caractériser les problèmes susceptibles d'être traités [CHARON 96, Chap. 12].

Enormément de travaux de recherche ont été développés dans le domaine de la recherche opérationnelle afin de résoudre des problèmes qui sont maintenant devenus des standards : "*Problème du sac à dos (Knapsack Problem)*", "*Bin Packing*", "*Problème du Voyageur de Commerce*", etc...., mais l'application de ces méthodes à des cas réels entraîne bien souvent une trop forte simplification du problème qui fait que, bien que le résultat de l'algorithme soit exact, il peut être très éloigné de l'optimum réel.

Pour l'optimisation d'un système de production, dont la complexité peut être très grande en terme de solutions possibles, et qui de plus est soumis à des perturbations aléatoires, l'utilisation de ces différentes méthodes ne semble pas être la plus appropriée.

□ III.2 - Méthodes approchées ou heuristiques

Une méthode approchée ou heuristique est un algorithme d'optimisation qui a pour but de trouver une solution réalisable de la fonction objectif, mais sans garantie d'optimalité. Le principal avantage de ces méthodes est qu'elles peuvent s'appliquer à n'importe quelle classe de problèmes, faciles ou très difficiles, bien ou mal formulés, avec ou sans contrainte. En particulier, elles ne nécessitent pas une modélisation mathématique du problème. Elles semblent être tout à fait adaptées à l'optimisation de systèmes de

production et c'est donc ce type de méthode qui sera utilisé dans la suite de ce travail. Ces méthodes peuvent être classées en 3 familles :

- les méthodes de recherche locale et globale définies par un voisinage,
- les méthodes de recherche aléatoire,
- les méthodes aveugles ou méta-heuristiques.

Avant de rentrer dans le détail de chacune de ces méthodes, il est nécessaire de formuler le problème d'optimisation que l'on se pose et de définir la notion de voisinage.

III.2-a) Formulation du problème d'optimisation (appliquée à la ligne expérimentale)

Soit une fonction objectif, **Fobj = f (TiLa, StAm, StAv, Prio,)**.

On appelle **{S}** l'ensemble fini, mais de grand cardinal, correspondant à toutes les combinaisons de toutes les variables d'entrée.

Le problème d'optimisation de la fonction objectif peut être mis sous la forme :

Minimiser (ou Maximiser) Fobj = f(TiLa, StAm, StAv,....)

Un élément $s_{(i)}$ de **{S} qui donne à Fobj sa valeur minimale (ou maximale) est donc une solution optimale.**

III.2-b) Principe du voisinage d'une solution

On définit le voisinage **V(s_(i))** d'une solution par une transformation élémentaire permettant de passer d'une solution **s_(i)** de **{S}** à une autre solution **s_(j)** de **{S}** avec une faible modification de la structure de la solution **s_(i)**.

Exemples de transformations élémentaires :

- **Complémentation** (cas d'un codage binaire de la solution) :
on remplace un bit quelconque du code par son complément.

Exemple : **1 0 1 0** = solution initiale

$$V(s_{(i)}) = \underline{0} 0 1 0, 1 \underline{1} 1 0, 1 0 \underline{0} 0, 1 0 1 \underline{1}$$

= ensemble des solutions voisines définies par complémentarité

- Autres transformations possibles :

Echange, Insertion / décalage, Inversion [CHARON 96, page 169]... dont le choix dépend du problème à traiter.

[PRINS 97] précise qu'avant de transformer une solution, il faut commencer par choisir initialement cette solution. Cela est généralement fait aléatoirement, mais il est préférable de partir d'une solution initiale assez bonne.

Si $V(s_{(i)}) = \{S\}$, on se ramène à une énumération complète du voisinage d'une solution, ce qui peut s'avérer (très) long en calcul.

Si par contre $V(s_{(i)})$ est trop petit devant $\{S\}$, la recherche est piégée dans un minimum local.

On peut déjà noter le niveau de subjectivité de ce type de méthode, puisqu'il faut non seulement choisir une solution initiale, qui peut être quelconque, mais il faut aussi choisir une transformation élémentaire. On est en droit de se poser la question de savoir si ces choix vont avoir une influence sur le résultat final....

III.2-c) Méthodes de recherche locale définies par un voisinage

On peut définir principalement 2 méthodes qui diffèrent uniquement par l'énumération du voisinage d'une solution :

- **Algorithme du Grimpeur "Hill Climbing" (ou de la descente stochastique) :**

On part d'une solution initiale et on lance l'exploration de son voisinage jusqu'à ce que l'on rencontre une solution meilleure, à partir de laquelle on applique le même principe. Dans ce cas, la recherche locale du voisinage d'une solution se termine

après un nombre variable d'itérations, quand elle n'arrive plus à trouver une solution meilleure que la solution courante. On n'explore donc pas systématiquement tout le voisinage d'une solution.

Présentation de cet algorithme en pseudo-code (minimisation de f) :

Soit $V_{(mini)}$ = voisinage de la solution courante, comportant k solutions.

Soit $s_{(i)} \in V_{(mini)}$, une solution parmi les k solutions de $V_{(mini)}$.

Début

mini = $s_{(0)} \in \{S\}$

Pour [i = 1 à k] et [$s_{(i)} \in V_{(mini)}$]

Si $f(s_{(i)}) < f(mini)$

Alors mini = $s_{(i)}$, i = 0

Fin pour

mini = solution finale

Fin

- Algorithme de la plus grande pente :

L'algorithme est le même que le précédent, mais toutes les solutions du voisinage correspondant à la transformation élémentaire sont énumérées et on ne retient que la meilleure.

Présentation de cet algorithme en pseudo-code (minimisation de f):

Soit $V_{(mini)}$ = voisinage de la solution courante, comportant k solutions.

Soit $s_{(i)} \in V_{(mini)}$, une solution parmi les k solutions de $V_{(mini)}$.

Début

mini = $s_{(0)} \in \{S\}$, mini_v = $s_{(0)}$

Répéter

mini = mini_v

Pour [i = 1 à k] et [$s_{(i)} \in V_{(mini)}$]

Si $f(s_{(i)}) < f(mini_v)$

Alors mini_v = $s_{(i)}$

Fin pour

Jusqu'à ce que mini_v = mini

mini = solution finale

Fin

Le principal avantage de chacun de ces deux algorithmes est qu'ils sont faciles à programmer. En contrepartie, ils ont l'inconvénient de ne pas être adaptés à des fonctions objectives multimodales (avec plusieurs pics et creux) car ils s'arrêtent dès qu'un minimum local est trouvé, à moins de les relancer à partir d'une autre solution initiale (voir figure 2). Pour des problèmes comportant beaucoup de variables, le temps de calcul peut être très long si l'on opte pour une transformation élémentaire exhaustive ou si l'on utilise l'algorithme de la plus grande pente.

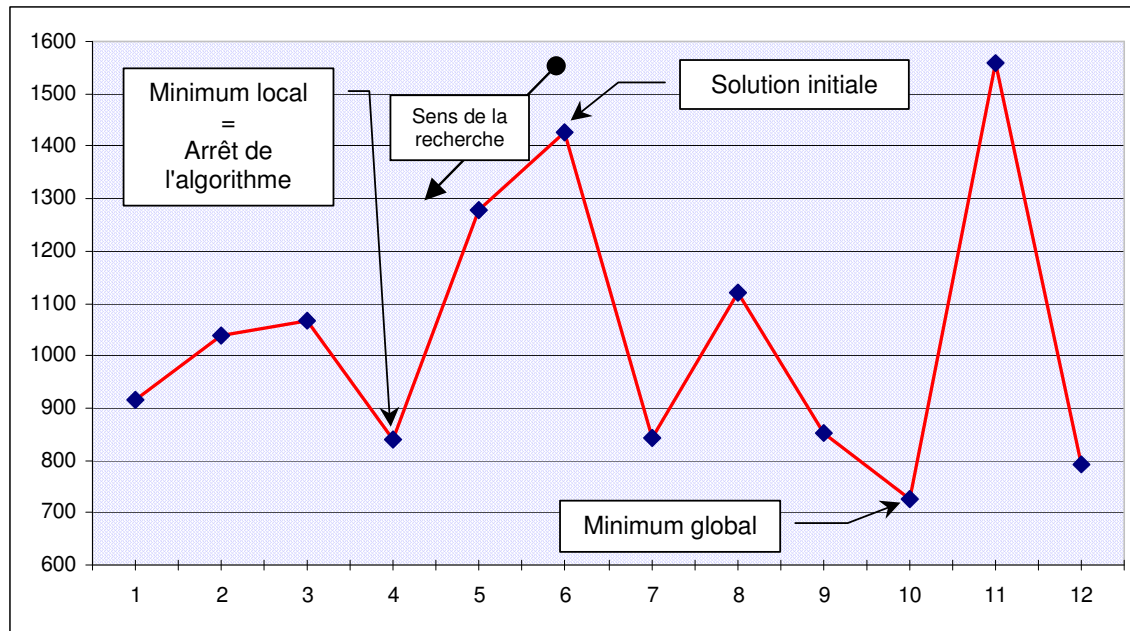


Figure 2 :Exemple de recherche locale

On constate sur ce graphique les limites d'un algorithme de recherche locale. Si la fonction objectif est très bruitée, on risque de tomber dans un minimum local qui est bien loin du minimum global, sans qu'il soit possible d'en sortir. Ces algorithmes sont donc utilisés de façon complémentaire à un autre type d'algorithme qui peut sortir d'un minimum local.

III.2-d) Méthodes de recherche globale définies par un voisinage

Le développement de ces méthodes est issu du constat d'échec des méthodes locales. Le principe recherché est donc de pouvoir sortir d'un optimum local, mais on s'appuie toujours sur une recherche locale avec une transformation élémentaire, que l'on fait ensuite évoluer si elle s'avère insuffisante. Deux méthodes ont été développées très récemment : le Recuit Simulé (Simulated Annealing) et la recherche Tabou (Taboo Search).

- Le Recuit Simulé :

Cet algorithme a été développé en 1982 par deux équipes indépendantes : aux Etats-Unis [Kirkpatrick 83] et en Slovaquie (V. Cernie). Il repose sur une analogie avec la métallurgie et le recuit des métaux : un métal refroidi trop vite présente de

nombreux défauts qui correspondent à des excédents d'énergie interne. L'objectif du recuit est de minimiser ces excédents de façon à obtenir une configuration d'énergie minimale. Pour le réaliser, on réchauffe le métal ce qui a pour effet d'augmenter encore l'énergie interne, mais un réglage judicieux de la température de refroidissement permet de sortir de l'état initial et d'obtenir finalement une énergie interne plus faible. L'application de ce principe à l'optimisation est le suivant : Il est possible, contrairement à un algorithme de recherche locale, d'accepter une dégradation de la fonction objectif avec une certaine probabilité, sachant que cette dégradation pourra entraîner une amélioration ultérieurement.

Présentation de cet algorithme en pseudo-code (minimisation de f) :

Début

Engendrer la solution initiale $s_{(0)}$

Mini = $s_{(0)}$

Répéter

Choix de $s_{(i)} \in V(s_{(i)})$

Calcul de $\Delta d = f(s_{(i)}) - f(\text{mini})$

Si $\Delta d < 0$

Alors mini = $s_{(i)}$! idem algorithme de la descente

Sinon tirer p dans $[0, 1]$ suivant une distribution uniforme

Si $p \leq e^{(-\Delta d / T)}$

Alors mini = $s_{(i)}$! solution moins bonne

Sinon $s_{(i)}$ est rejetée

$T = g(T)$! avec g décroissante, par exemple $T = 0,9995 \cdot T$

Jusqu'à ce que T proche de 0

Fin

Une autre analogie pour expliquer cet algorithme est proposée par [PRINS 97 , page 50] : On considère l'ensemble des solutions comme un paysage de montagnes, comme dans la figure 2. Une pierre lâchée dans un tel paysage va glisser et s'arrêter au fond du premier creux venu : c'est l'analogie à une recherche locale. Une balle élastique peut, par contre, rebondir et contourner les obstacles. L'énergie initiale de la balle va diminuer lentement à chaque rebond, et la balle aura de fortes chances de finir plus bas que la pierre : c'est l'analogie au recuit simulé.

Le principal avantage du Recuit simulé est donc de pouvoir sortir d'un minimum local, en fonction d'une probabilité d'acceptation liée à une fonction exponentielle, appelée transformation de Gibbs-Boltzmann. Par contre, elle présente l'inconvénient de faire intervenir des paramètres intrinsèques au recuit simulé : la valeur initiale de la température et le coefficient de décroissance de la température. [CHARON 96, page 177] présente plusieurs modèles de recuit en fonction de ses différents paramètres. Le réglage de ces paramètres est assez délicat et repose sur de nombreux essais. Toutefois, on constate que les utilisateurs de cet algorithme donnent au coefficient de décroissance une valeur comprise entre 0,85 et 0,95. Quant à la température initiale, celle-ci est déterminée empiriquement ou par des méthodes plus rigoureuses (C.R. Aragon, D.S. Johnson, L.A. McGeoch, C. Shevon dans [ARAGON 89]). C'est ce qui peut paraître paradoxal et déroutant dans cet algorithme, puisqu'il faut tenter d'optimiser ses paramètres intrinsèques tout en optimisant la fonction objectif.

Malgré la jeunesse de cet algorithme, les applications à l'optimisation des systèmes de production sont nombreuses. Deux articles sont particulièrement intéressants et en rapport avec ce travail de recherche :

[MUKHOPADHYAY 98] propose d'utiliser un algorithme de Recuit Simulé pour optimiser l'équilibrage des systèmes de fabrication flexibles (FMS). La variable d'entrée est le séquençement des ordres de fabrication. De plus, une évolution de l'algorithme de base est proposée, dans laquelle la variation de la température est parabolique. L'utilisation de cet algorithme permet de réduire de 50% le déséquilibre du système.

[YING-CHIN 98] utilise un Recuit Simulé pour déterminer l'implantation optimale des machines desservies par un système de manutention automatisée. Plusieurs stratégies de réduction de température sont proposées : décroissement linéaire, décroissement en paliers constants, décroissement en paliers concaves et décroissement en paliers convexes. Il propose de fixer la température initiale dans des valeurs comprises entre 3,5 et 5,5. Il utilise aussi un nouveau paramètre appelé température de refroidissement, afin d'empêcher la température de descendre trop bas, ce qui risquerait d'empêcher le déplacement vers une autre

solution. De nombreuses expérimentations ont été menées afin de déterminer les paramètres du Recuit Simulé les mieux adaptés en fonction des problèmes à traiter.

- La méthode TABOU :

Cette méthode est encore plus récente que le Recuit Simulé, puisqu'elle a été proposée par Glover en 1986. Dans [GLOVER 97], l'auteur de cette méthode détaille les différents algorithmes à base de listes Tabou. Depuis, de nombreuses contributions ont été apportées à cette méthode, notamment à l'Ecole Polytechnique Fédérale de Lausanne où [WIDMER 91] l'a appliquée à l'optimisation des ateliers flexibles.

Encore une fois, le point de départ de cette méthode est identique à une recherche locale avec une transformation de voisinage. La différence provient du fait que l'on va sur la meilleure solution voisine, même si elle est plus mauvaise que la solution initiale. L'autre originalité de cette méthode est la construction d'une liste de mouvements tabous \mathbf{T} , qui correspondent aux solutions testées dans un passé proche et sur lesquelles on s'interdit de revenir. A partir de la solution courante $\mathbf{s}_{(i)}$ et pour la remplacer, on choisit dans $\mathbf{V}(\mathbf{s}_{(i)}) - \mathbf{T}$ la solution $\mathbf{s}_{(j)}$ qui minimise la fonction objectif f , tout en ajoutant $\mathbf{s}_{(i)}$ à la liste \mathbf{T} .

La difficulté majeure de cette méthode est la gestion de cette liste Tabou. En pratique, cette liste est de taille réduite (comprise entre 3 et 12) et elle fonctionne en FIFO : quand la liste est pleine, on supprime la solution qui s'y trouve depuis le plus longtemps et on la remplace par la nouvelle solution à interdire.

Les recherches et applications dans le domaine de la production sont encore une fois nombreuses. On pourra trouver par exemple dans [TAILLARD 89] et dans [WIDMER 89] des applications de cette méthode à l'optimisation de l'ordonnancement dans une chaîne de montage et dans un atelier de type Job Shop.

Pour conclure sur ces méthodes de recherche globale définies par un voisinage, on peut noter leurs différences :

- Le Recuit Simulé est stochastique et sans mémoire, alors que la méthode Tabou est déterministe et possède une mémoire sous la forme d'une liste.
- Les paramètres intrinsèques sont moins nombreux pour la méthode Tabou que pour le Recuit Simulé.
- Le Recuit Simulé ne s'intéresse qu'à un seul voisin à chaque itération, la méthode Tabou nécessite l'exploration exhaustive du voisinage, ce qui peut prendre

beaucoup de temps si le cardinal de $V(\mathbf{s}_{(i)})$ est élevé. Dans un tel cas, on se contente donc d'un échantillon du voisinage, prélevé de façon aléatoire ou non.

Dans la littérature, les auteurs qui ont tenté de comparer ces deux méthodes ont des avis assez partagés :

[PRINS 97, page 52] pense, sans le démontrer, que la méthode Tabou est meilleure que le Recuit Simulé à temps d'exécution égal.

[CHARON 96, page 184] cite les travaux de S. de Amorin, J.P. Barthélemy et C.C. Ribeiro qui ont comparé la méthode du Grimpeur, le Recuit Simulé et la méthode Tabou sur un même problème. Ces travaux démontrent que les deux méthodes de recherche globale donnent de meilleurs résultats qu'une recherche locale, mais qu'il y a peu de différence entre le Recuit Simulé et Tabou, en tout cas quand le nombre de solutions n'est pas trop important. Quand le nombre de solutions est très important, le Recuit Simulé permet d'obtenir un optimum en un temps de calcul plus court que la méthode Tabou.

En conclusion, il semble donc très délicat de chercher à comparer chacune de ces méthodes. Le choix des paramètres ou du problème à traiter peut faire pencher la balance d'un côté ou de l'autre.

III.2-e) Méthode d'exploration aléatoire

On trouve peu de références dans la littérature au sujet de cette méthode. Cela est peut être dû au fait qu'il est impropre d'appeler "Méthode" une recherche qui est purement aléatoire. D'autre part, cette exploration aléatoire peut sembler trop peu efficace aux yeux des chercheurs en optimisation combinatoire. Cependant, la recherche aléatoire d'un optimum a été utilisée avec succès avant que les méthodes de recherche globale ne soient développées [GOLDBERG 94, page 6].

Le principe de cette recherche aléatoire est donc des plus simples : on génère "au hasard" une solution $\mathbf{s}_{(i)}$ puis on évalue la fonction objectif f avec cette solution et on recommence ces deux étapes jusqu'à ce qu'un nombre d'itérations (arbitraire ou pas) fixé par l'utilisateur soit atteint.

Malgré le manque d'intérêt de cette recherche par les spécialistes, on peut quand même penser que pour les problèmes de très grande taille et qui ont une fonction objectif multimodale, compte tenu du temps de calcul nécessaire à n'importe laquelle des méthodes globales vues précédemment, cette recherche aléatoire permet :

- soit de dégrossir le problème avant de l'explorer par une méthode globale ou locale de voisinage, et sur ce point, beaucoup de spécialistes sont d'accord;

- soit de se contenter de cette recherche si la nature et la complexité du problème est telle qu'une méthode de voisinage ne permet pas l'obtention d'un gain significatif par rapport à une solution obtenue aléatoirement.

Ce dernier point sera repris lors de l'optimisation du système expérimental, principalement dans le but de comparer l'efficacité des algorithmes de recherche (voir Partie 8).

Un autre intérêt de cette recherche aléatoire est qu'elle permet de se faire une idée de la "*surface de réponse*" de la fonction objectif. Après avoir généré des solutions "au hasard", on peut tenter de les représenter sur un réseau tridimensionnel. La condition de cette représentation est que le nombre de variables d'entrée soit de deux au maximum. La troisième dimension est réservée à la représentation de la réponse de la fonction objectif pour une combinaison donnée des 2 variables d'entrée. Si cette surface présente peu d'irrégularités, on peut extrapoler en disant que le problème n'est pas complexe. Par contre, si les irrégularités sont nombreuses et d'amplitude importante, le problème est certainement complexe. Mais cette démarche est difficilement généralisable à des problèmes comportant plus de deux variables d'entrée et surtout elle est très dépendante du nombre de solutions aléatoires générées. Malgré cela, elle constitue peut-être une piste à des travaux de recherche complémentaires.

III.2.f) Algorithmes génétiques

Les algorithmes génétiques ont été développés au début des années 60 par John Holland [HOLLAND 75 et 92] et ses travaux ont été poursuivis par un de ses anciens élèves [GOLDBERG 94].

L'idée des algorithmes génétiques est encore une fois assez simple et concrète : reproduire les mécanismes naturels de l'évolution des espèces en fonction de leur adaptation au milieu dans lequel elles évoluent. Les individus les mieux adaptés ont plus de chance de survivre et donc de se reproduire.

A chaque génération, les caractéristiques génétiques (les chromosomes) des individus les mieux adaptés seront répandues dans une majorité de la population.

Afin d'éviter que l'ensemble de la population converge vers une même caractéristique génétique (notion de consanguinité), certains individus n'hériteront pas complètement des caractéristiques de leurs parents, et subiront une mutation. Cette mutation permettra peut-être à ces individus de survivre à une évolution brutale de leur milieu naturel, là où les autres individus mourront. C'est peut-être l'absence de mutation dans les

chromosomes des dinosaures qui les a tous fait disparaître suite à une évolution brutale de leur environnement (la chute d'une météorite qui a bouleversé le climat de la planète). Cette espèce était peut-être trop parfaite et aucun individu n'a pu survivre...

Le mécanisme de reproduction des espèces est un mécanisme qui est aléatoire, mais il ne faut pas confondre les algorithmes génétiques avec une recherche purement aléatoire : "L'algorithmes génétique est un exemple de procédures d'exploration qui utilise un choix aléatoire comme outil pour guider une exploration hautement intelligente dans l'espace des paramètres codés" [GOLDBERG 94, page 6].

Les algorithmes génétiques correspondent donc plutôt à ce que l'on pourrait appeler des **méthodes Pseudo-Aléatoires**.

Par analogie avec la reproduction des espèces, la terminologie utilisée dans les algorithmes génétiques est la suivante :

- Une **POPULATION** est composée d'**INDIVIDUS** porteurs de **CHROMOSOMES** formés de **GENES**.
- La population évolue grâce à la **REPRODUCTION** vers une autre population formés des individus de la **GENERATION** suivante.
- La reproduction se décompose en 2 étapes : **SELECTION** des individus et **CROISEMENT** (ou **CROSSOVER**) des individus deux à deux.

La figure 3 illustre cette terminologie et présente la structure d'une espèce.

	Individu 1	Individu 2	Individu	Individu i
Génération 0	CHROM OSOME GENE1 GENE2 GENEK	CHROM OSOME GENE1 GENE2 GENEK	CHROM OSOME GENE1 GENE2 GENEK		CHROM OSOME GENE1 GENE2 GENEK
Génération 1	CHROM OSOME GENE1 GENE2 GENEK	CHROM OSOME GENE1 GENE2 GENEK			
.....					
Génération g	CHROM OSOME GENE1 GENE2 GENEK				

Figure 3 :Structure d'une espèce

Dans [GOLDBERG 94, page 25] et dans [CHARON 96, page 187], on peut trouver des termes complémentaires et une comparaison de ces termes utilisés par les spécialistes de la génétique naturelle et de la génétique artificielle : le **GENOTYPE** est l'ensemble du

matériel génétique, le **PHENOTYPE** correspond aux caractéristiques que l'on peut mesurer chez un individu. Les valeurs que prennent les gènes sont appelées les **ALLELES**.

Contrairement aux autres méthodes de recherche, les algorithmes génétiques ne considèrent pas qu'une seule solution à la fois, mais plusieurs solutions correspondant aux chromosomes des individus. Cette particularité des algorithmes génétiques est appelée "**PARALLELISME**" [RENDERS 95, page 74]. Ce parallélisme est essentiel pour le mécanisme de recombinaison des individus. Les algorithmes génétiques sont donc particulièrement bien adaptés au traitement sur des ordinateurs parallèles, ce qui permet en plus de diminuer notablement leur temps de calcul, qui peut paraître trop important aux yeux de certains auteurs [PRINS 97, page 54].

Les caractéristiques des algorithmes génétiques semblent tout à fait adaptées à l'optimisation des systèmes de production : ils ne nécessitent pas de modélisation mathématique du problème, ils peuvent être couplés à un simulateur de flux qui sera utilisé comme une boîte noire par l'algorithme génétique (voir figure 4). C'est ce principe qui sera utilisé dans la partie 8 pour l'optimisation de la ligne de production.

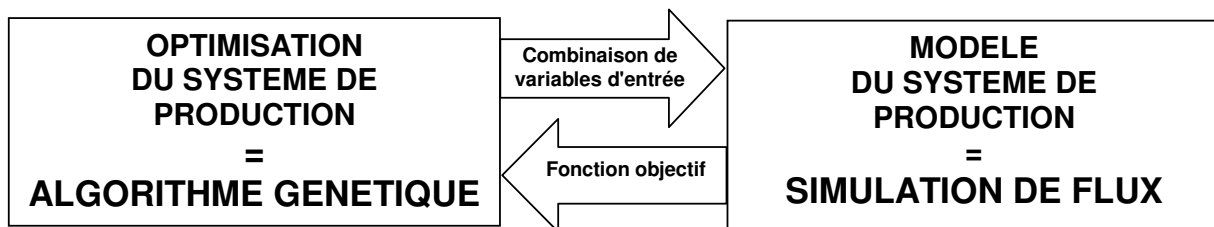


Figure 4 : Couplage entre Algorithme Génétique et Simulation de Flux

Ils sont capables, plus que tout autre algorithme, de traiter des problèmes très fortement combinatoires comme c'est le cas pour le système expérimental servant de support à ce travail. Le reproche que l'on pouvait leur faire jusqu'à présent était de consommer un temps de calcul assez important. L'arrivée de micro-ordinateurs surpuissants et des possibilités de traitement parallèle les rend compétitifs en terme de rapidité. Ce n'est que depuis le début des années 90 que ces algorithmes ont commencé à être utilisés pour l'optimisation des systèmes de production. Depuis, plusieurs équipes de recherche, françaises et internationales, développent et perfectionnent ces algorithmes à des cas de production très variés. En France, les équipes de M.C. Portmann à l'Ecole des Mines de Nancy et d'H. Pierreval, à l'Institut Français de Mécanique Avancée, sont les initiateurs de l'utilisation de ces algorithmes en production.

Les quelques références bibliographiques suivantes permettent de se donner une idée des applications rencontrées en production :

- [BERARD 97] propose d'utiliser un algorithme génétique couplé à un simulateur de flux pour optimiser l'ordonnancement d'un atelier de fabrication de semi-conducteurs et d'un atelier de chimie fine . Dans cet article, un graphique permet de comparer les différentes méthodes d'optimisation en fonction du nombre de produits à traiter et du nombre de machines dont on dispose. Ce graphique fait clairement ressortir que les algorithmes génétiques sont les seuls à pouvoir traiter des problèmes de très grande taille (70 produits et 20 machines). Ils sont meilleurs que le Recuit Simulé (50 produits et 10 machines) et que les méthodes arborescentes de type Branch and Bounds (15 produits et 15 machines). L'autre intérêt de cet article est l'utilisation de la simulation de flux, couplée à l'algorithme génétique. Cette liaison entre deux outils est aussi utilisée par [TAUTOU 95], [PARIS 96] et [PLAQUIN 97]. *Ce couplage est aussi utilisé dans ce travail de recherche (voir Partie 8).*
- [TAUTOU 95] présente le principe de l'optimisation des systèmes de production par les algorithmes génétiques en s'appuyant sur un exemple concret : la minimisation du coût de fabrication d'une ligne de fabrication de yaourts avec comme variables d'entrée la taille du silo, la taille de l'entrepôt, et le choix entre deux méthodes de fabrication. Plusieurs essais ont été réalisés avec des valeurs différentes de probabilité de mutation (entre 0,1 et 0,2) et de taille de population (entre 15 et 30). Dans ce travail, les auteurs montrent qu'une probabilité de mutation assez forte (0,2) associée à une taille de population importante (30), permettent d'obtenir le coût minimal en un minimum de générations.
- [BONNEVILLE 95] utilise les algorithmes génétiques pour optimiser les gammes d'assemblage. L'algorithme génétique est utilisé à la fois pour générer des gammes d'assemblage et pour les évaluer. L'algorithme démarre sur une population initiale de gammes d'assemblage données par un expert. La fonction objectif prend en compte le nombre de sous-ensembles générés par l'algorithme, l'optimisation consiste à minimiser le nombre de sous-ensembles d'une gamme d'assemblage.
- [WELLMAN 95] est l'un des premiers articles à proposer l'utilisation des algorithmes génétiques pour optimiser une ligne d'assemblage automatisée asynchrone. Une comparaison est faite avec la méthode du Quasi-Gradient. Dans ce travail, l'algorithme génétique est une nouvelle fois couplé à un simulateur de flux, et l'optimisation consiste à déterminer la taille des stocks tampons entre 2 machines sur une ligne comportant 3 machines. La fonction objectif correspond à une maximisation du taux de production (*Throughput*) en produits assemblés par unité de temps. Bien que l'algorithme génétique utilisé ne donne pas de meilleurs résultats que la méthode du Quasi-Gradient, il permet, contrairement au Gradient, de proposer plusieurs bonnes solutions. Le concepteur d'une ligne d'assemblage a donc le choix entre plusieurs configurations optimales. L'algorithme génétique a aussi l'avantage d'être entièrement automatique, alors que la méthode du Gradient fonctionne en interaction avec l'utilisateur. Cette application est assez proche du problème soulevé dans ce travail de

recherche, à l'exception qu'elle se situe au niveau de la conception d'une ligne d'assemblage.

- [GROPETTI 97] s'intéresse lui aussi à l'utilisation des algorithmes génétiques pour l'optimisation des gammes d'assemblage. La fonction objectif est multicritère et prend en compte le coût de l'assemblage qui doit être minimisé, le temps de cycle de l'assemblage qui doit être minimisé, la fiabilité du produit qui doit être maximisée, les coûts de maintenance qui doivent être minimisés, la reconception ("*redesign*") du produit qui doit être minimisée. Les résultats sont convaincants et l'auteur indique que son travail peut être appliqué à des tâches autres que l'assemblage.
- [KAZEROONI 97] propose d'utiliser un algorithme génétique pour constituer des familles de pièces et des îlots de production. La fonction objectif est obtenue en maximisant le nombre de zéro dans une matrice de similarité (MCS) conçue par l'auteur. Les individus de l'algorithme génétique correspondent aux différents circuits possibles que peut suivre une pièce.
- [ISLIER 98] présente un cas de logistique externe et utilise un algorithme génétique pour optimiser l'implantation de locaux industriels. La fonction objectif prend en compte 3 critères : minimisation de la charge de transport, maximiser la compacité des surfaces des locaux, et minimiser la différence entre les surfaces demandées et les surfaces disponibles. Les gains obtenus sont supérieurs de 30% par rapport à une méthode d'optimisation classique (non citée).
- [GERO 98] traite aussi d'un cas d'optimisation de l'implantation de locaux, mais de type tertiaire (bureaux). Cette application doit pouvoir être transposée sans trop de difficulté à un cas d'implantation industrielle.
- [GAURY 97] utilise un algorithme génétique couplé à un simulateur pour tester plusieurs stratégies de pilotage en flux tiré (Kanban , ConWip et Hybride) sur un modèle d'atelier générique. L'objectif de ce travail est de déterminer une configuration optimale en minimisant le nombre moyen de produits en encours. La configuration optimale trouvée par l'algorithme comporte des boucles Kanban seulement en début et fin de ligne, les machines du milieu fonctionnant en flux poussé. Une boucle ConWip permet en plus de contrôler le niveau d'encours.

La principale difficulté rencontrée avec les algorithmes génétiques est le **CODAGE** du problème sous une forme propre à la génétique : l'objectif est de pouvoir coder efficacement chaque solution du problème. Plusieurs possibilités existent :

- **Le Codage Binaire des solutions :**

C'est la méthode choisie par le créateur des algorithmes génétiques [HOLLAND 75] et développée ensuite par son élève [GOLDBERG 94, Chap. 2] sous le nom de "Théorie des Schémas". Cette théorie suggère de prendre les alphabets les plus réduits possibles pour coder le problème, donc par exemple un codage binaire pour coder des entiers ou des réels. Ce codage permet de faciliter la conception du croisement entre les individus. L'inconvénient majeur de ce codage est qu'une petite variation

consistant à changer le bit de poids fort correspond à une modification importante de la valeur entière qui est codée (quand le codage est de type binaire naturel).

- **Le Codage Réel des solutions :**

[RENDERS 95, page 67] propose une version des algorithmes génétiques appelée "Algorithmes Génétiques Codés Réels". Ceux-ci semblent être mieux adaptés aux besoins et aux pratiques industrielles [DAVIS 91] pour la résolution des problèmes réels. Parmi ces algorithmes génétiques Codés Réels, on distingue une classe particulière appelée "Algorithmes Evolutionnistes" [BEASLEY 98] qui ont la particularité d'introduire dans le chromosome des gènes supplémentaires représentant les amplitudes des mutations, si bien que ces dernières grandeurs sont elles aussi optimisées dans le processus de recherche. En général, ces algorithmes Codés Réels sont plus pratiques d'emploi, plus rapides et plus précis que les Algorithmes Génétiques Codés Binaires [JANIKOW 91].

[RENDERS 95, page 68] a comparé ces 2 types de codage et conclut que les Algorithmes Codés Réels sont bien adaptés à l'optimisation des problèmes industriels réels.

[DESSALLES 96, chap. 5] fait aussi une comparaison entre ces deux types de codage. Il note que l'efficacité de l'algorithme et son pouvoir d'invention sont directement liés au codage. La principale différence entre les deux codages se fait sentir au moment du croisement des chromosomes : un codage binaire présente un nombre de **SITES de CROISEMENT** beaucoup plus important qu'un codage réel.

Exemple :

Soit à coder un problème comportant 6 variables entières dont la valeur est comprise entre 0 et 256 :

Codage Réel : 202 - 17 - 145 - 94 - 224 - 102

Codage Binaire : 11001010 - 00010001 - 10010001 - 01011110 - 11100000 – 01100110

Dans le cas du codage réel, on dispose de 5 sites de croisement (entre chaque valeur de variable), alors que dans le cas du codage binaire, il y a 47 sites de croisement possibles.

Le second codage offre donc beaucoup plus de possibilités de recombinaisons à l'algorithme et Dessalles pense qu'il faut proscrire un codage qui considère les variables comme des entités atomiques que le croisement n'a pas la possibilité de couper.

Encore une fois, deux écoles s'opposent avec des arguments solides, mais il est difficile de faire un choix. A moins de tester les deux possibilités de codage, ce qui augmenterait encore plus le nombre d'itérations !

Le choix pertinent et scientifique du codage d'un problème d'optimisation par un Algorithme Génétique semble donc être une piste de recherche à explorer

Dans le cas de l'optimisation du système expérimental, le problème du codage consisterait donc à passer d'une solution composée de plusieurs variables, $TiLa(i)$, $StAm(j)$, $StAv(k)$, $PRio(m)$, etc..., à une structure de données correspondant à un chromosome.

Comme on vient de le voir, plusieurs possibilités sont offertes, mais pour les raisons évoquées par [RENDERS 95], c'est le codage réel qui semble être le plus approprié à ce type de problème, avec toutes les réserves émises par Dessalles.

Un développement plus important de l'exploitation d'un algorithme génétique pour optimiser le système expérimental est présenté en détail dans la Partie 8.

Il existe donc plusieurs formes d'Algorithmes Génétiques. Mais ils sont tous constitués des opérateurs de **SELECTION**, puis de **CROISEMENT** puis de **MUTATION**. Les différences entre les algorithmes génétiques résident dans le codage du problème (binaire ou réel), mais aussi dans les méthodes de sélection des individus (méthode probabiliste ou déterministe), dans le mécanisme de croisement des gènes et enfin dans le taux de mutation.

La Sélection peut s'effectuer de différentes façons :

- Généralement elle est réalisée à partir d'un tirage probabiliste basé sur le principe de la "**Roue de la fortune**" découpée en secteurs. Chaque secteur correspond à un individu de la population, la superficie de chaque secteur étant proportionnelle à la réponse de la fonction objectif : plus un individu est adapté, plus le secteur qui lui correspond est grand. C'est la méthode préconisée par [HOLLAND 75] et [GOLDBERG 94] et c'est aussi la plus employée.

Exemple : Cas d'un codage réel du problème et recherche d'un minimum

Individu	Chromosome	Fonction Obj. = f_i	Probabilité de sélection
1	1 3 5 2 1 8	185	$(1 - 185/1104)/3 = 0,277$
2	9 6 1 5 4 3	560	$(1 - 560/1104)/3 = 0,164$
3	8 7 6 3 7 5	101	$(1 - 101/1104)/3 = 0,303$
4	1 5 9 8 3 2	258	$(1 - 258/1104)/3 = 0,255$
		Somme =1104	Somme =1

On obtient donc une roue de la fortune découpée en 4 secteurs (pourcentages arrondis à la valeur entière) :

Pour réaliser la sélection, il suffit de faire "tourner" la roue de la fortune autant de

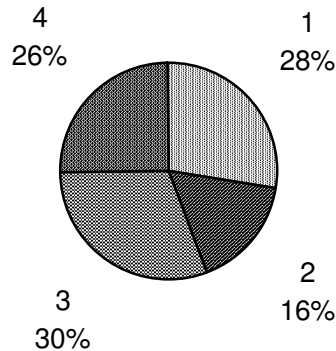


Figure 5 : Sélection par roue de la fortune

fois qu'il y a d'individus dans une population. Comme on le voit sur l'exemple, l'individu 3 qui a la meilleure fonction objectif, se voit attribuer une portion plus importante de la roue. Si la loi de distribution est répartie uniformément de 0 à 1, chaque tirage a une probabilité de 0,303 de tomber sur le secteur qui correspond à l'individu 3. De cette façon, les individus bien adaptés ont un plus grand nombre de chances d'être utilisés pour la reproduction de la génération suivante.

- **D'autres possibilités de sélection existent.** Pour éviter qu'un individu trop parfait ne constitue presque exclusivement la génération suivante, et afin de préserver une certaine diversité, il est possible de procéder à un changement d'échelle statique ou dynamique, linéaire ou exponentiel de la fonction objectif [BAKER 85]. On peut aussi décider de ne sélectionner qu'une partie de la population en fonction d'un taux de sélection [DAVID 91]. La stratégie élitiste [GOLDBERG 1994, page 131] consiste quant à elle à conserver le meilleur individu d'une génération et à le reproduire tel quel pour la génération suivante (clonage).

Le Croisement a pour objectif de recombinaison les chromosomes d'une paire d'individus sélectionnés (les parents) afin de créer une nouvelle paire d'individus (les enfants). Plusieurs méthodes de croisement existent :

- **Le croisement simple** dans lequel un seul site de croisement est choisi.

Exemple : (voir figure 6)

On suppose que les quatre tirages de la roue de la fortune ont sélectionné deux individus N°3, un individu N°1 et un individu N°4 (l'individu N°2 est donc éliminé pour le croisement).

On commence par "marier" au hasard les individus sélectionnés. Par exemple l'individu N°1 avec un des deux individus N°3, et l'individu N°4 avec le second individu N°3.

On détermine ensuite aléatoirement un site de croisement et on croise les chromosomes des deux parents pour obtenir les chromosomes des deux enfants de la génération suivante.

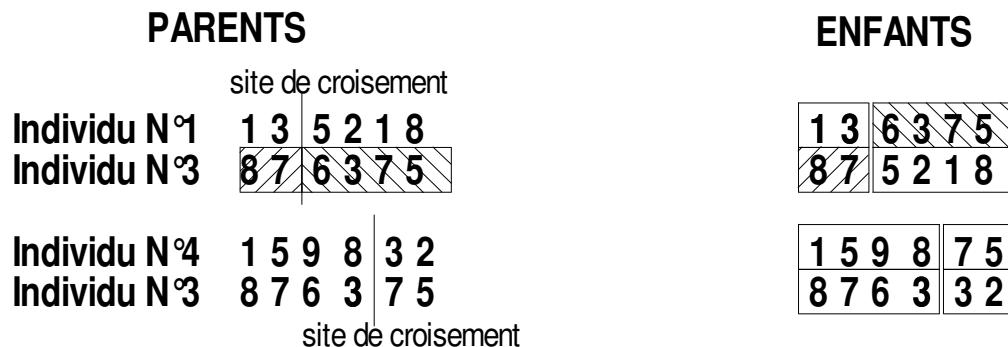


Figure 6 : Croisement simple

- **Le croisement double** et plus généralement à plusieurs sites, qui permet d'échanger des morceaux de chromosomes compris entre deux sites de croisement.
- **Le croisement moyenné** qui propose de calculer la moyenne de certains gènes choisis au hasard chez les deux parents.
- **Le croisement multiparental**, qui consiste à combiner les gènes de plus de deux parents en utilisant par exemple la méthode du simplexe pour orienter la recombinaison [RENDERS 96, page 70].

La Mutation permet de transformer au hasard le codage d'un individu afin d'apporter une certaine diversité dans la population et empêcher que celle-ci converge trop vite vers un seul type d'individu parfait, incapable de sortir d'un minimum local. La mutation est réalisée en modifiant un gène d'un individu pris au hasard. L'amplitude de la modification peut elle même être déterminée aléatoirement ou correspondre à une valeur fixée. Très

couramment, la mutation consiste à rajouter + ou - 1 à la valeur d'un gène, suivant une probabilité de mutation qui est en général assez faible : comprise entre 0,001 et 0,005.

[GOLDBERG 94, page 17] indique que l'opérateur de mutation joue un rôle secondaire dans la mise en œuvre des Algorithmes Génétiques, mais qu'il est cependant nécessaire pour éviter une perte de matière génétique.

Présentation d'un Algorithme Génétique en pseudo-code (minimisation de f) :

Début

Création aléatoire d'une population de k individus

! k paramètre de l'algorithme déterminé empiriquement

Répéter

***Evaluation* de la fonction objectif f de chaque individu**

***Sélection* des meilleurs individus**

! plusieurs processus de sélection possibles

***Reproduction/croisement* entre deux individus sélectionnés**

! obtention des deux enfants issus des deux gènes sélectionnés

! plusieurs types de croisement possibles

***Mutation* : transformation aléatoire d'un élément d'un gène d'un des individus de la population**

! permet de ne pas rester dans un minimum local

Jusqu'à condition d'arrêt

! plusieurs types de condition d'arrêt

Fin

Conclusion

Lors des expérimentations réalisées dans la partie 6, nous avons pu constater les difficultés qu'il y a à optimiser un système de production qui comporte de nombreuses variables de pilotage. Une optimisation basée uniquement sur l'utilisation de relations théoriques et de règles expérimentales n'est pas suffisante pour traiter des cas complexes.

Dans cette partie, nous avons donc présenté le principe de l'utilisation d'un algorithme d'optimisation couplé à un modèle de simulation de flux. Comme on l'a vu dans la partie 4 puis dans la partie 6, le modèle de simulation d'un système de production, une fois qu'il est validé, réagit comme une "boîte noire" : il reçoit en entrée des variables de pilotage, il fournit en sortie des variables d'analyse, définies par l'utilisateur. Utilisé seul, le modèle de simulation nécessite donc d'être "alimenté" par des combinaisons des variables de pilotage de façon répétitive par l'utilisateur, comme cela a été fait dans la partie 6. Pour des cas simples, où le nombre de variables et de combinaisons est faible, on peut, grâce au temps de réponse très court de la simulation, simuler de façon exhaustive toutes les combinaisons. Par contre, dès que le nombre de combinaisons devient important, il n'est plus question d'utiliser cette démarche. C'est dans ce cas de figure qu'il devient intéressant d'utiliser un algorithme d'optimisation.

Parmi tous les algorithmes d'optimisation, nous avons éliminé ceux qui nécessitent une modélisation mathématique du problème, pour ne retenir que les algorithmes de recherche approchée. Nous avons démontré que l'utilisation d'un algorithme de recherche locale, tel que l'algorithme de la descente stochastique, bien que simple à programmer, présente des risques de blocage dans un minimum local. Si la fonction objectif est fortement "bruitée", avec de nombreux "pics" et "creux", ce type d'algorithme peut se retrouver piégé sur une solution très éloignée d'une "bonne" solution. Nous avons donc orienté notre choix sur les algorithmes de recherche globale à base d'heuristiques, et en particulier sur les algorithmes génétiques. Dans la littérature, ils sont souvent comparés à l'algorithme du Recuit simulé et à l'algorithme de recherche Tabou, mais nous les avons choisis car ils semblent être plus efficaces pour traiter des problèmes de grande taille. Leurs principaux avantages par rapport au Recuit Simulé ou à la recherche Tabou est qu'ils explorent plusieurs solutions en parallèle et qu'ils sont très "flexibles" : on peut les adapter facilement au cas à traiter.

Dans la partie 8, nous allons donc développer et appliquer ces algorithmes génétiques sur la ligne expérimentale, dans le but de créer un outil d'optimisation couplé à la simulation.

Bibliographie de la partie 7 : "Optimisation de la production"

[MUKHOPADHYAY 98] : " FMS machine loading : a simulated annealing approach "
International Journal Of Production Research 1998, Vol. 36 N°6 pages 1529 - 1547
S.K. Mukhopadhyay, M.K. Singh, R. Srivastava

[WIDMER 89] : "A new heuristic method for the flow shop sequencing problem"
European Journal of Operational Research, vol. 41 (1989) pages 186 à 193
M. Widmer, A. Heritz

[WIDMER 91] : " Modèles mathématiques pour un gestion efficace des ateliers flexibles"
Presses Polytechniques et Universitaires Romandes, 1991
M. Widmer

[WELLMAN 95] : " A Genetic Algorithm Approach to Optimization of asynchronous automatic assembly systems "
International Journal of Flexible Manufacturing Systems, 7 (1995)
M. A. Wellman, D. D. Gemmill

[KAZEROONI 97] : " A genetic algorithm based cell design considerate alternative routing"
Computer Integrated Manufacturing Systems, Vol. 10 N° 2 pages 93-107 , 1997
M. Kazerooni Lee, H. S. Luong, K. Abhary

[TAUTOU 95] : "Using evolutionary algorithms and simulation for the optimization of manufacturing systems"
Actes du congrès IEEE 1995, pages 509 à 524
L. Tautou, H. Pierreval

[MOLET 97] : "Une nouvelle gestion industrielle"
Editions Hermes, 2^{ème} édition, 1997
H. Molet

[BONNEVILLE 95] : "A genetic algorithm to generate and evaluate assembly plans"

Actes du congrès IEEE 1995, pages 231 à 239

F. Bonneville, C. Perrard, J.M. Henrioud

[GERO 98] : "Space Layout planning using an evolutionary approach"

Artificial Intelligence in Engineering Vol. 12 / 1998 pages 149-162

J. S. Gero, Jun. H. Jo

[PARIS 96] : "Une méthode d'optimisation-simulation par algorithme évolutionniste en gestion de production juste-à-temps"

RAIRO-APII-JESA, vol. 30, N° 7/1996, pages 929 à 944

J.L. Paris, H. Pierreval, L. Tautou

[RENDERS 95] : "Algorithmes génétiques et réseaux de neurones"

Editions Hermes, 1995

J.M. Renders

[DESSALLES 96] : "L'ordinateur génétique"

Editions Hermes, 1996

J.L. Dessalles

[GOLDBERG 94] : "Algorithmes Génétiques"

Editions Addison-Wesley, 1994 (traduction française de l'ouvrage anglais paru en 1991 "Genetic Algorithms")

D. E. Goldberg

[PRINS 97] : "Algorithmes de Graphes"

Editions Eyrolles, 2ème tirage, 1997

C. Prins

[CHARON 96] : "Méthodes d'optimisation combinatoire"

Editions Masson, 1996

I. Charon, A. Germa, O. Hudry

[KIRKPATRICK 83] : "Optimization by Simulated Annealing"

Revue Science, N°220, pages 671 à 680, 1983

Kirkpatrick, Gelatt, Vecchi

[GLOVER 97] : "Tabu Search"

Kluwer academic Publishers, 1997

F. Glover, M. Laguna

[HOLLAND 75] : "Adaptation in Natural and Artificial Systems"

Ed. The University of Michigan Press, 1975 (2ème édition en 1992)

J.H. Holland

[ISLIER 98] : "A genetic algorithm approach for multiple criteria facility layout design"

International Journal Of Production Research, 1998, Vol. 36, N°6 pages 1549 à 1569

A. A Islier

[GROPETTI 97] : "On a genetic multiobjective approach for the integration and optimization of an assembly product design and process planning"

Actes IDMME 1997, pages 61 à 70

R. Gropetti, R. Muscia

[SURI 97] : "Performance analysis of flexible Manufacturing systems with single discrete material handling device"

International Journal of Flexible Manufacturing Systems, 1997

R. Suri, R. Desiraju

[PILLET 92] : "Introduction aux plans d'expériences par la méthode Tagushi"

Editions d'Organisation Université, 1992

M. Pillet

[TAGUSHI 87] : "System of Experimental Design"

Traduction anglaise Unipub Kraus International Publication, 1987

G. Tagushi

[LEMOIGNE 93] : "La modélisation des systèmes complexes"

Editions Dunod, AFCET Systèmes, 1993

J. L. Lemoigne

[GERSHWIN 94] : "Manufacturing Systems Engineering"

Editions PTR Prentice Hall, 1994

S. B. Gershwin

[BUZACOTT 93] : "Stochastic Models of Manufacturing Systems"

Editions PTR Prentice Hall, 1993

J. A . Buzacott, J.G. Shantikumar

[ALFARO 97] : "Vers la mise en évidence d'un régime chaotique d'un atelier flexible de production automatisé"

Actes du 2ème congrès franco-qubécois de Génie Industriel, Albi 1997

M. Alfaro, G. Bonhomme, P. Charpentier, P. Martin

[UNIVERSALIS 95] : "Simulation"

Encyclopédia Universalis, version CDROM, 1995

[SAKAROVITCH 84] : "Optimisation combinatoire - Graphes et programmation linéaire"

Editions Hermann, 1984, Volume 1

M. Sakarovitch

[THIEL 90] : "Recherche opérationnelle et management des entreprises"

Editions Economica, collection exercices et cas N°10, 1990

D. Thiel

[YING-CHIN 98] : "Machine Layout with a Linear Single-Row Flow Path in an automated manufacturing system"

Journal of Manufacturing System, Vol. 17 N°1/1998, pages 1 à 22

Ying-Chin Ho, C.L.Moodie

[ARAGON 89] : "Optimization by Simulated Annealing : an experimental evaluation"

Part 1, Graph partitioning, Operations research, Vol. 37, N°6, 1989, pages 865 à 892

C.R. Aragon, D.S. Johnson, L.A.McGeoch, C. Schevon

[GILL 72] : "Quasi-Newton Methods for unconstrained optimization"

Journal of the Institute of Mathematics and its Applications, Vol. 9, 1972, pages 91 à 108

[DAVIS 91] : "Handbook of Genetic Algorithms"

Editions Van Nostrand Reinhold, N.Y., 1991

L. Davis

[BEASLEY98] : "The Hitch-Hiker's Guide to Evolutionary Computation"

On line Ed. <http://www.cs.purdue.edu/coast/archive/clife/FAQ/www/>

D. Beasley, J. Heitkötter

[JANIKOV 91] : "An experimental comparison of binary and floating point representation in Genetic Algorithms"

4th International conference on Genetic Algorithms , 1991, pages 31 à 36

C.Z. Janikov, Z. Michalewicz

[BAKER 85] : "Adaptative selection methods for genetic algorithm"

1st International Conference on Genetic Algorithm, 1985, pages 101 à 111

J.E. Baker

[BERARD 97] : "Résolution de problèmes d'ordonnancement en génie des procédés"

Actes du congrès international Franco-Québécois de Génie Industriel, Albi 1997

F. Berard, C. Azzaro-Pantel, L. Pibouleau, S. Domenech

[PLAQUIN 97] : "An evolutionary manufacturing cell formation method incorporating certain decision maker's requirements"

International Conference on Industrial Engineering and Production Management, 1997, Tome 1, pages 209 à 218

M.F. Plaquin, H. Pierreval

[GAURY 97] : "Evolutionary optimization of a generic model to select pull control strategy"

International Conference on Industrial Engineering and Production Management, 1997, Tome 1, pages 541 à 550

E.G.A Gaury, H. Pierreval

- PARTIE 8 -
Choix et intégration
d'un outil d'optimisation

Sommaire

Introduction	268
I- Comparaison et validation des algorithmes d'optimisation.....	269
□ I.1- Rappel des conditions expérimentales choisies	269
□ I.2- Optimisation par un algorithme d'exploration aléatoire.....	271
□ I.3- Optimisation par un algorithme de descente.....	272
□ I.4- Optimisation par un algorithme de recuit simulé.....	274
□ I.5- Optimisation par un algorithme génétique.....	275
I.5-a) Structure de l'algorithme :	277
I.5-b) Analyse des solutions générées :	286
I.5-b) Améliorations possibles de l'algorithme génétique au vu des résultats :	288
□ I.6- Synthèse et conclusion.....	288
II- Optimisation du délai de production.....	290
□ II.1- Choix des variables et codage du problème	290
□ II.2- Optimisation en horizon fini.....	291
II.2-a) Rappel des résultats obtenus sans algorithme d'optimisation :	291
II.2-b) Résultats obtenus par l'algorithme d'optimisation :	291
□ II.3- Optimisation en horizon infini.....	294
II.3-a) Rappel des résultats obtenus sans algorithme d'optimisation :	294
II.3-b) Résultats obtenus par l'algorithme d'optimisation :	294
III- Optimisation multicritère	296
□ III.1- Choix des critères	297
□ III.2- Problématique de l'optimisation multicritère	298
□ III.3- Développement expérimental d'une méthode d'optimisation multicritère	300
III.3-a) Méthode 1 : Optimisation par Sommation (O.S)	300
III.3-b) Méthode 2 : Optimisation par Normage, Pondération et Sommation (N.P.S).....	304
III.3-c) Méthode 3 : Optimisation par Non Domination Pondérée (N.D.P).....	313
III.3-d) Conclusions sur les méthodes de sélection testées.....	323
IV- Proposition d'intégration des outils informatiques en gestion de production	323
□ IV.1- Analyse fonctionnelle des principaux outils informatiques	323
IV.1-a) Gestion de Production Assistée par Ordinateur (G.P.A.O) :	323
IV.1-b) Ordonnancement.....	324
IV.1-c) Simulation de flux.....	325
IV.1-d) Optimisation.....	325
IV.1-e) Supervision.....	326
IV.1-f) Analyse Statistique des Phénomènes Aléatoires (A.S.P.A)	327
IV.1-g) Suivi de production	327

□ IV.2- Modèle d'intégration des outils	328
Conclusion	330
Bibliographie	332

Introduction

Dans cette partie, nous allons développer le principe du couplage entre un modèle de simulation de flux et un algorithme d'optimisation. Nous exploiterons comme support de ce travail le modèle de simulation de la ligne expérimentale déjà utilisé dans la partie 6. Nous utiliserons aussi, pour valider les algorithmes proposés et avant de les appliquer sur des cas plus complexes, les résultats expérimentaux de la partie 6. L'objectif est de mettre en œuvre un algorithme d'optimisation couplé à la simulation pour traiter des cas de grande complexité. Plusieurs algorithmes sont possibles, du plus simple (exploration aléatoire de l'espace des solutions) au plus "intelligent" (recuit simulé ou algorithme génétique).

Dans le chapitre 1, une comparaison de quatre algorithmes (exploration aléatoire, descente stochastique, recuit simulé et algorithme génétique) est réalisée pour l'optimisation du délai pour le jeu d'essai N°1 (production monoproduit). Pour des raisons déjà développées dans la partie 7, nous retiendrons finalement l'utilisation d'un algorithme génétique. La structure de l'algorithme choisi est présentée, en mettant en évidence ses particularités : sélection élitiste, codage réel des paramètres, prise en compte de la consanguinité et utilisation d'un individu mutant à chaque génération.

Dans le chapitre 2, l'algorithme génétique et la simulation sont ensuite utilisés pour optimiser le délai de production du jeu N°2 (production multiproduit).

Le chapitre 3 présente un prolongement possible de cette méthodologie d'optimisation en faisant intervenir plusieurs critères (chacun à minimiser) : le délai, l'encours moyen, et les réglages. La principale difficulté en optimisation multicritère réside dans la sélection des solutions satisfaisant au mieux chaque critère. Plusieurs méthodes de sélection associées à l'algorithme génétique de base sont testées expérimentalement.

Pour conclure cette partie, et en s'inspirant de la méthodologie suivie pour ce travail, le chapitre 4 propose de mettre en place un modèle d'intégration des différents outils informatiques à la disposition du gestionnaire de production. L'objectif de ce modèle est à la fois de mettre en évidence les liens à établir entre chaque outil et aussi de structurer leur utilisation.

I- Comparaison et validation des algorithmes d'optimisation

Dans ce chapitre, nous allons commencer par tester quatre algorithmes d'optimisation :

- Un algorithme d'exploration aléatoire, basé, comme son nom l'indique, sur une recherche aléatoire dans l'espace des solutions;
- Un algorithme de descente qui s'arrête sur un optimum local;
- Un algorithme de recuit simulé qui fait une recherche globale en acceptant une dégradation de solutions locales;
- Un algorithme génétique que nous présenterons plus en détail pour montrer la contribution apportée.

L'objectif est d'étudier le comportement de chacun de ces algorithmes pour résoudre le même problème, afin de justifier le choix de l'un d'entre eux pour la suite du travail. Ils seront utilisés pour optimiser le délai de fabrication qui correspond au jeu N°1 (production monoproduit), à l'expérience N°1 (horizon fini), en fixant la capacité du stock amont (StAm) à 1. Nous agirons donc uniquement sur la variable (TiLa), mais contrairement à ce qui a été fait dans la partie 6, nous allons considérer que (TiLa) peut varier d'une palette à l'autre. Nous avons donc plusieurs variables homogènes ($TiLa_{(p,p+1)}$), avec $p = (\text{nombre de palettes} - 1)$ (voir partie 7, chapitre II.2). Pour le cas expérimental choisi, qui comporte 25 palettes, nous avons donc 24 variables ($TiLa_{(1,2)}$) à ($TiLa_{(24,25)}$), qui peuvent chacune prendre une valeur discrète comprise entre 1 et 20. Il y a donc $[(20 - 1) + 1]^{24}$ combinaisons possibles.

Nous avons démontré dans la partie 6 que le délai minimum théorique obtenu aussi en pratique, était de 577 secondes (voir partie 6, chapitre I.2-d), pour un (TiLa) constant entre chaque palette de 16 secondes.

En augmentant considérablement l'espace des solutions, l'idée est donc de "tromper" l'algorithme utilisé pour voir s'il est capable de trouver UNE bonne combinaison, si ce n'est LA bonne combinaison.

Chaque algorithme sera testé sur environ 140 itérations, ce qui correspond à peu près au nombre de simulations réalisées à chaque expérimentation de la partie 6.

□ I.1- Rappel des conditions expérimentales choisies

- Lancement de type "MONOPRODUIT".
- Nombre de palettes en réserve : $R_p = 25$, égal à la quantité à lancer (N_p).
- Séquencement des opérations dans un ordre imposé, sans redondance avec :
 $Pi[j]$ = opération sur le poste i ($i=1\dots 6$) pendant un temps de cycle de j secondes.

Les temps de cycle sont fixés arbitrairement aux valeurs suivantes :

P2[7] – P3[12] – P5[8] – P6[3]

- Horizon de fabrication FINI : la ligne est vide en fin de simulation.

Les figures 1 et 2 rappellent le routage des palettes et l'allure de la fonction objectif qui nous sert de test, dans le cas où (TiLa) est constant entre chaque palette.

Une représentation graphique d'une fonction de 24 variables étant impossible, on peut quand même imaginer que celle-ci sera au moins aussi "bruitée" que la fonction à une variable.

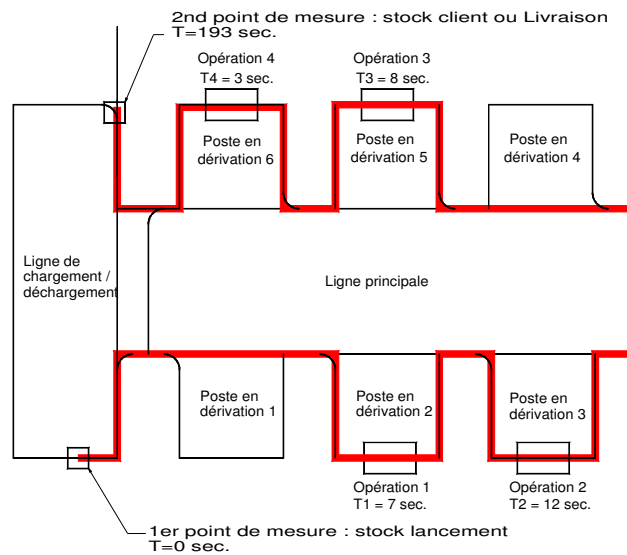


Figure 1 : Routage des palettes

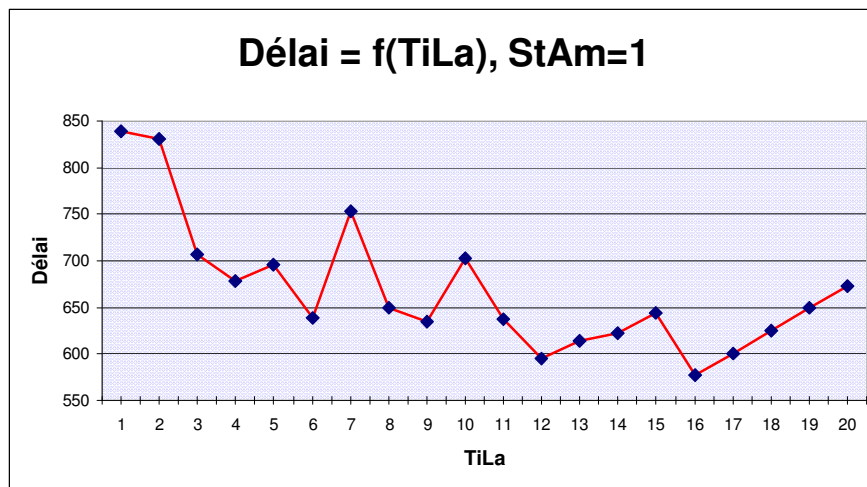


Figure 2 : Allure de la fonction objectif (cas où TiLa est constant)

□ I.2- Optimisation par un algorithme d'exploration aléatoire

Le graphique de la figure 3 présente le résultat de cet algorithme. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont présentées sur la figure 4.

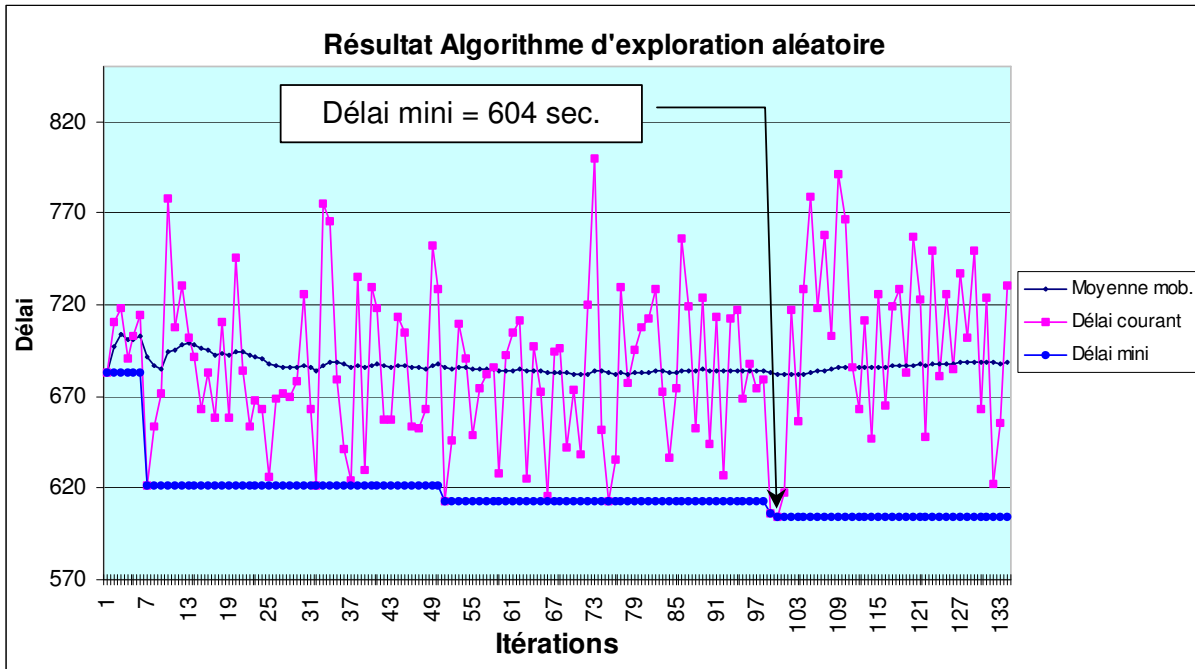


Figure 3 : Délai =f(itération) pour algorithme d'exploration aléatoire

Iter.	Mini	Moy	Délai	TiLa 1 à TiLa 24																							
1	683.4	683	683.4	1	5	14	17	13	5	5	7	13	6	10	12	19	9	13	6	6	5	7	19	6	20	1	9
7	621.2	691	621.2	10	9	18	12	6	4	5	14	12	8	4	17	7	13	12	3	14	10	4	1	17	11	6	9
51	612.6	686	612.6	6	15	20	18	11	14	13	10	5	3	9	7	9	20	3	3	10	18	3	9	16	11	8	15
99	606.5	683	606.5	12	1	10	7	17	6	2	2	15	19	1	6	1	6	13	13	8	18	3	2	6	8	4	14
100	604.1	682	604.1	6	19	18	16	5	3	12	13	3	17	14	14	5	13	15	18	5	19	15	11	6	1	12	15

Figure 4 : Meilleurs résultats aux paliers

La meilleure solution qui donne le délai minimal (D = 604 sec.), obtenue à la 100^{ème} itération est :

6-19-18-16-5-3-12-13-3-17-14-14-5-13-15-18-5-19-15-11-6-1-12-15

Comparativement au délai minimal théorique de 577 secondes, l'écart est de :

$$(604 - 577) = 27 \text{ secondes.}$$

Même si ce résultat semble correct pour une exploration aléatoire, on voit très bien sur le tracé du délai courant (figure 3) qu'il est le fruit du pur hasard. L'allure du tracé du délai ainsi que de la moyenne mobile (*) montrent aussi la dispersion importante non seulement entre les valeurs extrêmes du délai (de 604 sec. à 800 sec.), mais aussi entre chaque solution, preuve de la complexité du problème.

□ I.3- Optimisation par un algorithme de descente

Le graphique de la figure 5 présente le résultat de cet algorithme. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont présentées sur la figure 6.

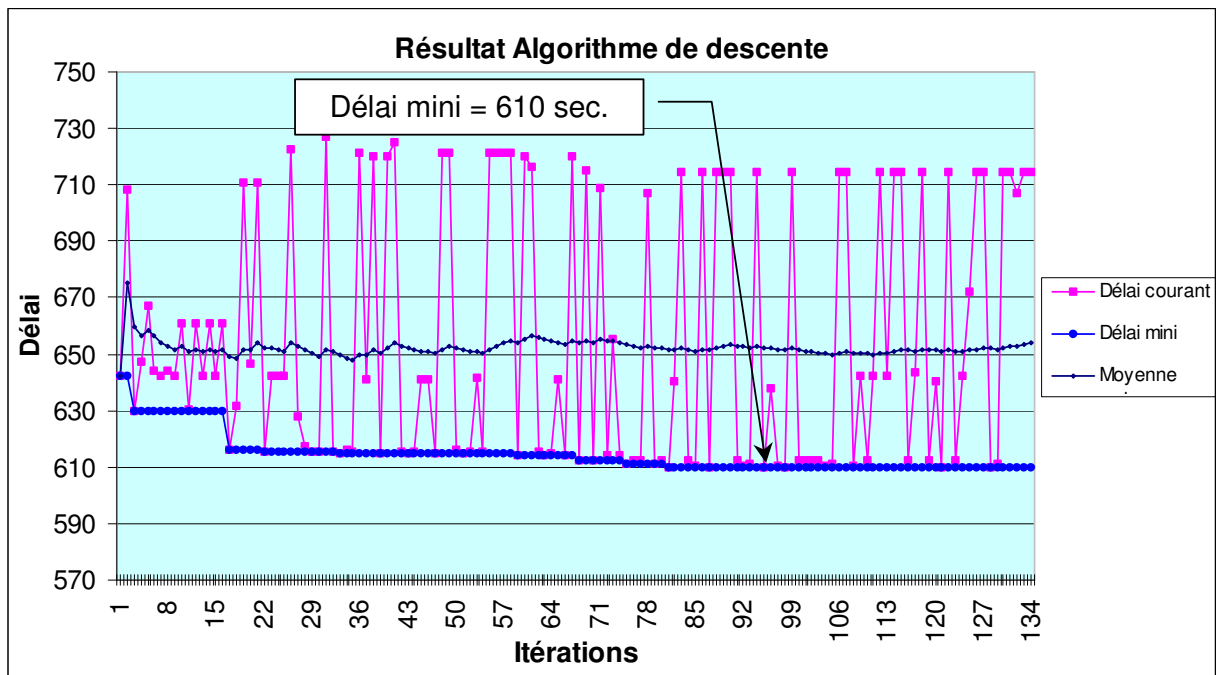


Figure 5 : Délai = f(itération) pour algorithme de descente

(*) La moyenne mobile est la moyenne calculée depuis la première itération jusqu'à l'itération courante.

Iter.	Mini	Moy	Délai	TiLa 1 à TiLa 24																							
				10	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
1	642	642	642	10	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
3	629.5	659.9	629.5	10	4	5	2	13	15	19	17	10	18	18	11	12	20	7	8	13	5	14	7	8	20	19	14
17	615.9	649.4	615.9	10	4	5	2	13	15	19	17	10	18	18	11	12	20	7	8	12	5	14	7	8	20	19	14
30	615.3	649.4	615.3	10	4	5	2	13	14	19	17	10	18	18	11	12	20	7	8	12	4	14	7	8	19	19	14
59	613.9	654.3	613.9	10	4	5	2	13	14	18	17	10	18	18	11	12	20	7	8	12	4	14	6	8	20	19	14
68	612.4	654	612.4	10	4	5	2	13	13	18	17	10	18	18	11	12	20	7	8	12	4	14	6	8	20	19	14
75	610.8	653.3	610.8	10	4	5	2	13	13	18	17	10	17	18	11	12	20	7	8	12	4	14	7	8	20	19	14
81	610	651.4	610	10	4	5	2	13	13	18	17	10	17	18	11	12	20	7	8	12	4	14	7	7	20	19	14

Figure 6 : Meilleurs résultats aux paliers

La meilleure solution qui donne le délai minimal ($D = 610$ sec.), obtenue à la 81^{ème} itération est :

10-4-5-2-13-13-18-17-18-11-12-20-7-8-12-4-14-7-7-20-19-14

Comparativement au délai minimal théorique de 577 secondes, l'écart est de :

$$(610 - 577) = 33 \text{ secondes.}$$

Le tableau de la figure 6 montre bien le principe du "voisinage" utilisé par cet algorithme. La transformation élémentaire d'une solution consiste à rajouter aléatoirement +/- sur certaines variables tirées au hasard. Si cette transformation produit une amélioration de la fonction objectif, la solution obtenue est utilisée comme solution de départ au processus de transformation.

Le résultat final de la fonction objectif est moins bon que pour l'algorithme d'exploration aléatoire, ce qui peut à la fois être surprenant et décevant. Le choix de la solution de départ est primordial dans ce type d'algorithme. Pour ce test, la solution de départ est choisie au hasard. On peut espérer, en combinant l'algorithme d'exploration aléatoire et l'algorithme de descente, améliorer la performance de celui-ci. Un algorithme "hybride", qui commencerait par une exploration aléatoire puis, à partir de la meilleure solution, tenterait de "creuser" un peu plus profond pour améliorer cette solution, semble être une piste intéressante.

□ I.4- Optimisation par un algorithme de recuit simulé

Le graphique de la figure 7 présente le résultat de cet algorithme. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont présentées sur la figure 8.

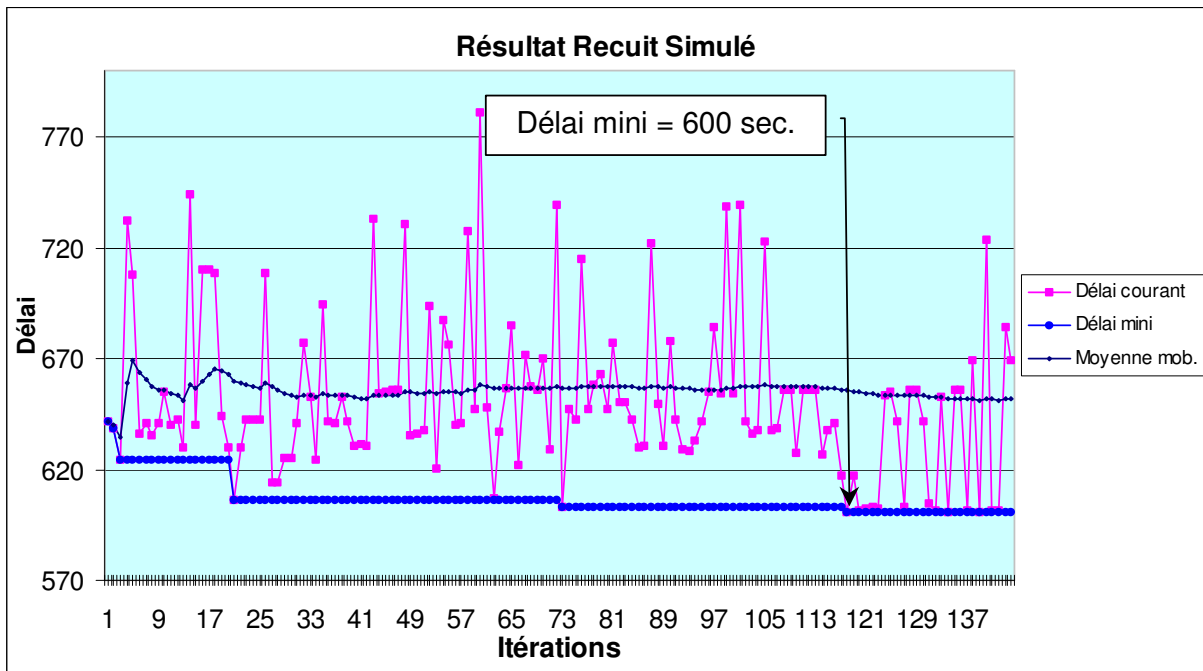


Figure 7 : Délai = f(itération) pour algorithme du recuit simulé

Iter	Mni	Mby	Délai	TiLa 1 à TiLa 24																							
				10	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
1	642	642	642	10	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
2	639	640	638.5	20	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
3	624	635	624.3	10	14	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	7	8	20	19	14
21	606	660	605.9	10	4	5	2	13	15	19	17	10	18	18	10	12	20	7	8	13	5	14	17	8	20	19	14
73	603	657	603.3	9	3	5	4	4	15	19	7	9	18	16	9	13	9	7	8	3	16	14	8	8	20	20	14
118	600	656	600.3	10	2	15	5	14	15	19	8	9	17	6	19	12	9	9	19	2	17	15	7	6	10	19	13

Figure 8 : Meilleurs résultats aux paliers

La meilleure solution qui donne le délai minimal ($D = 600$ sec.), obtenue à la 118^{ème} itération est :

10-2-15-5-14-15-19-8-9-17-6-19-12-9-9-19-2-17-15-7-6-10-19-13

Comparativement au délai minimal théorique de 577 secondes, l'écart est de :

$$(600 - 577) = 23 \text{ secondes.}$$

Les paramètres de cet algorithme ont été fixés aux valeurs suivantes :

- Température initiale = 100°
- Taux de refroidissement = 0,91
- Nombre maximal d'acceptations de solutions dégradées = 25

Pour l'instant, c'est cet algorithme qui donne le meilleur résultat, même si la différence n'est pas flagrante. Il faut rappeler que le principe de cet algorithme est de pouvoir sortir d'un minimum local, et d'explorer d'autres "paysages", ce qui le rend plus "gourmand" en itérations qu'un algorithme de descente. Pour qu'il soit plus efficace, on peut penser qu'il faudrait le laisser travailler sur un nombre d'itérations plus élevé.

On voit sur le tableau de la figure 8 que l'évolution des bonnes solutions est parfois marquée par des variations assez importantes de certaines variables. Ceci prouve que ces solutions ont été obtenues à partir d'une solution dégradée. Comme pour l'algorithme de descente, la solution de départ est choisie au hasard et conditionne peut-être l'obtention de solutions meilleures.

□ 1.5- Optimisation par un algorithme génétique

Le graphique de la figure 9 présente le résultat de cet algorithme. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont présentées sur la figure 10.

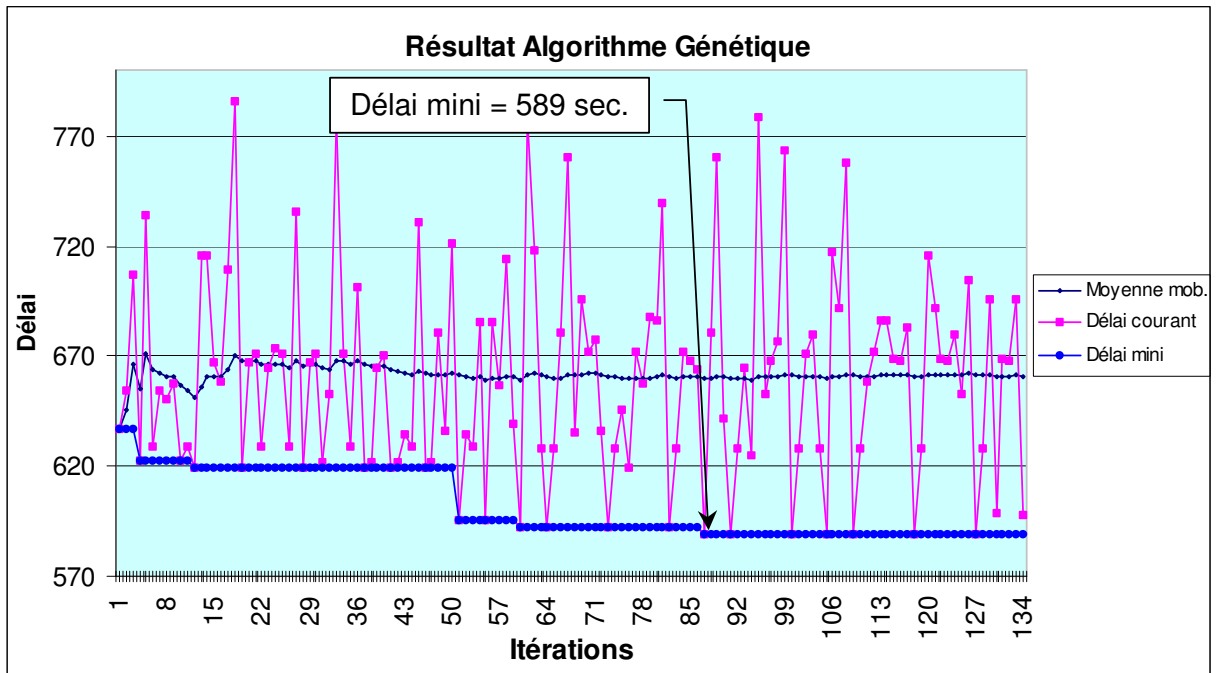


Figure 9 : Délai =f(itération) pour algorithme génétique

Iter	Mni	Moy	Délai	TiLa 1 à TiLa 24																							
				2	11	18	9	10	13	17	16	11	8	6	6	8	10	16	4	6	12	9	9	3	10	10	18
1	636.6	636.6	636.6	2	11	18	9	10	13	17	16	11	8	6	6	8	10	16	4	6	12	9	9	3	10	10	18
4	622.2	655	622.2	15	2	1	10	11	4	9	7	10	15	14	12	2	15	10	2	16	13	3	10	20	8	7	18
12	619.1	651.3	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
51	595.7	661.2	595.7	7	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
60	591.9	659.2	591.9	5	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
87	589.1	659.7	589.1	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4

Figure 10 : Meilleurs résultats aux paliers

La meilleure solution qui donne le délai minimal (D = 589 sec.), obtenue à la 87^{ème} itération est :

3-6-13-12-16-6-9-7-10-15-14-12-19-15-14-4-1-16-17-15-9-5-20-4

Comparativement au délai minimal théorique de 577 secondes, l'écart est de :

$$(589 - 577) = 12 \text{ secondes.}$$

Sur les quatre algorithmes testés, c'est donc celui-ci qui donne le meilleur résultat. Il faut encore préciser, mais cela va être détaillé par la suite, que ces résultats sont fonction de paramètres intrinsèques à chaque algorithme, à l'exception de l'exploration aléatoire. Pour l'algorithme de la descente, il y a de nombreuses possibilités de transformations élémentaires permettant de définir le voisinage. Le choix de l'une d'elles peut influencer sur le fonctionnement de l'algorithme et sur les résultats obtenus. Pour l'algorithme du recuit simulé, il faut choisir non seulement une transformation élémentaire, mais aussi la température initiale et le taux de refroidissement.

L'objectif de ces tests n'est donc pas de généraliser ou de conclure hâtivement sur la performance de tel ou tel algorithme. Un choix différent des paramètres intrinsèques aurait pu modifier le résultat. On cherche seulement à analyser leur comportement sur un problème donné, à proposer des possibilités d'amélioration et à disposer d'éléments permettant de justifier l'utilisation de l'un d'eux.

En ce qui concerne l'algorithme génétique utilisé, nous allons voir en détail sa structure et son fonctionnement, car il est issu d'un développement propre à ce travail qui tient compte de certaines remarques déjà faites sur les autres algorithmes.

1.5-a) Structure de l'algorithme :

L'algorithme utilisé peu être qualifié d'hybride. En effet, il est constitué à la fois d'une recherche purement aléatoire et d'une recherche pseudo-aléatoire comportant une structure d'algorithme génétique. La partie génétique ne correspond pas exactement à celle proposée par [GOLDBERG 94], notamment pour la sélection des individus choisis pour le croisement. Dans ce travail, il a été décidé d'opter pour une sélection de type ELITISTE, qui consiste à ne garder que les deux meilleurs individus d'une génération. Le principe de la "roue de la fortune" de Goldberg n'est donc pas retenu. D'autre part, comme le recommande [RENDERS 97], les gènes sont codés en réel, plus précisément en entier, et non pas en binaire.

Des développements d'algorithmes génétiques, comportant plusieurs variantes de sélection, de croisement et de mutation, sont proposés par [VINCENT 98] et appliqués sur les modèles de simulation de la ligne expérimentale pour optimiser le délai de fabrication d'un lot de palettes. La structure du code est récursive, ce qui lui confère une plus grande compacité et surtout améliore la gestion de la mémoire. On trouve dans [FONTANILI 97], [FONTANILI 98a] et [FONTANILI 98b] les résultats obtenus pour optimiser le délai de fabrication à partir de ces algorithmes.

L'algorithme est codé en Pascal sous l'environnement de développement DELPHI 3™ pour WINDOWS 95™. Seules les parties de code standard au Pascal sont présentées, indépendamment du code nécessaire à l'interface graphique.

L'algorithme manipule des gènes qui correspondent dans le cas étudié aux différentes variables de pilotage, donc aux vingt-quatre variables TiLa (TiLa_1 à TiLa_24).

La figure 11 montre les constantes utilisées : le nombre de gènes donc de variables (homogènes dans ce cas), le nombre de générations (fixé à 15 pour obtenir un total d'environ 140 itérations) et le nombre d'individus par génération (voir explications plus loin).

Pour revenir sur le nombre d'individus par génération, celui-ci comporte :

- deux parents qui sont les deux meilleurs individus sélectionnés dans la génération précédente (choix élitiste),
- n paires d'enfants issus du croisement des gènes des deux. Pour ce test, on a choisi $n = 3$ (nombre choisi arbitrairement),
- un individu dont les gènes sont tous générés au hasard, appelé "Mutant".

(*Nombre. de variables = Nombre. de Gènes, par exemple TiLa_(1) à TiLa_(24) *)
Nbre_genes=24;

(* Nombre maxi de générations*)
Nbre_generation=15;

(* Obligatoirement Nombre IMPAIR d'individus/génération*)
(* = 2 premiers individus = clones de la génération précédente *)
(* + 2 x n enfants issus des 2 parents clonés *)
(* + 1 individu Mutant *)
Nbre_individus=9;

Figure 11 : Constantes utilisées dans l'algorithme génétique

Le programme doit commencer par générer aléatoirement une population composée d'individus dont les gènes sont pris au hasard. C'est cette première génération qui va servir de point de départ à l'algorithme génétique. Contrairement aux autres algorithmes, celui-ci démarre avec plusieurs individus (9 dans le cas étudié) tirés aléatoirement, ce qui lui permet de choisir les deux meilleures solutions pour démarrer. La figure 12 montre la procédure permettant de générer cette population au hasard.


```

(* procédure permettant de générer aléatoirement les individus de la 1ère génération *)

begin
(* pour chaque individu d'une génération*)
  for i:=1 to nbre_individus do

(* un individu correspond à un chromosome composé de nbre_genes*)
(* dans le cas de la ligne, chaque gène est un Tila(i), chaque chromosome est donc un
(* ensemble de Tila(i) *)
  begin

(* pour chaque gène d'un chromosome (donc d'un individu i) *)
    for j:=1 to (nbre_genes) do

      begin
(* génération aléatoire de la valeur d'un gène j comprise entre 1 et 20 sec. pour Tila_(i)
(* exemple : gene[i][1] = Tila_(1), gene[i][2] = Tila_(2) etc... *)
        gene[i][j]:=random(20)+1;

      end;

    end;

  (* on lance la simulation pour avoir la réponse (la valeur de la fonction objectif*)
  (* on récupère la valeur de la fonction objectif*)
  reponse(i);

  end;
end;

```

Figure 12 : Génération Aléatoire (Génération initiale)

La procédure REponse(i) n'est pas détaillée ici, mais il s'agit simplement d'une procédure qui permet de lancer le modèle de simulation avec comme variables d'entrée celles générées aléatoirement dans la première génération et de récupérer la valeur de la fonction objectif, c'est-à-dire le délai de fabrication. On peut cependant préciser que cet échange de données entre l'outil d'optimisation et l'outil de simulation est obtenu en utilisant les liens d'échanges dynamiques O.L.E entre applications WINDOWS 95™ (voir figure 13).

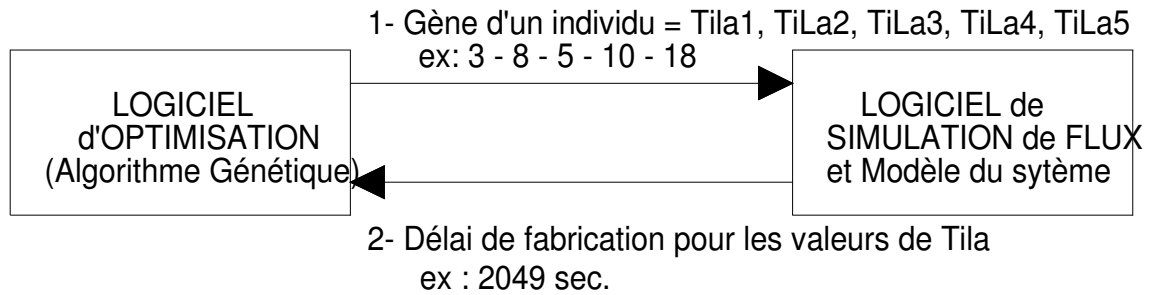


Figure 13 : Utilisation des liens O.L.E pour le couplage Simulation/Optimisation

La procédure SELECTION qui utilise les fonctions FIRST et SECOND permet de sélectionner uniquement les deux meilleurs individus d'une population, c'est-à-dire les deux individus qui ont la fonction objectif la plus faible. Une fois que ces individus ont été identifiés, on initialise leur indice à 1 et 2 pour qu'ils deviennent les parents de la génération suivante (les parents ont toujours les deux premiers indices d'une population). La figure 14 montre cette procédure de sélection, basée sur le principe de l'élitisme. C'est à ce niveau que l'algorithme utilisé présente une différence importante avec celui proposé par Goldberg.

```
(* Procédure permettant d'identifier les indices des 2 meilleurs individus*)

Begin

(*Utilisation de fonctions qui renvoient les indices des 2 meilleurs individus*)
Gold:=First;
Silver:=Second;

(* on fait une image des gènes des 2 meilleurs individus *)

image_1:=gene[gold];
image_2:=gene[silver];

(* Les 2 meilleurs individus de la génération précédente deviennent les PARENTS *)
Gene[1]:=image_1;
Gene[2]:=image_2;

end;
```

Figure 14 : Sélection des 2 meilleurs individus

Il est important d'expliquer en détail la structure des fonctions qui permettent de trouver les 2 meilleurs individus. La fonction de la figure 15 montre la sélection du meilleur individu.

```
(* fonction renvoyant l'indice du meilleur individu d'une génération*)
begin
(* initialisation des variables*)
    mini:=10000;
    indice_mini:=1;

(*pour chaque individu d'une génération*)
    for i:=1 to nbre_individus do
        begin

(* si la valeur de la fonction objectif est inférieure à la valeur courante*)
(* Attention : Pour qu'en cas d'égalité de "force" entre 1 parent et 1 enfant, ce soit
(*l'enfant qui soit conservé dans la génération suivante : placer <=
(* Si on veut garder le parent : placer < *)
            if fct_obj[i]<=mini then
                begin
(*on place dans la variable mini la valeur de la fonction obj. *)
                    mini:=fct_obj[i];

(* on mémorise l'indice de l'individu qui a la valeur minimale de la fonction objectif *)
                    indice_mini:=i;
                end;
            end;

(*la fonction renvoie l'indice du meilleur individu d'une génération*)
            first:=indice_mini;

        end;

```

Figure 15 : Sélection du 1^{er} parent

La fonction qui détermine le second parent fait appel à la notion de **CONSANGUINITE** des gènes : pour sélectionner le second individu, on détermine le nombre de gènes qu'il possède à l'identique avec le premier parent. Au delà d'une valeur de consanguinité fixée arbitrairement dans le programme (12 pour le cas étudié), un individu, même s'il correspond à la seconde meilleure fonction objectif, ne sera pas sélectionné.

Dans certains cas extrêmes, l'algorithme peut être amené à retenir comme second parent l'individu qui a la plus mauvaise fonction objectif.

Cette partie de l'algorithme est une garantie de ne pas obtenir, au bout de quelques générations, des individus absolument identiques. La figure 16 présente la fonction qui sélectionne le second parent.

```
(*fonction renvoyant l'indice du 2nd meilleur individu d'une génération*)
begin
  mini:=10000;
  indice_mini:=1;
  for i:=1 to nbre_individus do
    begin
      consanguin:=0;
      (* si la valeur de la fonction objectif est inférieure à la valeur courante et que l'indice est
      (* différent de celui du meilleur individu et que le chromosome est plus ou moins proche
      (* consanguinité)*)

      for Kb:=1 to Nbre_genes do
        begin
          if gene[i][kb]=gene[first][kb] then
            consanguin:=consanguin+1;
          (* plus consanguin a une valeur élevée, plus les 2 chromosomes sont proches *)
        end;

      (* Pour qu'en cas d'égalité de "force" entre 1 parent et 1 enfant, ce soit l'enfant qui soit
      (* conservé dans la génération suivante : placer <=
      (* Si on veut garder le parent : placer <*)
      if ((fct_obj[i]<=mini) and (i<>first)) and (consanguin<=12)
      then
        begin
          mini:=fct_obj[i];
          indice_mini:=i;
        end;

      end;

      (* la fonction renvoie l'indice du 2nd meilleur individu d'une génération*)
      second:=indice_mini;

    end;
end;
```

Figure 16 : Sélection du 2nd parent avec consanguinité

Pour le croisement entre les parents, le hasard intervient puisqu'il faut choisir aléatoirement un site de croisement, qui est en fait un nombre entier compris entre 1 et 23, puisqu'il y a 24 gènes. Cet algorithme comporte un croisement simple (à un seul site), qui va servir à créer n paires d'enfants (n choisi arbitrairement, ici fixé à 3) dont le chromosome comporte une partie des gènes provenant du parent 1 et l'autre partie provenant du parent 2. La procédure de croisement est présentée sur la figure 17.

```
(* Processus de CROISEMENT des 2 meilleurs individus sélectionnés *)
  Begin
    site_precedant:=0;

    (* le croisement permet d'obtenir des enfants à partir de l'individu 3*)
    i:=3;
    Repeat
      (* pour chaque individu i d'une génération*)
      repeat
        (* choix aléatoire d'un site de croisement = position entre 2 gènes*)
        (* ce site doit être différent de celui du croisement précédent*)
        site:=random(nbre_genes-1)+1;
        until site<>site_precedant;

        site_precedant:=site;

      (* pour chaque gène jusqu'au site de croisement*)
      For j:=1 to Site do
        begin
          gene[i][j]:=gene[1][j];
          gene[i+1][j]:=gene[2][j];
        end;

      (* A partir du gène situé après le site de croisement jusqu'au dernier gène*)
      For j:=(site+1) to nbre_genes do
        begin
          gene[i][j]:=gene[2][j];
          gene[i+1][j]:=gene[1][j];
        end;

      i:=i+2;

    until i=nbre_individus;

  end;
```

Figure 17 : Croisement des deux parents et génération des enfants

Afin de ne pas enfermer la population dans un minimum local, celle-ci comporte systématiquement un individu dont les gènes ne sont pas issus du croisement des gènes des parents, mais uniquement du hasard. La procédure de création de cet individu est similaire à la procédure qui a permis de générer les individus de la première population. Grâce à cet individu, à chaque génération l'algorithme fait une exploration aléatoire de l'espace des solutions.

On utilise aussi cette idée de mutation, mais localement, sur un seul des gènes d'un seul des enfants issus du croisement des deux parents. Ceci permet de modifier très légèrement (en ajoutant ou en enlevant 1) un gène d'un individu pris au hasard parmi la nouvelle population. Cette mutation est réalisée en fonction d'une probabilité qui suit une loi de distribution uniforme entre 0 et 1. On se fixe un taux de mutation compris entre 0 et 1 (0,6 dans le programme) et on réalise à chaque génération un tirage aléatoire dont le résultat est compris entre 0 et 1. Si ce résultat est inférieur ou égal au taux de mutation, on prend un individu au hasard, soit parmi les enfants uniquement, soit en prenant en compte aussi le second parent, et on modifie un de ses gènes au hasard. La figure 18 présente cette procédure.

```
(* Procédure de mutation aléatoire d'un des gènes *)
  Begin
  (*Tire un nombre aléatoire compris entre 0 et 1*)
    Jet:=random(1);

    if jet<=taux_mute then
      begin
      (* On tire au hasard l'individu qui va subir une mutation d'un de ses gènes*)
      (* Mutation possible du 2nd parent : individu_mute : random(nbre_individus-1)+2*)
      (*Pas de mutation possible du 2nd parent individu_mute : random(nbre_individus-2)+3*)
        individu_mute:=random(nbre_individus-2)+3;

      (*On tire au hasard le gène qui va subir une mutation*)
        gene_mute:=random(nbre_genes)+1;

      (*La mutation est tirée au hasard entre 0 et 1*)
        mutation:=random(2);

      (* Si la valeur est 0, alors on retire 1 à la valeur du gène*)
        if mutation=0 then mutation:=-1;

      (* Sinon on ajoute 1*)
        gene[individu_mute][gene_mute]:=gene[individu_mute][gene_mute]+mutation;

      end;
    end;
```

Figure 18 : Mutation aléatoire d'un gène d'un individu

La figure 19 montre une partie du corps du programme principal de l'algorithme utilisé. Pour chaque génération, on retrouve l'appel aux procédures **SELECTION**, **CROISEMENT**, **MUTANT** et **MUTATION**. A chaque individu, l'algorithme envoie le chromosome, c'est-à-dire dans ce cas les valeurs des 24 variables $TiLa_{(i)}$, à la simulation et celle-ci renvoie la valeur de la fonction objectif, c'est-à-dire le délai de fabrication.

```

(* Corps du programme principal de la recherche génétique/hybride*)
  begin

(* on commence par créer les individus de la 1ère génération au hasard*)

      init_random;

(* pour chaque génération *)
      for i:=1 to nbre_generation do
        begin

(* Sélection des 2 meilleurs individus de la génération *)
          SELECTION;

(* Croisement des 2 meilleurs individus *)
          CROISEMENT;

(* Création d'un individu MUTANT (gènes aléatoires)*)
          MUTANT;

(* Mutation aléatoire d'un des gènes d'un des individus sauf mutant *)
          MUTATION;

(* Calcul de la fonction objectif, renvoyée par la simulation *)
          For j:=1 to nbre_individus do
            begin
              reponse(j);
            end;
          end;
        end;
      end;
  end;

```

Figure 19 : Programme principal

1.5-b) Analyse des solutions générées :

Afin de vérifier et de comprendre le fonctionnement de cet algorithme génétique, nous présentons les résultats des cinq premières générations (figure 20).

Iter	Mini	Moy	Délai	TiLa 1 à TiLa 24																							
				2	11	18	9	10	13	17	16	11	8	6	6	8	10	16	4	6	12	9	9	3	10	18	
1	636.6	636.6	636.6	2	11	18	9	10	13	17	16	11	8	6	6	8	10	16	4	6	12	9	9	3	10	18	
2	636.6	645.3	654	14	15	17	20	4	18	10	7	3	15	8	11	14	15	9	8	3	5	2	15	17	7	9	1
3	636.6	665.9	707.2	12	6	11	13	9	14	1	15	15	11	11	14	3	19	14	16	15	15	3	13	7	3	12	7
4	622.2	655	622.2	15	2	1	10	11	4	9	7	10	15	14	12	2	15	10	2	16	13	3	10	20	8	7	18
5	622.2	670.8	734.2	18	11	20	16	15	2	1	5	4	17	11	2	2	7	1	1	16	17	16	20	1	1	10	13
6	622.2	663.8	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
7	622.2	662.5	654.7	6	11	19	11	19	18	13	14	8	9	20	11	17	1	20	3	8	8	5	14	15	19	12	4
8	622.2	661	650.2	8	19	20	19	17	15	18	2	16	19	14	17	5	8	16	5	9	12	1	9	13	8	9	9
9	622.2	660.7	657.9	17	1	3	4	14	11	16	18	14	16	3	15	13	18	17	5	5	1	20	9	4	8	18	2
10	622.2	656.8	622.2	15	2	1	10	11	4	9	7	10	15	14	12	2	15	10	2	16	13	3	10	20	8	7	18
11	622.2	654.3	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
12	619.1	651.3	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
13	619.1	656.3	715.6	7	6	13	12	16	6	7	7	16	7	3	16	2	15	10	2	16	13	3	10	20	8	7	18
14	619.1	660.5	715.7	15	2	1	10	11	4	9	7	10	15	14	12	2	15	10	2	16	13	3	15	9	5	20	4
15	619.1	661	667.2	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	10	20	8	7	18
16	619.1	660.8	658.1	15	2	1	10	11	4	9	7	10	15	14	16	19	15	14	4	1	16	17	15	9	5	20	4
17	619.1	663.6	709.1	7	6	13	12	16	6	7	7	16	7	3	12	2	15	10	2	16	13	3	10	20	8	7	18
18	619.1	670.4	785.5	14	3	9	7	11	4	13	8	3	3	2	13	13	1	10	2	12	12	6	11	7	8	3	2
19	619.1	667.7	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
20	619.1	667.7	667.2	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	10	20	8	7	18
21	619.1	667.8	670.8	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	10	20	8	7	18
22	619.1	666	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
23	619.1	666	664.4	15	2	1	10	11	6	7	7	16	7	3	16	19	15	14	4	2	16	17	10	20	8	7	18
24	619.1	666.3	673.1	7	6	13	12	16	4	9	7	10	15	14	12	19	16	14	4	2	16	17	15	9	5	20	4
25	619.1	666.5	670.8	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	10	20	8	7	18
26	619.1	665	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
27	619.1	667.6	735.9	20	6	8	11	20	20	13	14	13	5	4	19	12	8	18	13	15	8	14	14	9	6	4	9
28	619.1	665.9	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
29	619.1	665.9	667.2	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	10	20	8	7	18
30	619.1	666.1	670.8	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	10	20	8	7	18
31	619.1	664.7	621.5	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	19	4
32	619.1	664.3	652.7	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	7	18
33	619.1	667.7	777.2	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	10	20	8	20	4
34	619.1	667.8	670.8	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	10	20	8	7	18
35	619.1	666.7	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
36	619.1	667.7	701.4	13	20	12	15	8	8	12	6	15	3	6	15	9	4	19	9	12	15	12	5	4	12	3	10
37	619.1	666.3	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
38	619.1	665.2	621.5	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	19	4
39	619.1	665.1	664.4	15	2	1	10	11	4	9	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	19	4
40	619.1	665.3	670.5	7	6	13	12	16	6	7	7	10	15	14	12	19	15	14	4	2	16	17	15	9	4	20	4
41	619.1	664.2	619.1	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	20	4
42	619.1	663.1	621.5	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	19	4
43	619.1	662.5	634.6	15	2	1	10	11	4	9	7	10	15	14	12	19	15	14	4	2	16	17	15	9	5	19	4
44	619.1	661.7	628.8	7	6	13	12	16	6	7	7	16	7	3	16	19	15	14	4	2	16	17	15	9	5	20	4
45	619.1	663.2	730.6	5	6	18	10	17	2	4	5	6	9	9	18	9	11	3	15	5	4	14	20	14	7	13	6

Figure 20 : Résultats pour les cinq premières générations

Les flèches indiquent les individus sélectionnés pour la génération suivante.

- Les gènes des neuf individus de la 1^{ère} génération sont obtenus aléatoirement. Les individus N°4 et N°6 sont sélectionnés pour devenir les parents de la génération suivante. Tous les autres individus sont éliminés.
- Dans la 2^{ème} génération, l'individu N°12 est sélectionné car il possède la meilleure fonction objectif. L'individu N°10 qui est pourtant le deuxième meilleur individu n'est pas sélectionné pour cause de consanguinité trop forte (13 gènes identiques sur 24). C'est finalement l'individu N°15 qui est retenu comme 2^{ème} parent.
- Dans la 3^{ème} génération, le même processus se produit puisque l'individu N°22 n'est pas sélectionné pour cause de consanguinité trop élevée.
- L'individu N°31 (Délai = 621,5 sec.) de la 4^{ème} génération a subi une mutation : son gène TiLa_(23,24) (avant dernier du chromosome) est passé de 20 à 19. Cette mutation a permis d'obtenir le 2nd meilleur parent de cette génération, que l'on retrouve dans la génération suivante. En effet, l'individu N°35, qui a le même site de croisement que l'individu N°31, a une performance moins bonne (Délai = 628,8 sec.).
- Les parents des individus N°41 et N°42 de la 5^{ème} génération ont un site de croisement qui est à une position où les gènes qui suivent sont identiques pour les deux parents. Les enfants sont donc des reproductions parfaites, c'est-à-dire des "Clones" de leurs parents, d'autant plus qu'ils n'ont subi aucune mutation.

La figure 21 montre les deux dernières générations de l'algorithme.

Iter	Mini	Moy	Délai	TiLa 1 à TiLa 24																							
				3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
127	589.1	661.5	589.1	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
128	589.1	661.2	628.1	3	19	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	15	17
129	589.1	661.5	695.5	3	6	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	15	17
130	589.1	661	598.4	3	19	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
131	589.1	661	669	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	15	17
132	589.1	661.1	668.2	3	19	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	20	4
133	589.1	661.4	695.5	3	6	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	15	17
134	589.1	660.9	597.8	3	19	13	13	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
135	589.1	661	677	14	8	8	20	8	18	7	11	20	19	10	1	16	9	4	18	20	3	15	5	18	4	18	1
136	589.1	660.5	589.1	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
137	589.1	660.2	628.1	3	19	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	15	17
138	589.1	660.3	669	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	9	7	17	15	17
139	589.1	660.3	657.8	3	19	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	15	9	5	20	4
140	589.1	660.5	695.5	3	6	4	18	12	16	4	16	12	12	20	15	9	1	12	4	14	6	10	9	7	17	15	17
141	589.1	660.1	598.4	3	19	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
142	589.1	660.7	739.6	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
143	589.1	660.9	695.4	3	19	4	18	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4
144	589.1	661.3	722	12	2	9	10	6	5	11	10	4	5	18	15	7	12	5	16	5	3	6	20	9	11	11	13

Figure 21 : Les deux dernières générations

1.5-b) Améliorations possibles de l'algorithme génétique au vu des résultats :

Compte tenu du nombre de paramètres intrinsèques et des multiples possibilités d'adaptation de cet algorithme, nous nous contenterons de suggérer des améliorations et de vérifier si elles apportent un gain. L'objectif de cette thèse est plus orienté vers la méthodologie et la validation d'un processus d'optimisation intégrant plusieurs outils, plus que vers le développement et le perfectionnement d'un algorithme particulier. Il s'agit essentiellement de justifier de l'intérêt ou pas d'utiliser un tel algorithme pour optimiser un système de production. Dans [VINCENT 98], on trouve, de façon complémentaire à ce travail, une comparaison sur le perfectionnement d'algorithmes génétiques.

Les propositions d'améliorations sont les suivantes :

- En regardant attentivement les combinaisons de gènes, on remarque que pour certaines générations, on obtient des enfants identiques (des "jumeaux"), mais obtenus à partir de croisements différents. Les individus N°129 et N°133 de l'avant dernière génération illustrent ce phénomène. Cela est dû au site de croisement qui est le même pour ces deux individus. L'algorithme choisit un site de croisement au hasard, excepté le site du croisement immédiatement précédent (voir procédure de croisement, figure 17). On peut donc imaginer un algorithme qui ne choisirait pas plusieurs fois le même site de croisement dans une génération. Une autre possibilité consisterait à garder le même site, mais à effectuer une mutation aléatoire d'un des gènes.
- On trouve aussi des enfants qui sont des clones des parents. Pour éviter ce phénomène, on peut placer le site de croisement à une position où la séquence de gènes est différente entre les deux parents.

□ 1.6- Synthèse et conclusion

Les meilleures solutions obtenues par chacun des algorithmes sont rassemblées dans le tableau de la figure 22. La dernière ligne correspond à la solution théorique.

	Iter	Mri	Mby	Déjà	TiLa 1 à TiLa 24																								
Exploration aléat.	100	604.1	682	604.1	6	19	18	16	5	3	12	13	3	17	14	14	5	13	15	18	5	19	15	11	6	1	12	15	
Descente	81	610	651.4	610	10	4	5	2	13	13	18	17	10	17	18	11	12	20	7	8	12	4	14	7	7	20	19	14	
Recuit simulé	118	600.3	656	600.3	10	2	15	5	14	15	19	8	9	17	6	19	12	9	9	19	2	17	15	7	6	10	19	13	
Algo. Génétique	87	589.1	659.7	589.1	3	6	13	12	16	6	9	7	10	15	14	12	19	15	14	4	1	16	17	15	9	5	20	4	
Déjà théorique				577	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16

Figure 22 : Synthèse des meilleures combinaisons obtenues

Le graphique de la figure 23 présente les tracés des délais minimums obtenus par chaque algorithme.

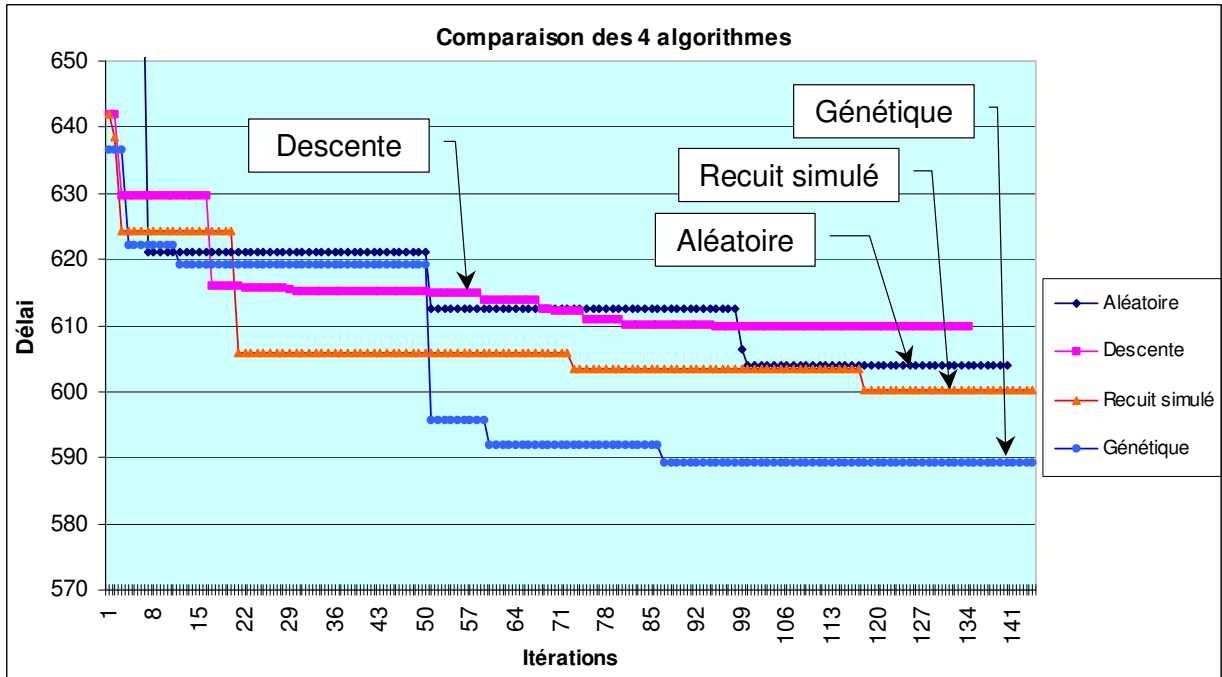


Figure 23 : Tracés des délais minimums de chaque algorithme

On remarque dans le tableau de synthèse que les solutions de chaque algorithme sont bien distinctes les unes des autres. Il n'y a pas de corrélation directe avec la solution théorique.

Le graphique fait ressortir la supériorité de l'algorithme génétique mais, comme on l'a déjà remarqué, la performance de l'exploration aléatoire est à la fois meilleure que l'algorithme de descente et proche du recuit simulé. Cela peut s'expliquer par le fait que le problème qui sert de support à ce test est très complexe en terme de résolution et que, sur un espace de recherche aussi grand, les différents algorithmes de recherche "organisée" ont du mal à se démarquer par rapport à une exploration totalement aléatoire. L'algorithme génétique, qui est le plus performant, combine, comme on l'a vu, une exploration aléatoire à une recherche organisée. C'est vraisemblablement ce qui justifie son efficacité. Au vu de ces résultats, on peut considérer que pour traiter un problème d'optimisation un peu moins complexe que celui-ci, avec une fonction objectif moins "bruitée", cet algorithme donnera de toute façon, une "bonne" solution. Pour la suite de ce travail, nous allons donc utiliser un algorithme génétique et revenir sur le problème d'optimisation d'une ligne d'assemblage multiproduit.

II- Optimisation du délai de production pour le cas multiproduit

Les conditions expérimentales et les caractéristiques du jeu d'essai N°2 et des expérimentations associées (horizon fini et horizon infini) sont données dans la partie 6, chapitre 2.

□ II.1- Choix des variables et codage du problème

Pour les expérimentations réalisées dans la partie 6, nous avons pris en compte deux variables de pilotage :

- le Temps Inter Lancement (TiLa) qui était constant entre chaque palette de la campagne;
- la capacité des stocks amont (StAm) qui était la même sur chaque poste.

L'objectif de ce chapitre est le même que dans la partie 6 : Trouver une combinaison des variables de pilotage qui minimise le délai de production (D) d'une campagne dans le cas d'une production à horizon fini, ou qui minimise le "Takt Time" (voir partie 6, chap. 2.3-a) dans le cas d'une production à horizon infini.

Mais nous avons surtout un autre objectif qui est d'**améliorer les résultats trouvés dans la partie 6**. Pour tenter d'y arriver et pour exploiter les possibilités de l'algorithme génétique, les contraintes placées sur les variables vont être enlevées :

- En ce qui concerne le Temps Inter Lancement (TiLa), sans aller jusqu'au cas extrême d'un (TiLa) variable entre chaque palette, il pourra varier d'un lot de palettes appartenant à un Ordre de Fabrication (O.F) à un autre. Dans ce jeu d'essai, la campagne comportant cinq ordres de fabrication, il y a donc cinq variables homogènes ($TiLa_{(k)}$), avec k = indice de l'O.F, qui peuvent chacune prendre une valeur comprise entre 1 et 12 secondes. Le nombre de combinaisons de ces variables est de $[(12 - 1) + 1]^5 = 248\ 832$. Ce nombre est beaucoup plus petit que celui obtenu pour les tests du chapitre 1, mais il faut le rapporter à la durée que prendrait une vérification exhaustive de ces combinaisons avec le modèle de simulation. Sur un ordinateur de type PC Pentium II à 300 Mhz, une simulation dure à peu près 1 seconde. Pour tester toutes les combinaisons, il faudrait donc $248\ 832 / (60 \times 60 \times 24) = 2,88$ jours, ce qui est inacceptable.
- La capacité des stocks amont peut être différente d'un poste à l'autre. Il y a donc 6 variables homogènes ($StAm_{(i)}$) avec i = indice du poste, qui peuvent chacune prendre 6 valeurs discrètes comprises entre 1 et 6 palettes. Le nombre total de combinaisons de ces variables est de $[(6 - 1) + 1]^6 = 46\ 656$.

Bien que toutes les contraintes ne soient pas levées et que l'étendue de variation de (TiLa) soit plus petite que pour le test du chapitre 1, le nombre total de combinaisons est de $(248\ 832 \times 46\ 656) \cong 1,16 \cdot 10^{10}$. En durée de simulation, il faudrait presque quatre siècles pour tout tester !

La priorité de l'optimisation est d'atteindre **rapidement** une performance de niveau satisfaisant. On fixe donc arbitrairement une durée maximale du processus de simulation/optimisation à 5 minutes, soit environ 300 itérations de 1 seconde chacune sur l'ordinateur utilisé.

□ II.2- Optimisation en horizon fini

II.2-a) Rappel des résultats obtenus sans algorithme d'optimisation :

Avec un nombre de palettes à lancer (N_p) limité à 25, nous avons obtenu, en testant les 120 combinaisons de variables avec le modèle de simulation, un délai minimal (D_{pr}) de 579 secondes (solution : $StAm=6$ et $TiLa = 9$ sec.).

Le délai minimal estimé, qui n'a été atteint par aucune des combinaisons, a été calculé de deux façons (voir partie 6, chapitre II.2-c) :

- soit en utilisant la méthode de calcul d'une production de type monoproduit. Dans ce cas, on obtient un délai (D_{th1}) de 531 secondes;
- soit en utilisant une méthode basée uniquement sur la charge du poste goulet, sans prendre en compte les autres postes. On obtient alors un délai (D_{th2}) de 397 secondes.

L'objectif de cette optimisation est de vérifier s'il est possible, en utilisant un algorithme génétique tel que celui présenté dans le chapitre 1, de se rapprocher du délai minimal théorique, ou au moins de réduire le délai obtenu avec la méthodologie employée dans la partie 6 et ceci en un minimum de temps.

II.2-b) Résultats obtenus par l'algorithme d'optimisation :

Le graphique de la figure 24 présente les résultats obtenus avec l'algorithme génétique présenté dans le chapitre 1. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont données dans le tableau de la figure 25.

Les paramètres de l'algorithme choisis pour cette expérimentation sont les suivants :

- Nombre d'individus par génération : 9
- Nombre de gènes : 11 (5 pour TiLa(1) à TiLa(5), 6 pour StAm(1) à StAm(6))
- Nombre de générations : 15 + 1 (génération aléatoire)
- Taux de mutation par génération : 0,6
- Consanguinité maximale entre les parents sélectionnés : 8 gènes

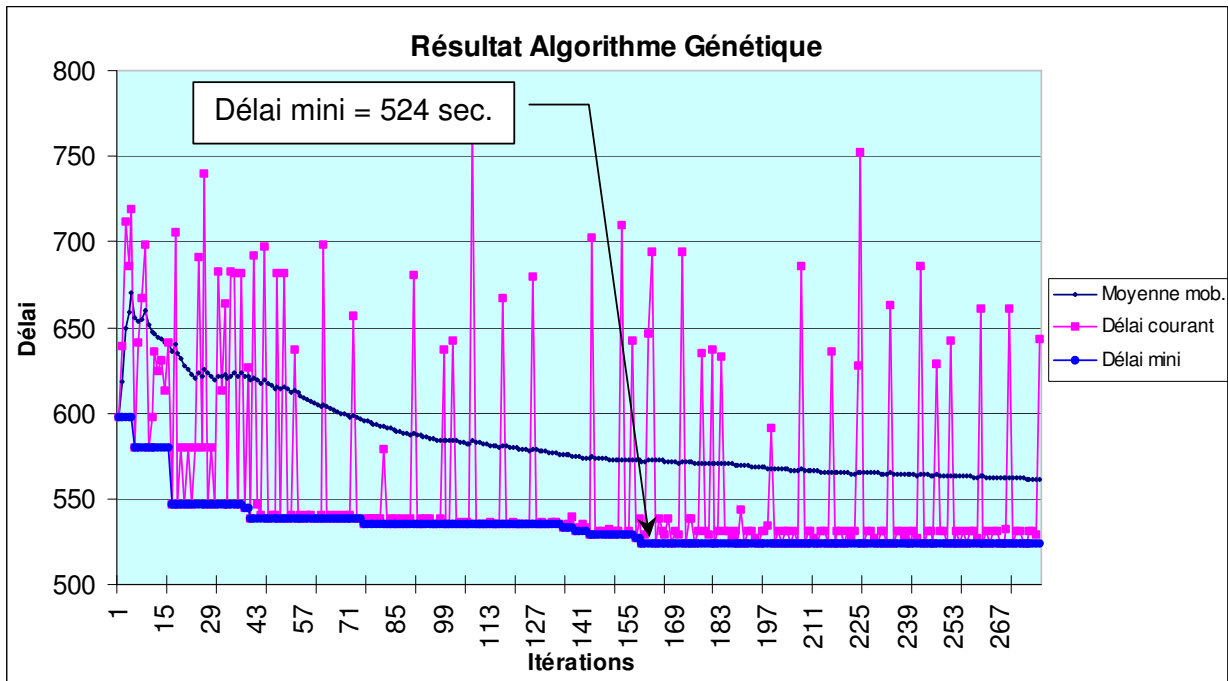


Figure 24 : Délai = $f(\text{itération})$ pour Horizon Fini

Iter	Mini	Moy	Délai	TiLa 1 à TiLa 5					StAm 1 à StAm 6					
1	597.8	597.8	597.8	2	7	11	5	6	4	6	5	4	3	2
6	579.6	655.5	579.6	9	9	7	7	8	1	6	4	5	5	5
17	547.1	635.8	547.1	2	7	11	7	8	1	6	4	5	5	5
39	544.2	621	544.2	2	5	3	8	9	6	4	2	3	6	5
41	538.3	619.2	538.3	2	7	3	8	9	6	4	2	3	6	6
75	535.1	596	535.1	2	7	4	8	9	6	4	4	3	5	5
96	535	585	535	2	8	4	8	9	6	4	4	5	5	6
135	532.8	575.9	532.8	3	11	5	6	1	4	4	2	6	5	5
139	531.1	574.8	531.1	2	11	5	6	1	4	4	2	6	5	5
143	529.1	573.6	529.1	2	8	4	8	9	5	4	4	6	5	5
150	528.8	572.7	528.8	2	11	5	8	9	5	4	4	6	5	5
157	526.8	572.6	526.8	2	11	5	6	9	5	4	4	6	5	5
159	523.9	572.1	523.9	2	11	5	6	1	5	4	4	6	5	5

Figure 25 : Meilleurs résultats aux paliers

La performance de l'algorithme est visible sur la figure 24 : il trouve une solution ($D_{opt} = 524$ sec.) qui permet de réduire sensiblement le délai de la campagne par rapport au délai obtenu dans la partie 6 ($D_{pr} = 579$ sec.). Cette solution est obtenue à la 159^{ème} itération, c'est-à-dire à peine plus que le nombre nécessaire pour tester toutes les combinaisons (120) dans la partie 6. La solution qui donne le délai minimal est :

2 - 11 - 5 - 6 - 1 / 5 - 4 - 4 - 6 - 5 - 5

Les écarts relatifs aux différents résultats obtenus en pratique ou calculés sont :

- par rapport au délai obtenu dans la partie 6 : $D_{pr} - D_{opt} = (579 - 524) = 57$ secondes, soit une amélioration de 9,5%.
- par rapport au délai estimé (relation 1) : $D_{opt} - D_{th1} = (524 - 531) = -7$ secondes
- par rapport au délai estimé (relation 2) : $D_{opt} - D_{th2} = (524 - 397) = 127$ secondes

Le tableau de la figure 26 récapitule les différentes solutions trouvées en pratique et en théorie :

	TiLa(1)	TiLa(2)	TiLa(3)	TiLa(4)	TiLa(5)	StAm(1)	StAm(2)	StAm(3)	StAm(4)	StAm(5)	StAm(6)
D _{pr} = 579 sec.	9	9	9	9	9	6	6	6	6	6	6
D _{th1} = 531 sec.	8	8	9	9	7	de 1 à 6 (sans effet sur le délai)					
D _{th2} = 397 sec.	Solution non déterminée					Solution non déterminée					
D _{opt} = 524 sec.	2	11	5	6	1	5	4	4	6	5	5

Figure 26 : Récapitulation des solutions trouvées

Pour optimiser un problème comme celui-ci, l'algorithme génétique utilisé montre son efficacité, puisqu'il permet d'obtenir un délai plus faible que celui trouvé dans la partie 6, en un temps pratiquement identique. Ce résultat permet aussi de valider, sur cette campagne, la relation (D_{th1}), qui fournit un délai proche de (D_{opt}). Par contre, nous sommes encore loin du résultat donné par la relation (D_{th2}). Cela peut être dû au fait que l'on n'a pas libéré toutes les contraintes sur la variable (TiLa) qui ne peut pas varier d'une palette à l'autre, mais seulement d'un O.F à l'autre. Mais c'est plus vraisemblablement dû à la quantité de palettes utilisées (N_p) pour la campagne. Comme cela avait été montré dans la partie 6 où les expérimentations avec trente puis quarante-quatre palettes avaient donné de meilleurs résultats, les vingt-cinq palettes utilisées ne sont sûrement pas suffisantes pour assurer la production de la campagne sans attente. En contrepartie, on avait vu qu'un nombre plus élevé de palettes augmentait l'encours moyen.

Il faudrait donc essayer de trouver une solution qui permette de minimiser le délai en utilisant une quantité minimale de palettes, tout en minimisant l'encours moyen. Il faut dans ce cas que l'algorithme d'optimisation puisse évoluer en fonction non pas d'un seul critère, mais de plusieurs critères. C'est ce problème qui sera développé dans le chapitre III.

□ II.3- Optimisation en horizon infini

II.3-a) Rappel des résultats obtenus sans algorithme d'optimisation :

Dans le cas d'une production en horizon infini, on mesure le délai entre la livraison des 44 palettes de la campagne N° 5, en régime stabilisé. Avec un nombre de palettes à lancer (N_p) limité à 25, nous avons obtenu, en testant les 120 combinaisons de variables avec le modèle de simulation, un délai minimal (D_{pr}) de 422 secondes (solution : $StAm=6$ et $Tila = 9$ sec.).

L'objectif de l'algorithme d'optimisation est donc d'améliorer le résultat obtenu dans les expérimentations exhaustives de la partie 6.

II.3-b) Résultats obtenus par l'algorithme d'optimisation :

Le graphique de la figure 26 présente les résultats obtenus avec l'algorithme génétique présenté dans le chapitre 1. Les solutions qui correspondent aux différents paliers du tracé du délai minimal sont données dans le tableau de la figure 27.

Les paramètres de l'algorithme choisis pour cette expérimentation sont les suivants :

- Nombre d'individus par génération : 7
- Nombre de gènes par individu : 11
- Nombre de générations : 42
- Taux de mutation par génération : 0,6
- Consanguinité maximale entre les parents sélectionnés : 8 gènes

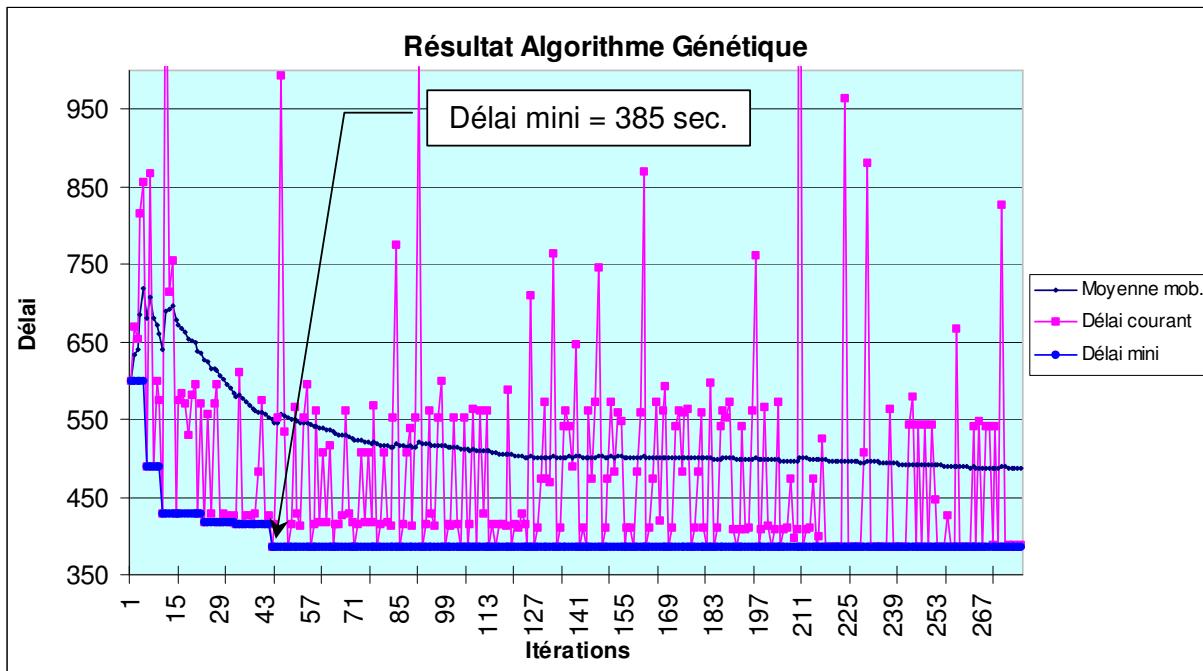


Figure 27 : Délai = $f(\text{itération})$ pour Horizon Infini

Iter	Mini	Moy	Délai	TiLa 1 à TiLa 5					StAm 1 à StAm 6					
1	598.7	598.7	598.7	2	7	11	5	6	4	6	5	4	3	2
6	488.8	680.2	488.8	9	9	7	7	8	1	6	4	5	5	5
11	429.3	640.1	429.3	2	7	7	7	8	1	6	4	5	5	5
24	418.1	626.5	418.1	2	9	11	5	8	1	6	4	5	5	5
45	385	549.3	385	2	8	7	7	8	1	6	4	5	5	5

Figure 28 : Meilleurs résultats aux paliers

Comme pour la production à horizon fini, l'algorithme d'optimisation améliore sensiblement le délai : il trouve une solution ($D_{opt} = 385 \text{ sec.}$) à la 45^{ème} itération, ce qui est assez remarquable. Au vu de la figure 27, il peut paraître intéressant d'inclure dans l'algorithme une procédure permettant de l'arrêter au bout d'un certain nombre d'itérations sans amélioration. Mais il faut garder à l'esprit qu'à chaque génération, une mutation d'un gène ou la naissance d'un mutant peut produire une amélioration. L'algorithme peut donc trouver une solution après être resté pendant un grand nombre d'itérations sur un plateau. La solution qui donne le délai minimal est :

2 - 8 - 7 - 7 - 8 / 1 - 6 - 4 - 5 - 5 - 5

L'écart relatif au résultat obtenu en pratique dans la partie 6 est :

$$D_{pr} - D_{opt} = (422 - 385) = 37 \text{ secondes, soit une amélioration de } 8,8\%.$$

Le tableau de la figure 29 compare les deux solutions.

	TiLa(1)	TiLa(2)	TiLa(3)	TiLa(4)	TiLa(5)	StAm(1)	StAm(2)	StAm(3)	StAm(4)	StAm(5)	StAm(6)
D _{pr} = 422 sec.	9	9	9	9	9	6	6	6	6	6	6
D _{opt} = 385 sec.	2	8	7	7	8	1	6	4	5	5	5

Figure 29 :Récapitulation des solutions trouvées

La conclusion de cette optimisation en horizon infini est sensiblement la même que pour l'horizon fini. L'efficacité de l'algorithme est indéniable et l'obtention d'une bonne solution est très rapide. La solution obtenue par l'algorithme est remarquable, surtout en ce qui concerne la valeur de StAm(1). En effet, les résultats de la partie 6 montraient une réduction du délai proportionnelle à la taille des stocks amont. La solution de l'algorithme montre que tous les stocks n'ont pas la même influence et que, même si globalement les capacités des stocks doivent être proches de leur valeur maximale (6), on peut minimiser le délai avec des capacités de certains stocks à leur valeur minimale (1), ce qui est le cas du poste N°1.

III- Optimisation multicritère

L'optimisation multicritère a été réalisée avec les deux jeux d'essais, mais seul le jeu N°2 est présenté dans ce chapitre car il est plus représentatif à cause de sa complexité.

Les conditions expérimentales et les caractéristiques du jeu d'essai N°2 et des expérimentations associées (horizon fini et horizon infini) sont données dans la partie 6, chapitre 2.

Comme dans le chapitre 2, nous prendrons en compte les mêmes variables de pilotage, avec les mêmes contraintes :

- le Temps Inter Lancement (TiLa), pouvant varier d'un lot de palettes appartenant à un O.F. à un autre, ce qui donne donc cinq variables homogènes et indépendantes (TiLa_(k)), pouvant chacune prendre une valeur comprise entre 1 et 12 secondes.
- la capacité des stocks amont (StAm), pouvant être différente d'un poste à l'autre, ce qui donne six variables homogènes et indépendantes (StAm_(i)), pouvant chacune prendre une valeur discrète comprise entre 1 et 6 palettes.

Nous allons, pour ce chapitre, introduire une autre variable qui avait été évoquée lors des précédentes expérimentations de la partie 6 :

- le nombre de palettes en réserve (R_p), dont la valeur peut être comprise entre 20 et 35 palettes.

Jusqu'à présent, dans toutes les expérimentations qui ont précédé, que ce soit avec ou sans algorithme d'optimisation, nous n'avons considéré qu'un seul critère (ou objectif) – le délai (D) – pour optimiser le pilotage de la ligne. Bien que ce critère soit déterminant, le pilotage optimal d'une ligne de production vise d'autres objectifs, plus ou moins importants, dont il faut pouvoir tenir compte.

□ III.1- Choix des critères

Comme cela a déjà été évoqué, il serait intéressant de considérer, simultanément au délai, l'encours moyen (EC). En effet, dans le chapitre II.2-d de la partie 6, on avait constaté qu'une augmentation du nombre de palettes en réserve permettait une réduction sensible du délai. Mais l'augmentation du nombre de palettes en réserve entraîne un surcoût d'exploitation (il faut acheter plus de palettes), qui peut dans certains cas ne pas se justifier par rapport à la réduction de délai obtenu, même si les palettes sont réutilisables d'une production à une autre. Sans entrer dans des considérations économiques, de toute façon difficiles à évaluer dans un contexte de production flexible, on a pu remarquer précédemment que l'augmentation du nombre de palettes en réserve, si elle a l'avantage de permettre une réduction du délai (D), présente l'inconvénient d'augmenter l'encours moyen (EC) qui s'accompagne d'une perte de fluidité du processus (voir Ratio de fluidité (R_f), chapitre I.2-b de la partie 6). Nous allons donc aussi prendre en compte l'encours moyen (EC), qu'il faudra minimiser.

Sur ce type de ligne, chaque poste de travail peut être amené à réaliser des opérations différentes, sur des produits différents. En général, chaque changement d'opération entraîne un changement d'outillage, un réglage, un changement de programme ou, s'il s'agit d'un poste manuel, des manipulations différentes. Le passage d'une opération à une autre sera appelé sans distinction "Réglage" dans la suite de ce travail. Ce réglage a pour conséquence de prendre du temps, donc de baisser la productivité de la ligne. C'est une notion que l'on retrouve dans le concept du Juste-à-Temps sous l'acronyme S.M.E.D, qui signifie "Single Minute Exchange of Die" (voir pour plus de détails à ce sujet [SHINGO 90]) et dont l'objectif est de réduire les temps de réglage entre deux ordres de fabrication. Comme pour l'aspect économique, la durée des réglages peut difficilement être prise en compte dans ce contexte de production flexible car elle peut être très variable d'un réglage à un autre. L'hypothèse qui est faite ici est donc de ne pas considérer directement les temps de réglage, mais le nombre total de réglages (Reg) ou de changement d'O.F sur l'ensemble des postes de travail. On cherchera donc à réduire le nombre de réglages (Reg), c'est-à-dire à se rapprocher d'une ligne de production fonctionnant par lot (Batch models line) alors que la ligne travaille en production multiproduit (Mixed models line).

L'optimisation du pilotage de cette ligne prendra donc en compte trois critères, chacun à minimiser :

- le délai de production (D),
- l'encours moyen (EC),
- le nombre total de réglage (Reg).

Il serait possible de considérer encore d'autres critères, comme par exemple la charge des postes de travail (à maximiser), le ratio de fluidité (à maximiser), le temps de défilement moyen (à minimiser), le temps d'attente moyen devant les postes de travail (à minimiser), etc.... Certains de ces critères sont directement dépendants de ceux retenus et n'apportent donc pas d'information complémentaire. D'autre part, quand le nombre de critères augmente, le problème de l'importance relative de chacun d'eux se fait encore plus difficile. Il est donc souhaitable de se limiter à deux ou trois critères indépendants, comme cela est fait ici.

□ III.2- Problématique de l'optimisation multicritère

L'optimisation multicritère est vraisemblablement un problème plus en rapport avec la réalité industrielle que l'optimisation monocritère. Comme le souligne [MOLET 97] déjà cité dans la partie 1, la gestion de production est souvent une question de compromis entre plusieurs objectifs qu'il faut pouvoir prendre en compte simultanément. En optimisation monocritère, on recherche l'extremum d'une fonction objectif. En optimisation multicritère, la notion d'optimalité n'est pas aussi évidente. En effet, pour un même problème, une solution peut être bonne pour un critère et mauvaise pour un autre, pendant qu'une autre solution peut s'avérer moyenne pour chaque critère. Dans un cas comme celui-ci, il est difficile de dire objectivement laquelle est la meilleure.

Plusieurs méthodes d'optimisation multicritère ont été développées par les chercheurs. Dans [COHON 78], on trouve une description détaillée de ces méthodes. [BENTLEY 98] en teste six à partir d'un algorithme génétique optimisant quatre fonctions mathématiques. [SHAW 96] et [MARCO 99] présentent dans leurs articles des méthodes multicritères utilisant un algorithme génétique. Ces méthodes sont issues des travaux de [GOLDBERG 94] qui a énoncé le principe de la création d'espèces et de "niches écologiques". Dans la suite de ce travail aussi, c'est le même algorithme génétique que celui présenté dans le chapitre 1 qui sera utilisé pour l'optimisation multicritère. Les principales différences se situent au niveau de la description de la fonction objectif ou de la sélection des individus en fonction des critères auxquels ils répondent.

On trouve principalement deux méthodes de classement des solutions en optimisation multicritère, chacune comportant plusieurs variantes :

- La première consiste à faire une combinaison linéaire des différents critères pour obtenir une fonction objectif multicritère. On peut appliquer une pondération à chacun de ces critères en fonction de leur importance. Mais c'est l'étape la plus délicate et la plus subjective car un choix inadapté des coefficients peut entraîner l'obtention de solutions incorrectes. Exprimé mathématiquement, on obtient :

Soit $F(x)$, fonction objectif globale, avec $[x]$ = vecteur solution du problème.

Soit $f_i(x)$, critère i .

Soit W_i , poids du critère i et n = nombre de critères.

On peut écrire $F(x) = W_1 \cdot f_1(x) + W_2 \cdot f_2(x) + \dots + W_n \cdot f_n(x) \Leftrightarrow$

$$F(x) = \sum_{i=1}^n W_i \cdot f_i(x)$$

Avec cette méthode, simple à programmer mais assez controversée par [GOLDBERG 94] entre autres, qui "refuse de comparer les pommes avec les oranges", on obtient une solution unique. Ce sera la meilleure solution au terme des itérations de l'algorithme, c'est-à-dire celle qui aura donné le minimum de $F(x)$.

- La deuxième méthode utilise le concept d'optimalité de Pareto ou de "Non Domination". Une solution est dite "**Pareto-Optimale**" (P.O) ou "**Non dominée**" s'il n'en n'existe pas d'autre qui permette d'améliorer la valeur d'un ou plusieurs critères sans dégrader celle d'au moins un autre critère. Exprimé mathématiquement, on obtient :

Soit $[F(x)]$, vecteur des différents critères (ou fonctions objectif indépendantes), avec :

$$[F(x)] = [f_1(x), f_2(x), \dots, f_n(x)] \text{ avec } f_i(x) = \text{critère } i \text{ et } n = \text{nombre de critères.}$$

Soit $[X]$, vecteur solution dont les p composantes sont les p variables du problème, avec :

$$[X] = [a_1, a_2, \dots, a_p]$$

Soit $[X_j]$, une solution j particulière du problème, avec $j = 1..m$, m = nombre total de solutions.

Pour un problème de minimisation, une solution $[x_1]$ est dite partiellement meilleure qu'une solution $[x_2]$ (ou dominante au sens de Pareto) si et seulement si :

$$\forall i, f_i(x_1) \leq f_i(x_2), \exists i / f_i(x_1) < f_i(x_2)$$

Avec cette méthode, on peut obtenir un ensemble de solutions non dominées au terme de l'optimisation, avec à l'extrême tous les individus de la dernière génération satisfaisant l'optimalité de Pareto. Si l'on doit disposer de plusieurs bonnes solutions, cela est avantageux par rapport à la méthode précédente, mais dans le cas où une seule solution est nécessaire, l'utilisateur devra choisir plus ou moins arbitrairement une solution parmi

toutes les solutions Pareto-Optimales. D'autre part, cette méthode ne permet pas, contrairement à la première, de pondérer les critères afin de leur attribuer plus ou moins d'importance.

Comme on vient de le voir, l'optimisation multicritère peut être traitée d'au moins deux façons sans qu'il soit possible de dire objectivement laquelle est la meilleure. Ce sujet est donc très ouvert au niveau de la recherche, et dans la suite de ce travail, nous allons donc apporter une contribution à ce problème à partir de l'analyse d'expérimentations successives menées sur le modèle de simulation de la ligne expérimentale, en utilisant à la fois l'algorithme d'optimisation présenté dans le chapitre 1 et les méthodes de sélection des solutions décrites précédemment ou des variantes de celles-ci.

□ III.3- Développement expérimental d'une méthode d'optimisation multicritère

Dans ce chapitre, l'objectif est de mettre en œuvre, par affinements successifs et sur le cas d'une production multiproduit à horizon fini, une méthode d'optimisation multicritère. Les paramètres de l'algorithme génétique utilisé pour développer cette méthode sont les suivants :

- Nombre de gènes : 12 (5 pour TiLa(1) à TiLa(5), 6 pour StAm(1) à StAm(6) et 1 pour Rp)
- Nombre d'individus par génération : 9
- Nombre de générations : 30 + génération initiale aléatoire
- Taux de mutation par génération : 0,6
- Consanguinité maximale entre les parents sélectionnés : 8 gènes

III.3-a) Méthode 1 : Optimisation par Sommation (O.S)

La fonction objectif est tout simplement égale à la somme de tous les critères, soit :

$$F(x) = D(x) + EC(x) + Reg(x)$$

Les résultats (courant, mini, moyenne) obtenus pour chacun des critères et pour la fonction objectif au cours des différentes itérations de l'algorithme sont présentés sur les graphiques des figures 30 à 33. La figure 34 montre l'évolution de la valeur minimale de la fonction objectif et les valeurs des critères correspondants.

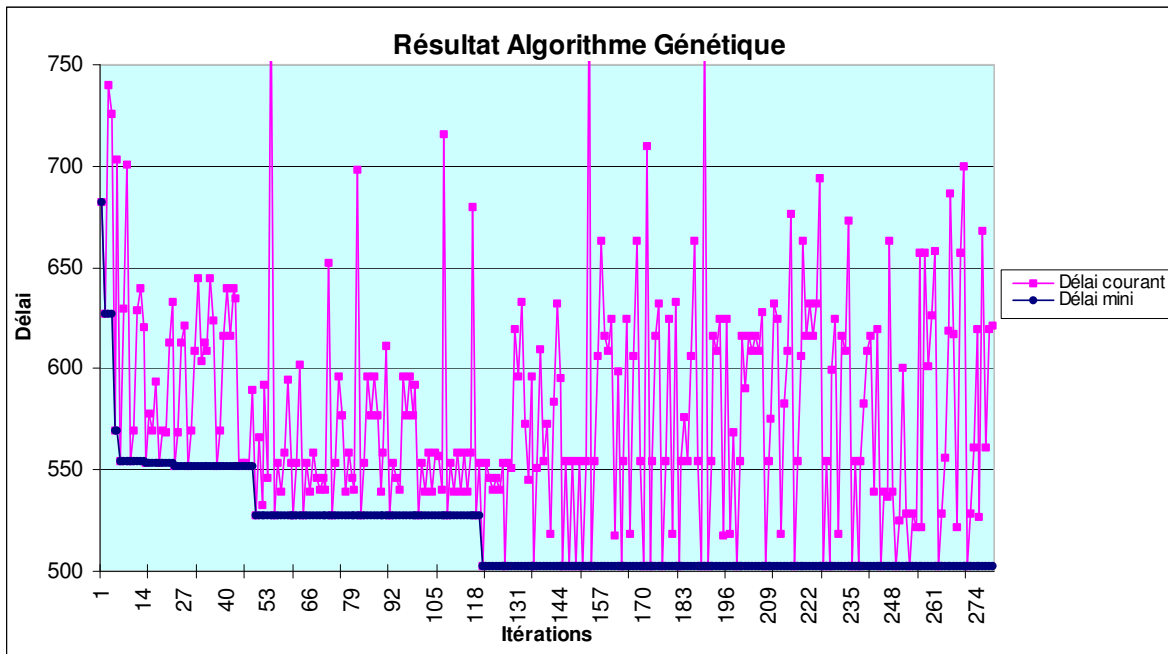


Figure 30 : Délai = $f(\text{itérations})$

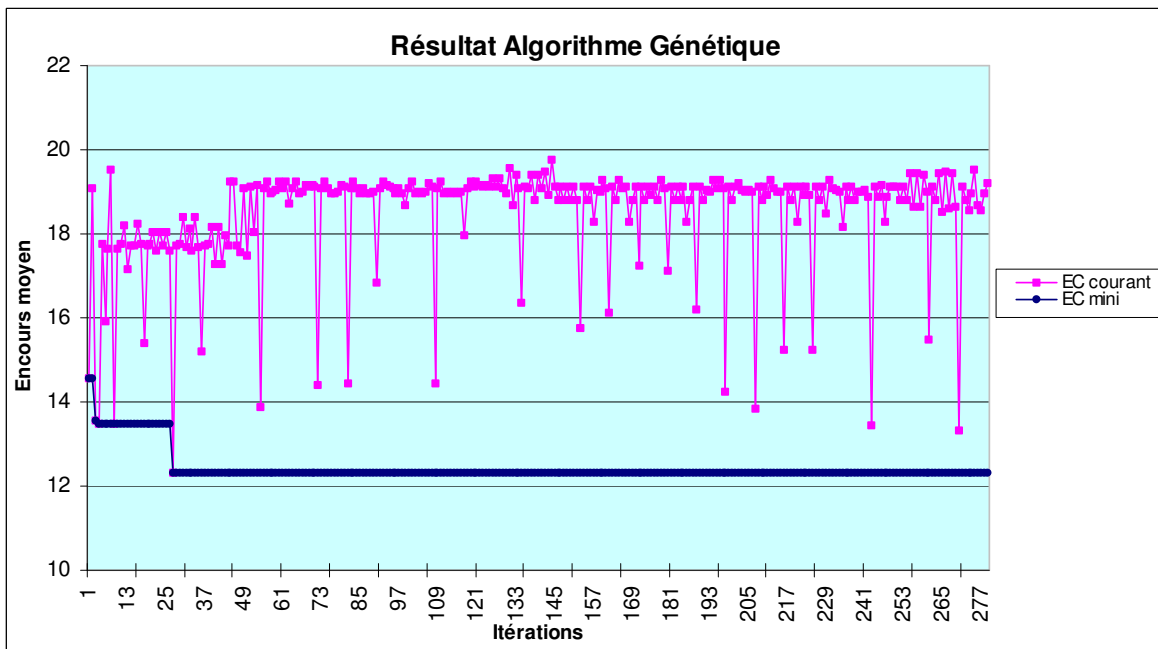


Figure 31 : Encours moyen = $f(\text{itérations})$

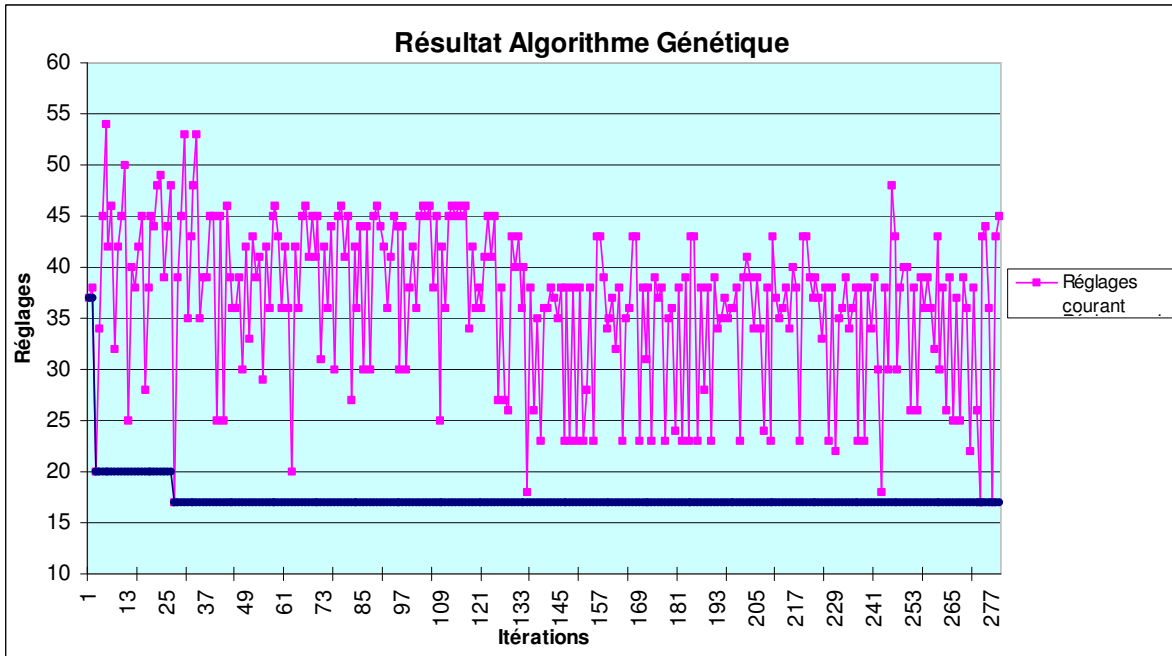


Figure 32 : Réglages = $f(\text{itérations})$

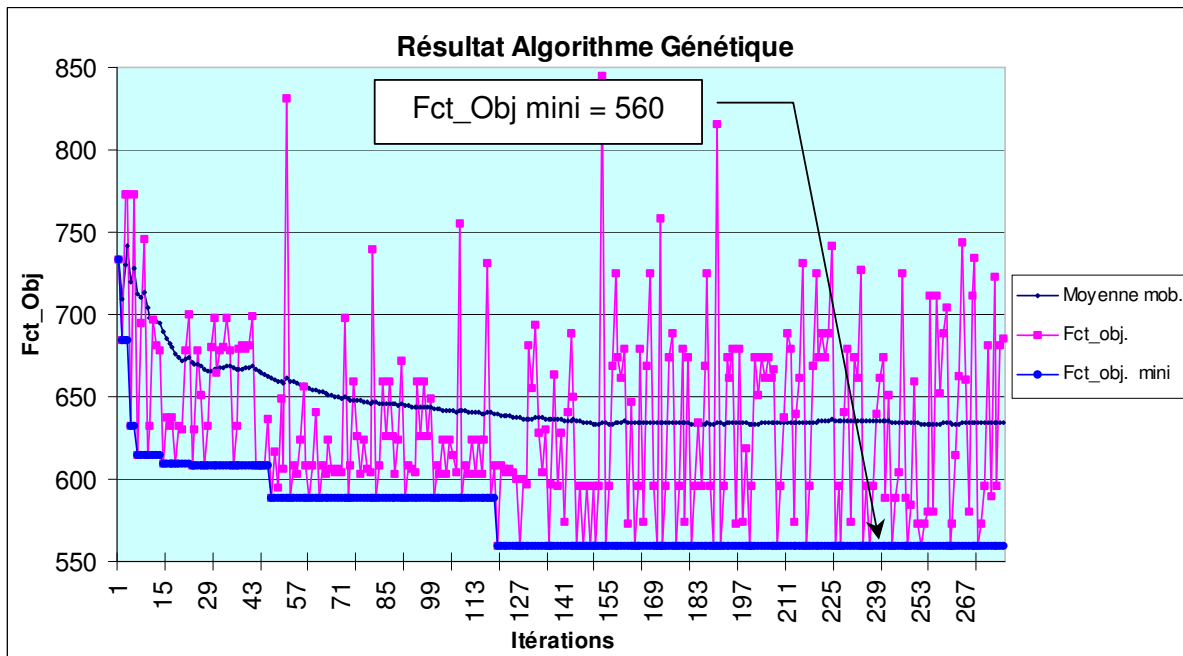


Figure 33 : Fonction Objectif = $f(\text{itérations})$

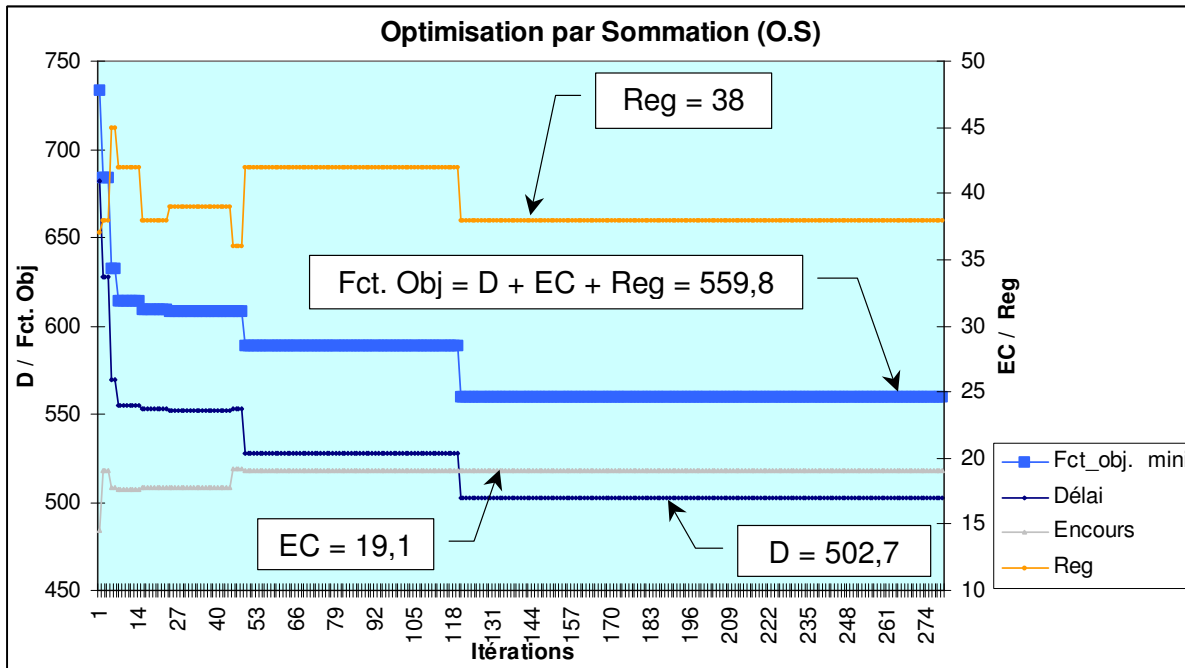


Figure 34 : Evolution de la fonction objectif et des critères

L'analyse des graphiques des figures 30 à 34, et en particulier la synthèse de la figure 34 met en évidence le point faible et dommageable de cette méthode : chaque critère a une étendue de variation différente des autres, ce qui déséquilibre l'importance des uns et des autres. Au cours des itérations, le délai (D) varie de 503 secondes à 801 secondes (étendue = $801 - 503 = 298$ secondes), l'encours moyen varie de 12,3 palettes à 19,8 palettes (étendue = $19,8 - 12,3 = 7,5$ palettes), et les réglages varient de 17 à 54 (étendue = $54 - 17 = 37$ réglages). Ces différences d'étendue de variation impliquent que le critère qui a la plus grande étendue, ici le délai (D), atténue l'effet des autres critères d'étendue moindre. Cela se vérifie bien sur le graphique des critères (figures 30 à 32) où l'on constate que l'algorithme maintient le délai minimal de chaque génération à la valeur minimale trouvée, alors que pour les réglages (Reg) et encore plus pour l'encours (EC), l'algorithme ne converge pas vers cette valeur minimale. Même si ce phénomène est normal pour une optimisation multicritère (on ne peut pas avoir simultanément un délai minimal avec un encours minimal et en plus un nombre de réglages minimal), la différence d'étendue biaise les résultats en avantageant le critère de plus grande étendue.

Cette méthode, très rudimentaire, nécessite donc une adaptation de l'étendue de chaque critère. Cela va être réalisé en appliquant un "normage" des critères, c'est-à-dire une adaptation permettant à chacun d'eux d'évoluer entre 0 et 1. Une fois que ce normage sera fait, on pourra pondérer les critères en fonction de l'importance qu'on décide de leur

attribuer. La méthode suivante est donc un développement de la méthode O.S faisant apparaître ces améliorations.

III.3-b) Méthode 2 : Optimisation par Normage, Pondération et Sommation (N.P.S)

Nous allons tout d'abord détailler le principe du normage des critères.

Le formalisme utilisé est le suivant :

Soit $f_i(x)$ = critère i et n = nombre de critères.

Soit $[X_j]$, une solution j particulière du problème, avec $j = 1..m$, m = nombre total de solutions.

Soit $f_i(x_j)$, la valeur "brute" du critère i pour une solution x_j .

Soit $\min(f_i)$ et $\max(f_i)$, les valeurs "brutes" minimales et maximales du critère i , avec étendue "brute" de $f_i(x) = \max(f_i) - \min(f_i)$.

Soit $\|f_i(x_j)\|$, la valeur normée (de 0 à 1) du critère i pour la solution $[X_j]$.

Le normage du critère $f_i(x)$ pour la solution $[X_j]$ est défini par :

$$\|f_i(x_j)\| = \frac{f_i(x_j) - \min(f_i)}{\max(f_i) - \min(f_i)}$$

pour un critère à minimiser, ou par :

$$\|f_i(x_j)\| = \frac{\max(f_i) - f_i(x_j)}{\max(f_i) - \min(f_i)}$$

pour un critère à maximiser.

Le principal problème rencontré avec le normage est la détermination des valeurs minimales et maximales des critères. Il existe plusieurs possibilités :

- Avoir une bonne connaissance du problème et donc connaître, à l'avance et par expérience, les valeurs minimales et maximales.
- Lancer une exploration aléatoire sur un nombre limité d'itérations, mais suffisant pour permettre de récolter des valeurs minimales et maximales de chaque critère. L'inconvénient est que lors de l'optimisation par l'algorithme génétique, si une valeur minimale ou maximale est dépassée, la fonction objectif sera faussée car le critère normé prendra une valeur négative.
- Utiliser la valeur minimale et la valeur maximale de chaque critère à chaque génération.

- Réactualiser la valeur minimale et la valeur maximale de chaque critère depuis le premier individu jusqu'à l'individu courant. C'est cette solution que nous allons utiliser pour effectuer le normage car elle ne nécessite pas de connaissance préalable du problème ou d'exploration aléatoire.

Le deuxième complément à la méthode O.S est la pondération des critères, qui est un problème délicat, à tel point qu'elle est évitée par certains chercheurs au profit de l'optimalité de Pareto. Il est vrai que le choix des coefficients appliqués à chaque critère est empirique, principalement quand il y a plus de deux critères. Mais dans une optimisation multicritère, les critères peuvent chacun avoir une importance variable, et la seule façon d'en tenir compte est de les pondérer. La méthode proposée dans ce travail commence par comparer les critères deux à deux dans une matrice de pondération pour finalement aboutir à des coefficients normés à appliquer à chacun des critères, eux mêmes normés, afin que la fonction objectif ait des valeurs comprises entre 0 et 1. Les différentes étapes de la pondération sont les suivantes :

- 1- Comparer les critères deux à deux dans une matrice [critères x critères] en fixant un coefficient allant de 1 à une valeur quelconque. En règle générale, cette valeur est rarement supérieure à 10 car, au delà, l'un des critères a trop d'importance par rapport à l'autre qui devient donc négligeable. Par exemple, on décidera que $f_1(x) = 3 \cdot f_2(x)$ et de façon équivalente, $f_2(x) = 1/3 \cdot f_1(x)$. Dans le cas de notre problème, on définit une première matrice de pondération avec les coefficients relatifs suivants (figure 35) :

		Poids relatifs		
		Délai	Encours	Reg
Délai		1	2	3
Encours		0.5	1	1
Reg		0.3333	1	1

Annotations :
 - Délai = 3 x Reg (pointant vers la cellule 3 dans la ligne Délai, colonne Reg)
 - Reg = (1/3) x Délai (pointant vers la cellule 0.3333 dans la ligne Reg, colonne Délai)

Figure 35 : Matrice de pondération, poids relatifs

L'interprétation de la matrice est la suivante :

- le délai (D) est deux fois plus important que l'encours (EC),
 - le délai (D) est trois fois plus important que le nombre de réglages (Reg),
 - l'encours (EC) est aussi important que le nombre de réglages (Reg).
 - etc...
- 2- Additionner les coefficients des critères par colonne.
 - 3- Diviser chaque cellule de chaque colonne par le résultat de l'addition des coefficients. La figure 36 montre le résultat obtenu dans notre cas.

Poids relatifs			
	Délai	Encours	Reg
Délai	1	2	3
Encours	0.5	1	1
Reg	0.3333	1	1
Somme	1.8333	4	5
	0.5455	0.5	0.6
	0.2727	0.25	0.2
	0.1818	0.25	0.2

Division de 0,5 par 1,8333

Addition de 3 + 1 + 1

Figure 36 : Coefficients corrigés

4- Additionner chaque ligne obtenue

5- Diviser le résultat par le nombre de critères. On obtient les poids normés des critères (la somme des poids normés est toujours égale à 1, de façon à ce que la fonction objectif avec les critères normés soit comprise entre 0 et 1). La figure 37 montre le résultat final et les poids normés obtenus.

Poids relatifs			
	Délai	Encours	Reg
Délai	1	2	3
Encours	0.5	1	1
Reg	0.3333	1	1
Somme	1.8333	4	5
	0.5455	0.5	0.6
	0.2727	0.25	0.2
	0.1818	0.25	0.2
Poids Normés			
	Délai	Encours	Reg
1	0.55	0.24	0.21

Division de (0,1818+0,25+0,2) par 3

Figure 37 : Obtention des poids normés

Pour le cas traité en exemple, la fonction objectif avec cette méthode de Normage, Pondération et Sommation (N.P.S) s'écrit donc sous la forme :

$$F(x) = 0,55.||D(x)|| + 0,24.||EC(x)|| + 0,21.||Reg(x)||$$

Avec :

$$||D(x)|| = (D(x) - \min(D)) / (\max(D) - \min(D))$$

$$\|EC(x)\| = (EC(x) - \min(EC)) / (\max(EC) - \min(EC))$$

$$\|Reg(x)\| = (Reg(x) - \min(Reg)) / (\max(Reg) - \min(Reg))$$

puisque tous les critères sont à minimiser.

Les résultats (valeur courante, mini, moyenne) obtenus pour chacun des critères "bruts" et pour la fonction objectif normée et pondérée au cours des différentes itérations de l'algorithme sont présentés sur les graphiques des figures 38 à 41. La figure 42 montre l'évolution de la valeur minimale de la fonction objectif et des valeurs des critères normés correspondants.

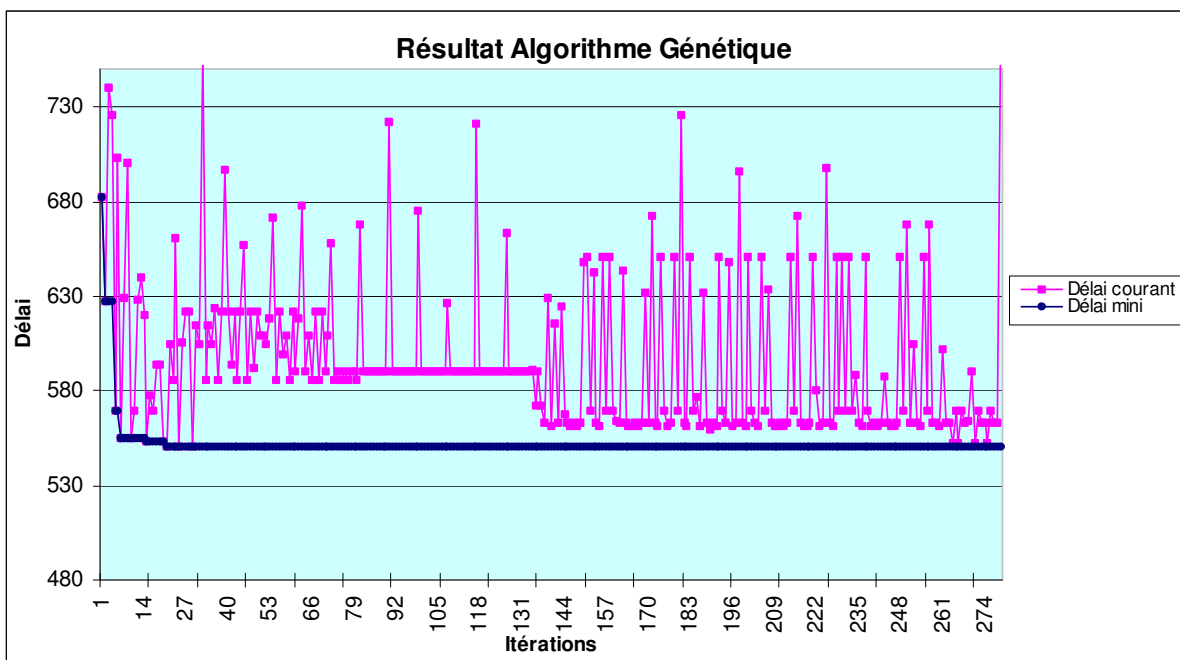


Figure 38 : Délai = f(itérations)

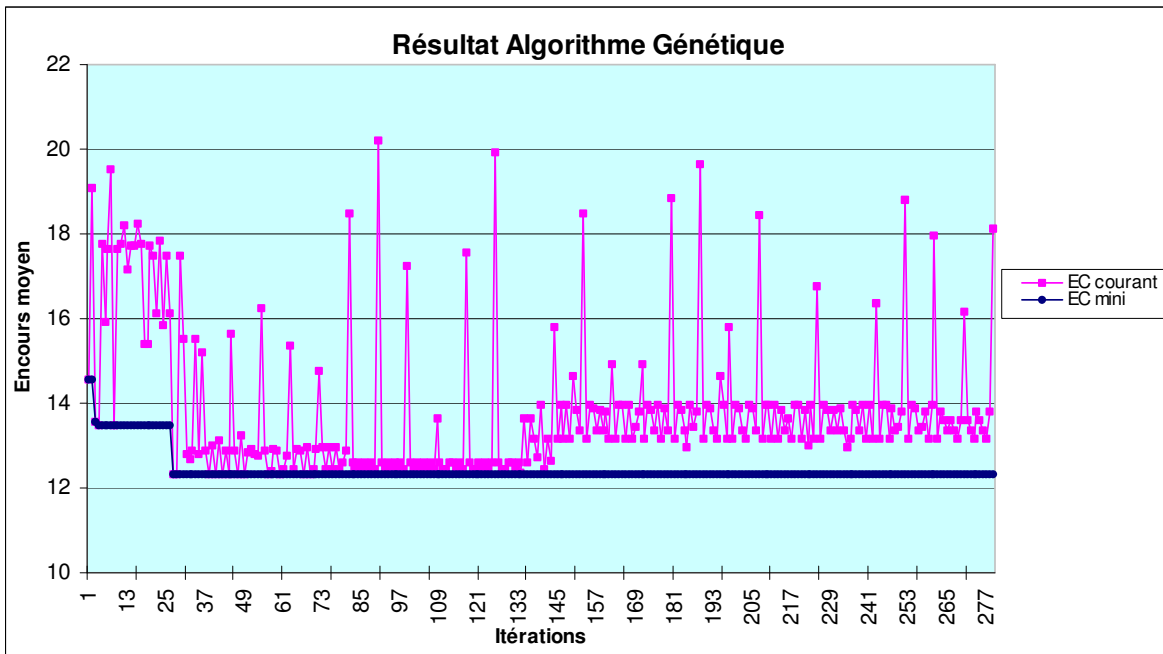


Figure 39 : $Encours = f(itérations)$

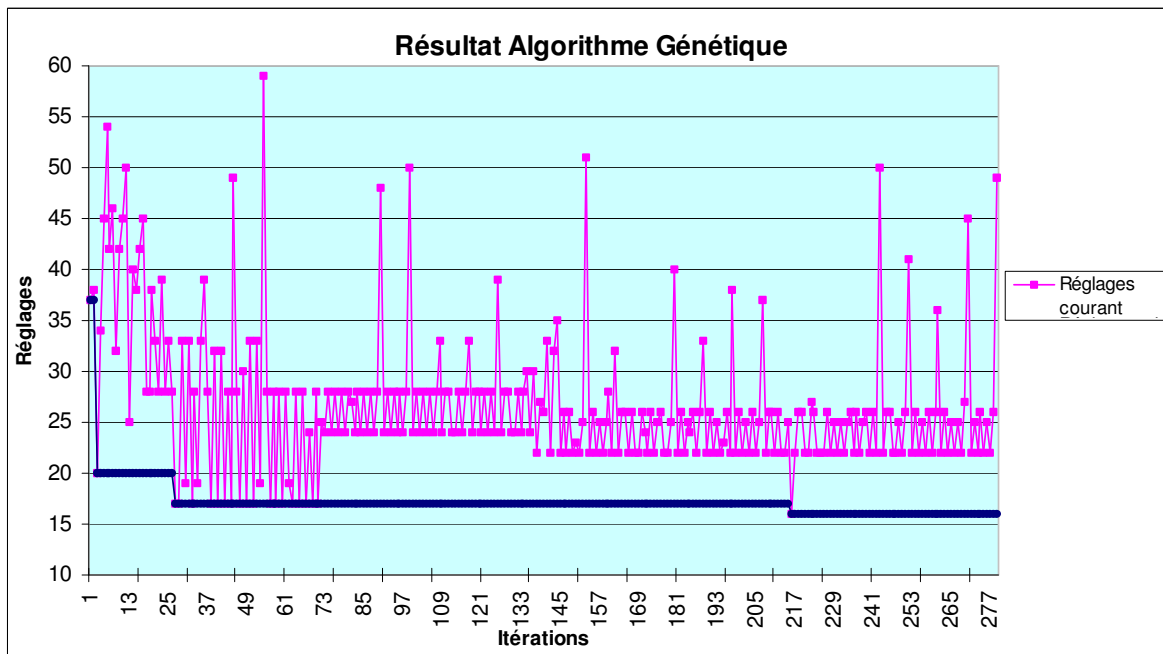


Figure 40 : $Réglages = f(itérations)$

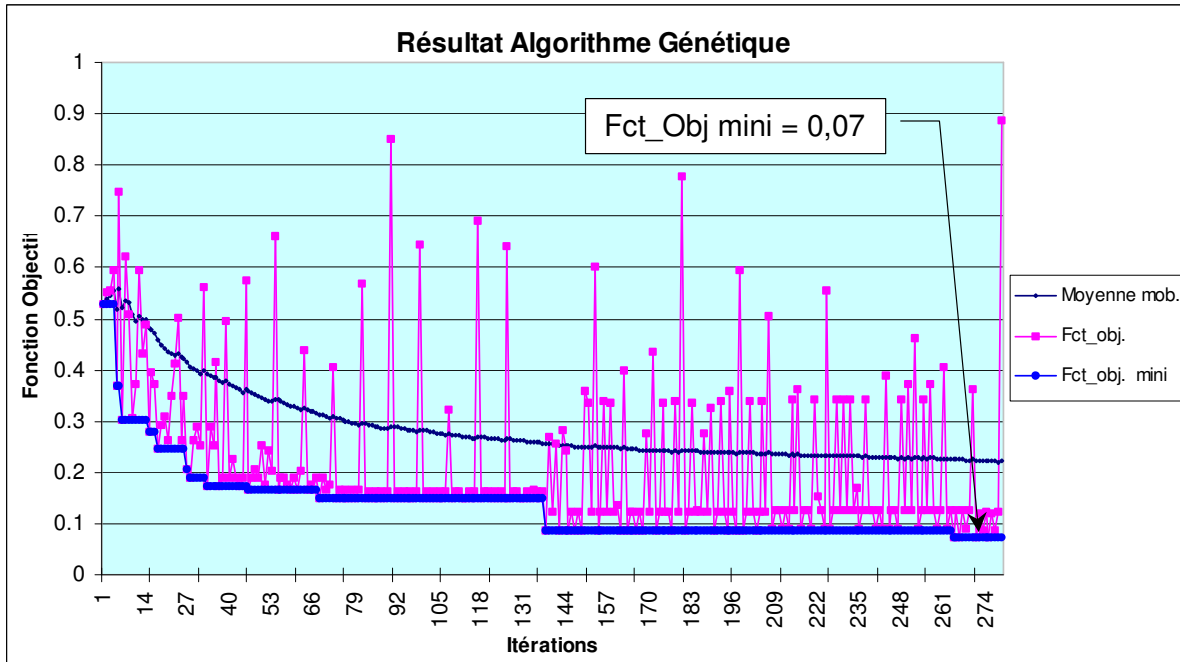


Figure 41 : Fonction objectif = $f(\text{itérations})$

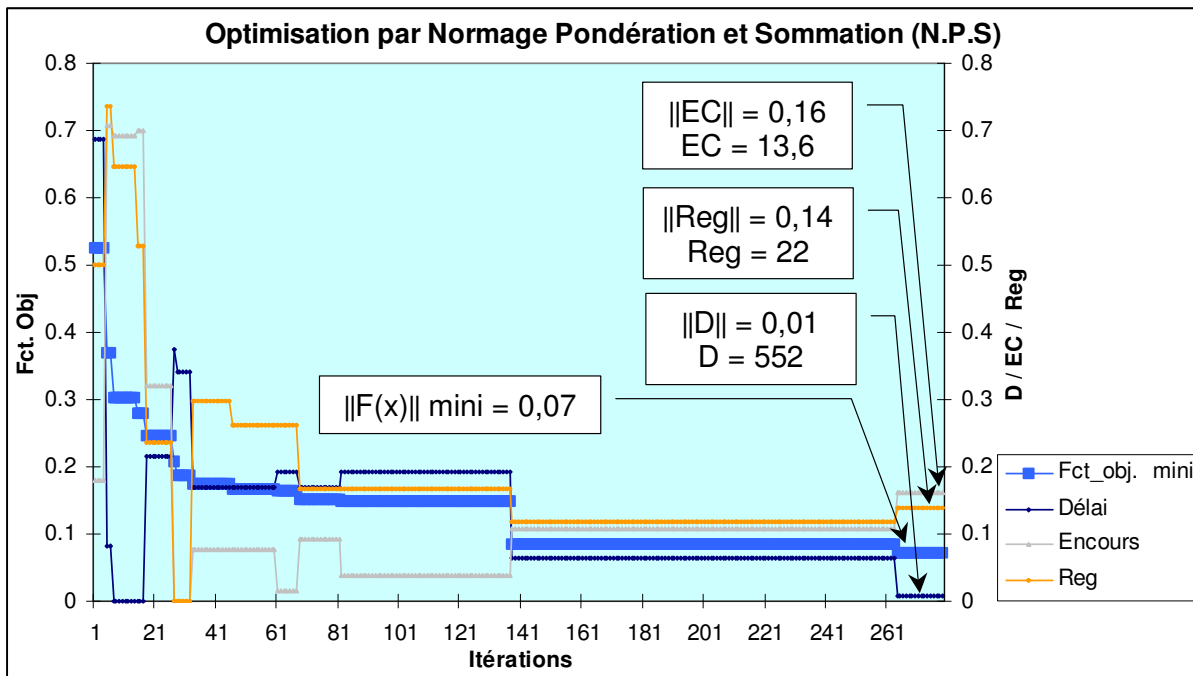


Figure 42 : Evolution de la fonction objectif avec la méthode NPS et des critères normés

Les différents graphiques permettent de visualiser le comportement de l'algorithme d'optimisation avec la méthode N.P.S. On peut notamment vérifier l'importance du délai par rapport aux autres critères, par exemple lors de la 264^{ème} itération où, malgré une dégradation de l'encours et des réglages, la réduction du délai entraîne une réduction de la fonction objectif. La meilleure solution obtenue à l'issue des itérations donne les résultats suivants :

- Délai (D) = 552 secondes,
- Encours moyen (EC) = 13,6 palettes,
- Nombre total de réglages (Reg) = 22 réglages.

Le chromosome correspondant à cette solution est :

6-9-5-9-8 / 6-5-5-4-3-3 / 22

Le tableau de la figure 43 montre les résultats obtenus aux paliers du tracé minimal de la fonction objectif.

Iter	Mini			Chromosome									Valeurs brutes			Mini			Valeurs normées					
	Mini	Moy	F. Obj	TiLa 1 à TiLa 5					StAm 1 à StAm 6				Rp	D	EC	Reg	D	EC	Reg	D	EC	Reg		
1	0.527	0.527	0.527	2	7	11	5	6	4	6	5	4	3	2	24	682	14.6	37	682	14.6	37	0.69	0.18	0.50
5	0.369	0.519	0.369	8	4	7	8	5	5	1	5	5	4	4	30	570	17.8	45	570	13.5	20	0.08	0.71	0.74
7	0.302	0.521	0.302	9	2	1	6	7	2	3	3	3	5	5	29	555	17.7	42	555	13.5	20	0.00	0.69	0.65
15	0.279	0.483	0.279	8	2	1	6	7	2	3	3	3	5	5	29	553	17.7	38	553	13.5	20	0.00	0.70	0.53
18	0.245	0.459	0.245	6	8	5	4	8	4	5	2	2	3	5	25	594	15.4	28	553	13.5	20	0.22	0.32	0.24
27	0.207	0.415	0.207	12	7	12	11	8	5	3	3	6	4	5	20	621	12.3	17	550	12.3	17	0.38	0.00	0.00
28	0.188	0.407	0.188	12	7	12	11	8	5	3	3	6	4	5	20	621	12.3	17	550	12.3	17	0.34	0.00	0.00
33	0.174	0.392	0.174	6	8	5	4	8	4	5	2	6	4	5	20	586	12.9	28	550	12.3	17	0.17	0.08	0.30
46	0.167	0.358	0.167	6	8	5	4	8	4	5	2	6	4	5	20	586	12.9	28	550	12.3	17	0.17	0.08	0.26
61	0.165	0.325	0.165	6	8	5	11	8	5	5	1	6	4	5	20	590	12.4	28	550	12.3	17	0.19	0.02	0.26
68	0.151	0.313	0.151	6	9	5	4	8	4	5	2	6	4	5	20	586	13	24	550	12.3	17	0.17	0.09	0.17
82	0.15	0.295	0.15	6	9	5	11	8	5	5	1	6	4	5	20	590	12.6	24	550	12.3	17	0.19	0.04	0.17
138	0.086	0.256	0.086	6	9	4	11	8	4	5	5	4	3	3	22	563	13.2	22	550	12.3	17	0.06	0.11	0.12
264	0.073	0.226	0.073	6	9	5	9	8	6	5	5	4	3	3	22	552	13.6	22	550	12.3	16	0.01	0.16	0.14

Figure 43 : Résultats aux paliers de la fonction objectif

Les graphiques de la figure 43 représente en deux dimensions les critères normés pour chaque individu de l'algorithme génétique. On voit aussi la position du meilleur individu (obtenu à la 264^{ème} itération) qui n'a dans aucun cas la meilleure valeur du critère, ce qui montre que la méthode N.P.S fournit une solution qui est en fait le meilleur compromis entre les critères pondérés. Sur chaque graphique, on peut aussi identifier une zone de concentration d'individus (en cerclé), dans laquelle se trouve le meilleur individu, que l'on peut considérer comme une zone d'optimalité vers laquelle l'algorithme converge.

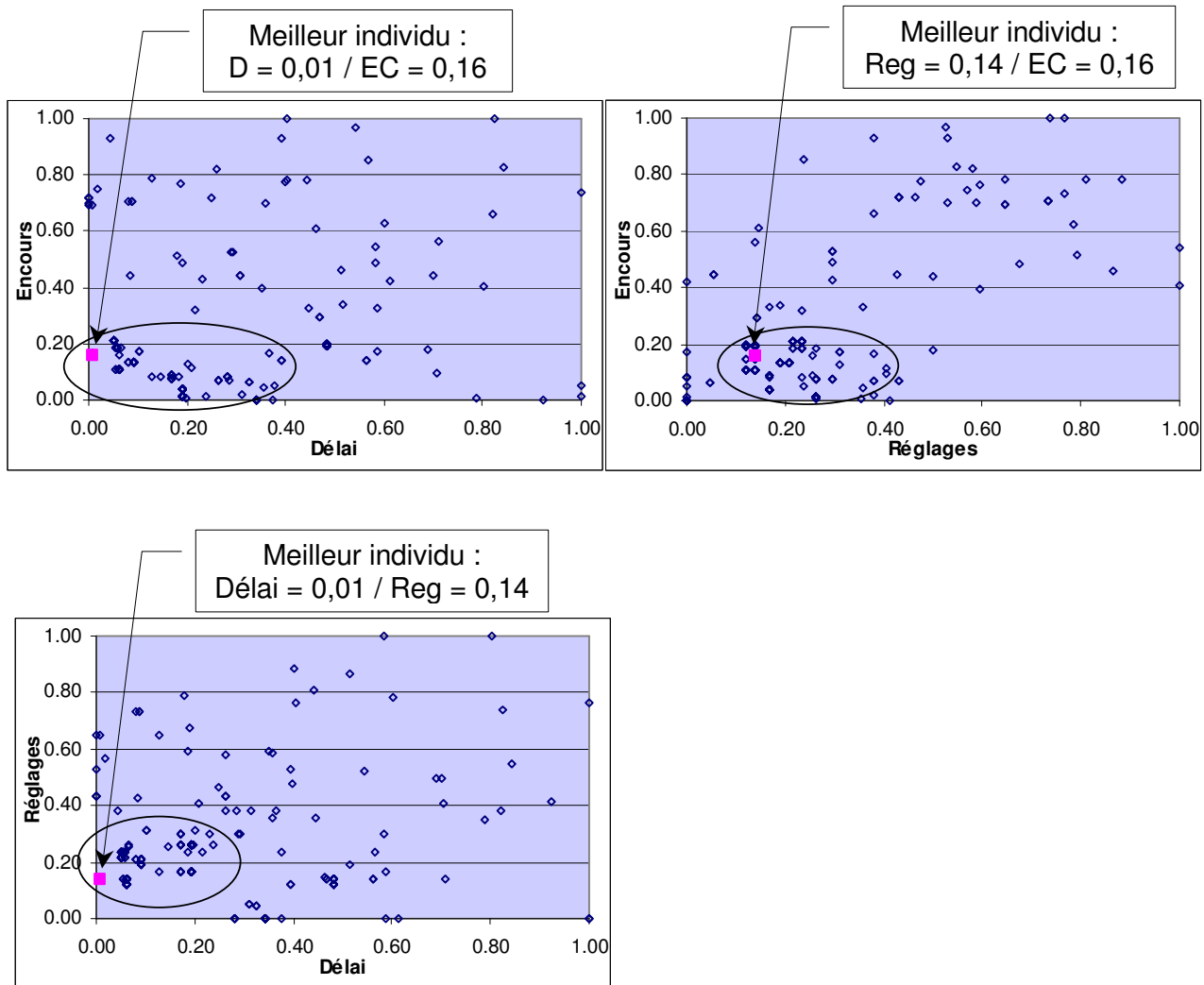


Figure 44 : Représentation des critères en 2 dimensions

Quatre autres expérimentations ont été réalisées avec cette méthode, mais avec d'autres coefficients de pondération (voir les matrices de pondération correspondant à ces expérimentations sur la figure 45). Mis à part pour l'expérimentation N°5 où tous les poids sont à 1 (coefficients à 0,33), le délai (D) est toujours supposé avoir le plus d'importance par rapport à l'encours (EC) et aux réglages (Reg). La synthèse des cinq expérimentations est donnée sous forme d'un tableau (figure 46) où l'on trouve les meilleures solutions. Les valeurs optimales du délai (D), de l'encours moyen (EC) et des réglages (Reg) sont dans les colonnes "valeurs brutes". On peut constater l'influence du poids des critères en comparant les résultats des différentes expérimentations.

Sauf pour l'expérimentation N°1 où le délai a un poids nettement supérieur aux deux autres critères, on peut voir dans les colonnes "mini" que les valeurs minimales obtenues

par chaque critère au cours des itérations ne correspondent pas à la fonction objectif optimale.

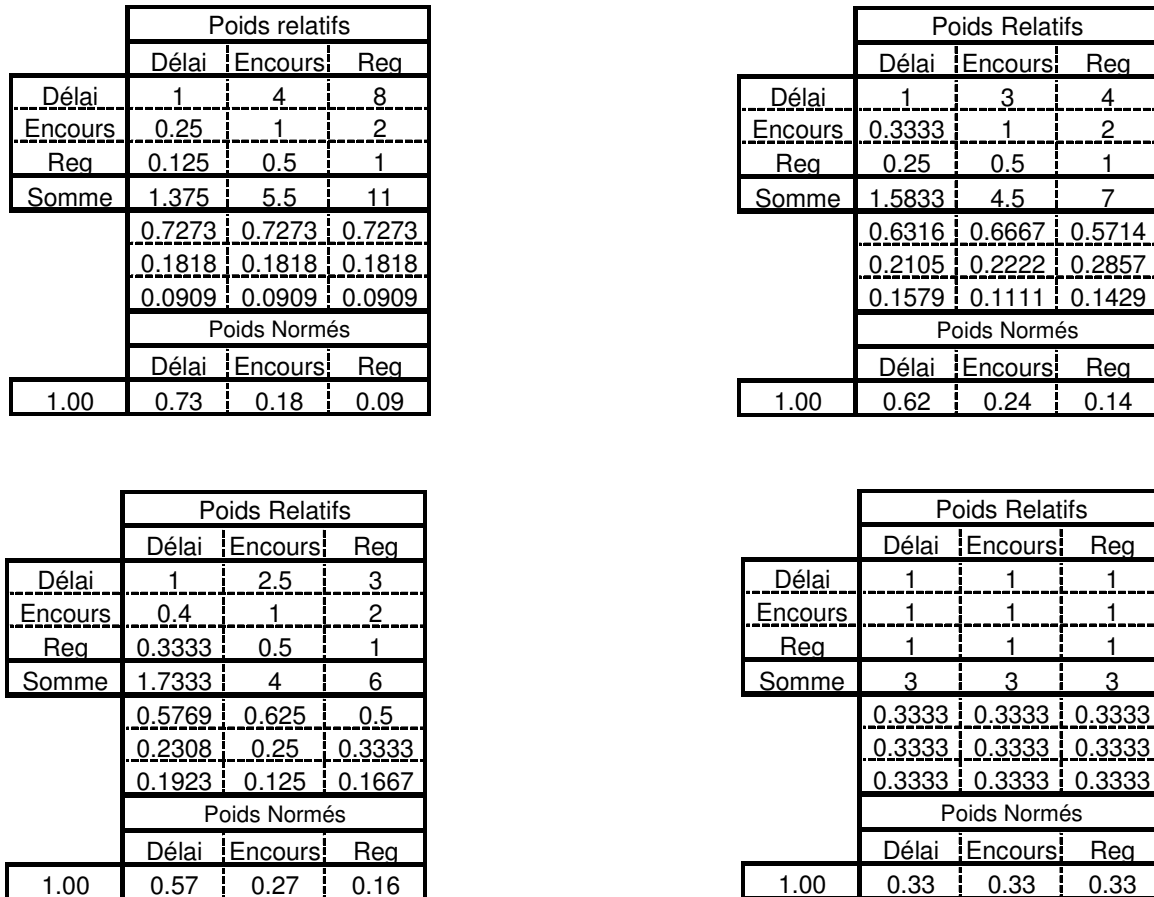


Figure 45 : Matrices de pondération

Exp	Poids normés			Chromosome									Valeurs brutes			Mini			Valeurs normées			F. Obj			
	W.D	W.EC	W.Reg	TiLa 1 à TiLa 5			StAm. 1 à StAm 6			Rp	D	EC	Reg	D	EC	Reg	D	EC	Reg						
1	0.73	0.18	0.09	3	11	1	8	8	3	6	6	6	6	6	34	486	19.5	41	486	12.7	18	0.00	0.92	0.65	0.224
2	0.62	0.24	0.14	6	11	1	8	8	4	2	6	6	5	4	27	525	16.8	32	508	12.7	17	0.06	0.55	0.41	0.216
3	0.57	0.27	0.16	6	10	10	11	8	2	6	1	2	4	5	34	568	14.4	19	526	12.3	16	0.16	0.26	0.07	0.171
4	0.55	0.24	0.21	6	9	5	9	8	6	5	5	4	3	3	22	552	13.6	22	550	12.3	16	0.01	0.16	0.14	0.072
5	0.33	0.33	0.33	8	9	9	12	8	6	3	3	6	5	5	21	587	12.7	16	544	12.1	14	0.16	0.07	0.05	0.224

Figure 46 : Synthèse des expérimentations réalisées avec la méthode N.P.S

L'observation de la zone de densité maximale dans laquelle se trouve le meilleur individu montre que plusieurs solutions intéressantes ont été explorées par l'algorithme. Si le problème le nécessitait, on pourrait donc utiliser n'importe laquelle de ces solutions en sachant qu'elle est proche de la solution optimale.

Contrairement à la méthode N.P.S, la méthode suivante comporte un mécanisme de sélection des solutions qui n'est pas basé sur le résultat d'une combinaison linéaire de plusieurs critères (fonction objectif).

III.3-c) Méthode 3 : Optimisation par Non Domination Pondérée (N.D.P)

Cette méthode s'inspire en partie du principe d'optimalité de Pareto. La première différence avec la méthode originale se situe au niveau de la sélection des individus non-dominés. La deuxième différence réside dans l'introduction de coefficients de pondération, ce que l'optimalité de Pareto ne permet pas. La méthode N.D.P tente donc d'utiliser une partie du concept d'optimalité au sens de Pareto et une partie du principe de pondération vu avec la méthode N.P.S.

Avec la méthode de Pareto, une solution est optimale si elle est non-dominée tous critères confondus (voir le principe de non-domination défini au début du chapitre III.2).

Exemple :

Les trois graphiques de la figure 48 montrent des solutions suivant trois critères normés (F1, F2 et F3). Le tableau de la figure 47 indique les valeurs des critères pour chaque solution et les solutions Pareto-Optimales (marquées d'un astérisque). Pour les autres solutions, les dominants sont indiqués.

	Critère F1	Critère F2	Critère F3	Dominants
Solution 1	5	9	7	2, 5
Solution 2	2	7	5	*
Solution 3	9	8	5	2, 4, 5, 6
Solution 4	8	6	1	*
Solution 5	3	3	2	*
Solution 6	7	1	4	*
Solution 7	10	2	8	6

Figure 47 : Valeurs des critères et dominants

La solution 1 est dominée par la solution 2 car la valeur de chaque critère de la solution 2 est inférieure à la valeur de chaque critère de la solution 1 (pour F1 : $2 < 5$, pour F2, $7 < 9$, pour F3 : $5 < 7$).

La solution 5 est non-dominée car aucune autre solution ne dispose de valeurs de critères inférieures aux valeurs des critères de la solution 5 (le critère F1 de la solution 2 est

inférieur au critère F1 de la solution 5, mais le critère F2 de la solution 2 est supérieur au critère F2 de la solution 5).

Dans cet exemple, il y a donc 4 solutions optimales au sens de Pareto ou non-dominées (en cerclé sur les graphiques de la figure 48).

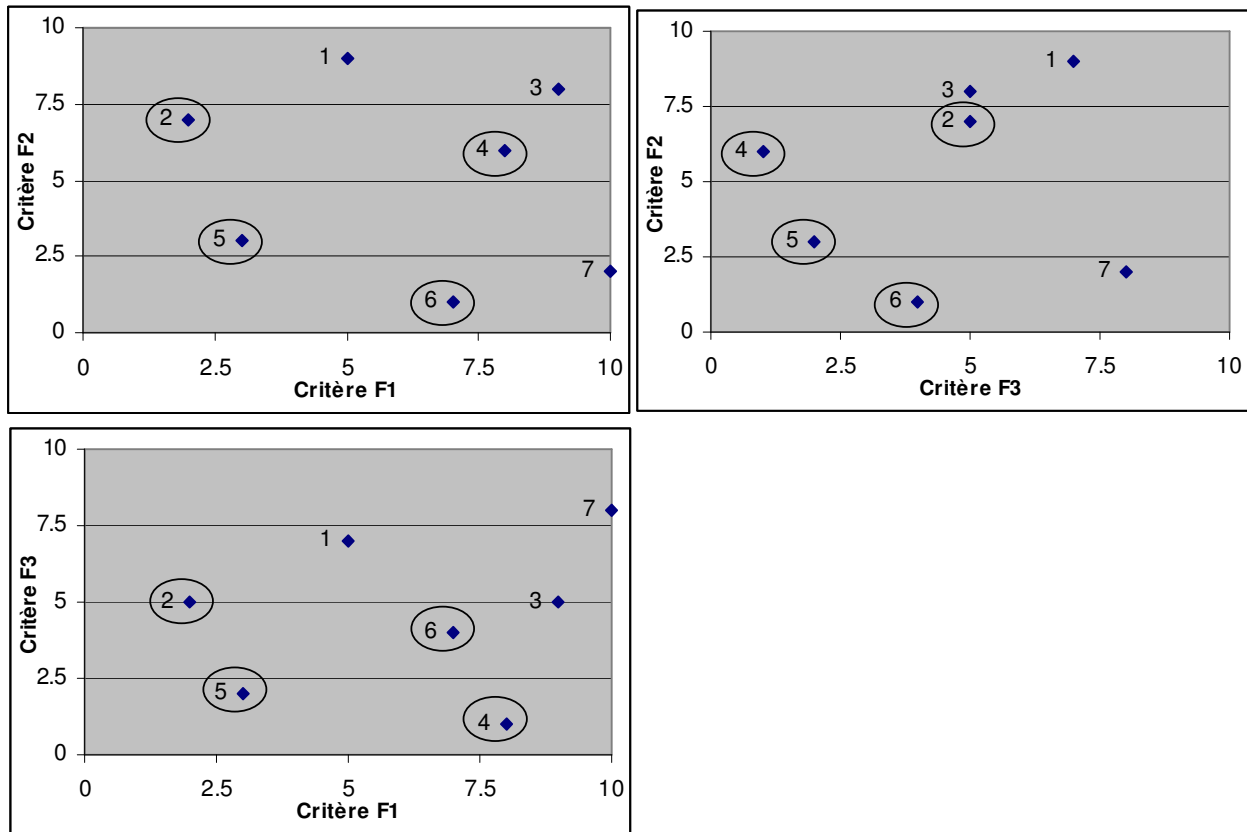


Figure 48 : Représentation graphique des solutions

Le principal inconvénient de la méthode de Pareto est de proposer un ensemble de solutions non-dominées, sans critère de choix entre chacune d'elles, alors que dans notre cas, à chaque génération de l'algorithme génétique, seulement deux solutions sont nécessaires pour effectuer la sélection.

La méthode proposée introduit une variable quantifiable qui permet d'établir un classement des différentes solutions.

Les différentes étapes de la méthode N.D.P sont les suivantes, en supposant que l'on travaille sur des critères normés :

- 1- Pour chaque critère et pour chaque solution, compter le nombre de dominants, c'est-à-dire le nombre de solutions ayant une valeur du critère inférieure à la solution courante.
- 2- Pour chaque critère et pour chaque solution, calculer la somme des écarts (ou des distances) par rapport aux solutions dominantes.
- 3- Pour chaque critère et pour chaque solution, calculer l'écart moyen, c'est-à-dire la somme des écarts par rapport aux solutions dominantes divisée par le nombre de solutions dominantes.
- 4- Pour chaque critère et pour chaque solution, multiplier l'écart moyen par le poids normé.
- 5- Pour chaque critère et pour chaque solution, additionner les écarts pondérés.
- 6- Faire le classement des solutions dans l'ordre croissant des écarts pondérés.

En reprenant les données de l'exemple précédant, l'application des différentes étapes de la méthode N.D.P apparaît dans le tableau de la figure 49 (voir la matrice de pondération sur la figure 50, obtenue d'après le même principe que pour la méthode N.P.S).

Solutions	1	2	3	4	5	6	7
Nb. Dominants / F1	2	0	5	4	1	3	6
Nb. Dominants / F2	6	4	5	3	2	0	1
Nb. Dominants / F3	5	3	3	0	1	2	6
Total Dominants	13	7	13	7	4	5	13
Σ Ecarts / Dominants (F1)	5	0	20	15	1	11	26
Σ Ecarts / Dominants (F2)	27	16	21	12	3	0	1
Σ Ecarts / Dominants (F3)	18	8	8	0	1	5	24
Ecart Moy. / Dominants (F1)	2.50	0.00	4.00	3.75	1.00	3.67	4.33
Ecart Moy. / Dominants (F2)	4.50	4.00	4.20	4.00	1.50	0.00	0.00
Ecart Moy. / Dominants (F3)	3.60	2.67	2.67	0.00	1.00	2.50	4.00
Ec. Moy. / D(F1)*poids F1	1.37	0.00	2.19	2.06	0.55	2.01	2.38
Ec. Moy. / D(F2)*poids F2	1.08	0.96	1.01	0.96	0.36	0.00	0.00
Ec. Moy. / D(F3)*poids F3	0.76	0.56	0.56	0.00	0.21	0.53	0.84
Σ (Ec. Moy. * poids)	3.21	1.53	3.77	3.02	1.12	2.54	3.22
Classt. Ec. Moy. pondérés	5	2	7	4	1	3	6

Figure 49 : Etapes de calcul de la méthode N.D.P

Poids Relatifs			
	F1	F2	F3
F1	1	2	3
F2	0.5	1	1
F3	0.333	1	1
total colonne	1.833	4	5
	0.545	0.5	0.6
	0.273	0.25	0.2
	0.182	0.25	0.2
Poids Normés			
	F1	F2	F3
	0.55	0.24	0.21

Figure 50 : Matrice de pondération

Avec cette méthode et l'algorithme génétique utilisé jusqu'à présent, la sélection des individus se fait en fonction du classement des écarts moyens pondérés. Le premier parent est systématiquement l'individu qui obtient le meilleur classement. Avec l'exemple précédent, c'est donc l'individu qui correspond à la solution 5 qui est sélectionné. Pour sélectionner le second parent, on applique en plus le critère de consanguinité défini dans notre algorithme.

Comme on le voit, la méthode N.D.P est un mélange d'optimalité de Pareto, mais critère par critère, et de pondération des critères, sauf que ce ne sont pas les valeurs des critères qui sont pondérées, mais les écarts moyens par rapport aux solutions dominantes.

Afin de vérifier l'efficacité de cette méthode, nous allons l'appliquer sur les cinq expérimentations réalisées avec la méthode N.P.S., en détaillant l'expérimentation N°4, afin de comparer les résultats à ceux du chapitre III.3-b.

Les résultats (valeur courante, mini, moyenne) obtenus pour chacun des critères "bruts" au cours des différentes itérations de l'algorithme sont présentés sur les graphiques des figures 51 à 53. La figure 54 montre l'évolution de la valeur minimale des critères normés correspondant aux individus sélectionnés pour la génération suivante. Les graphiques de la figure 55 représentent en deux dimensions les critères normés pour chaque individu de l'algorithme génétique. On voit aussi la position du meilleur individu de la dernière génération.

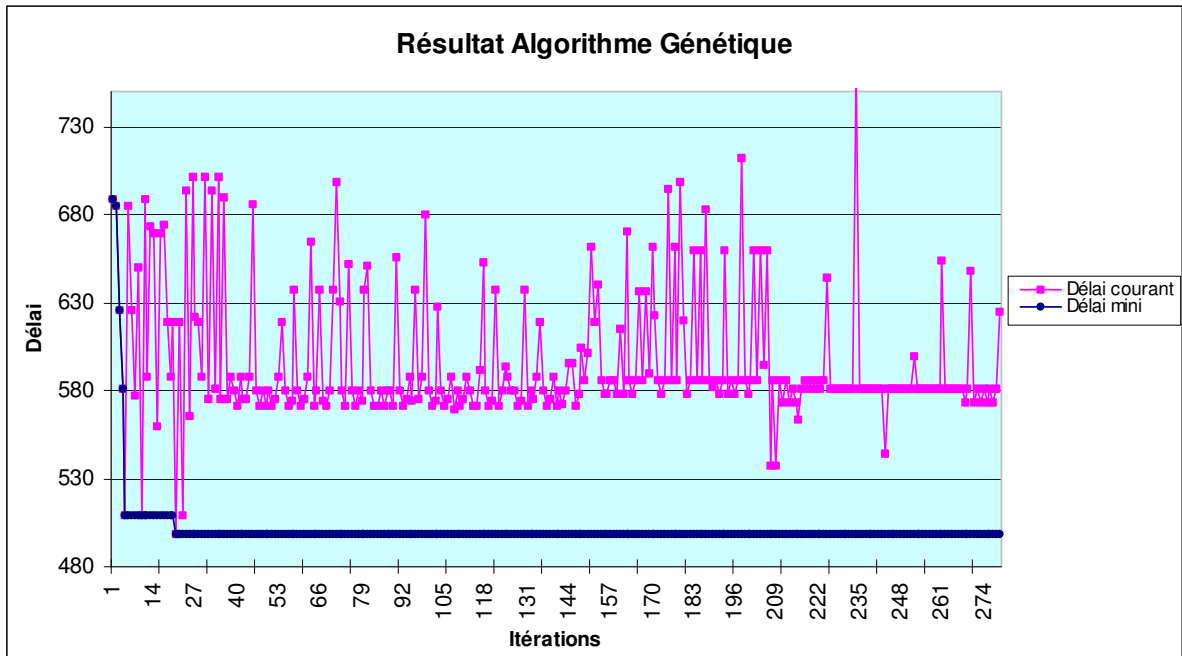


Figure 51 : Délai = $f(\text{itérations})$

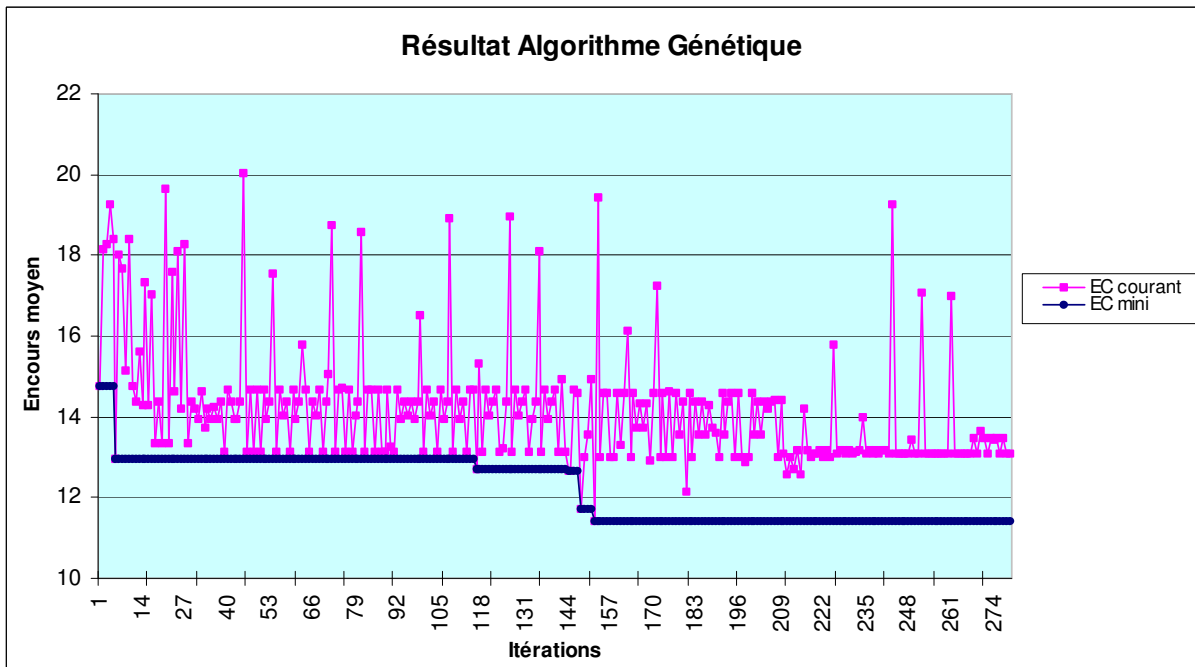


Figure 52 : Encours moyen = $f(\text{itérations})$

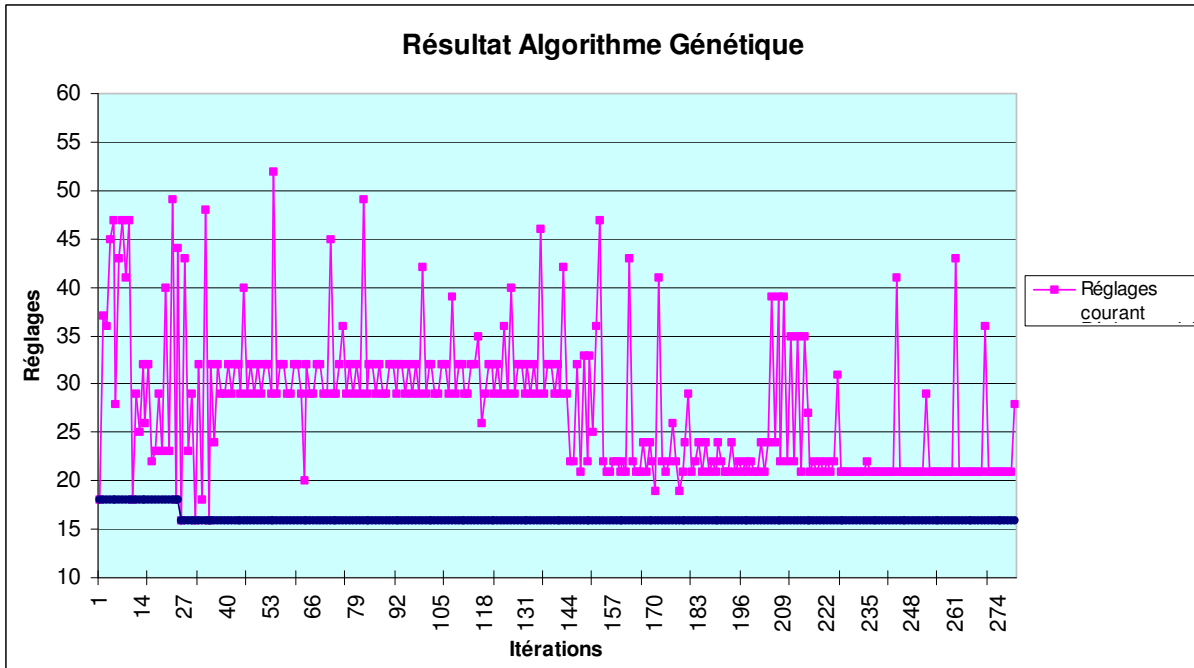


Figure 53 : Réglages = $f(\text{itérations})$

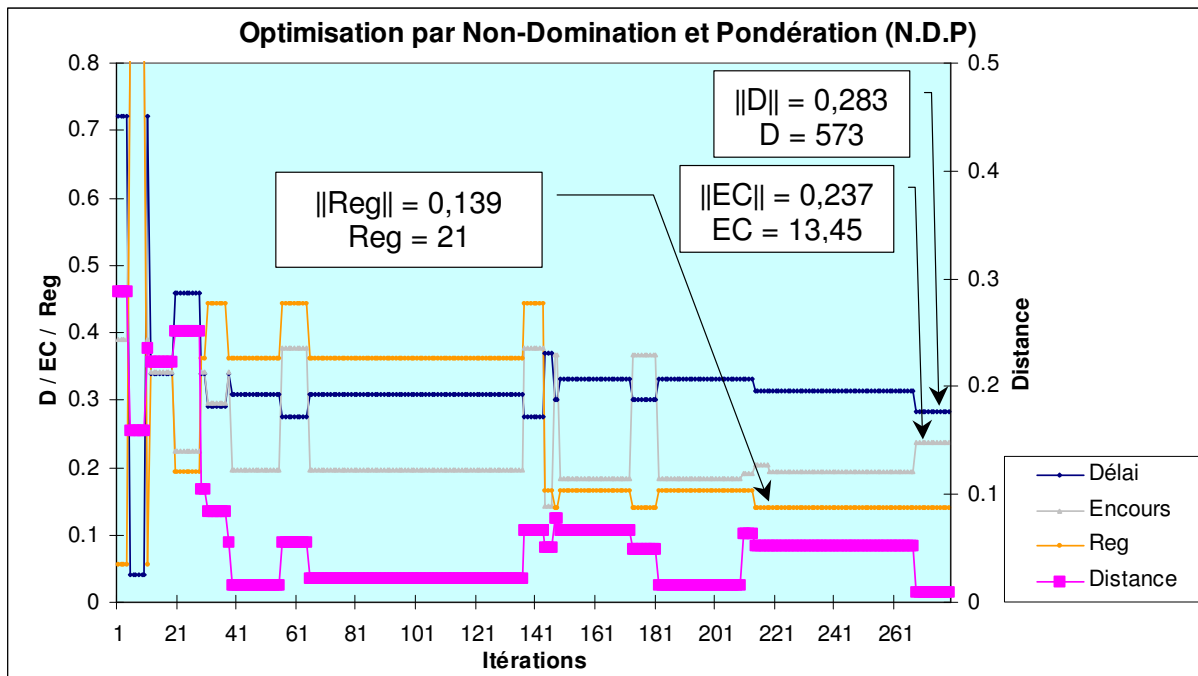


Figure 54 : valeur minimale des critères normés des individus sélectionnés

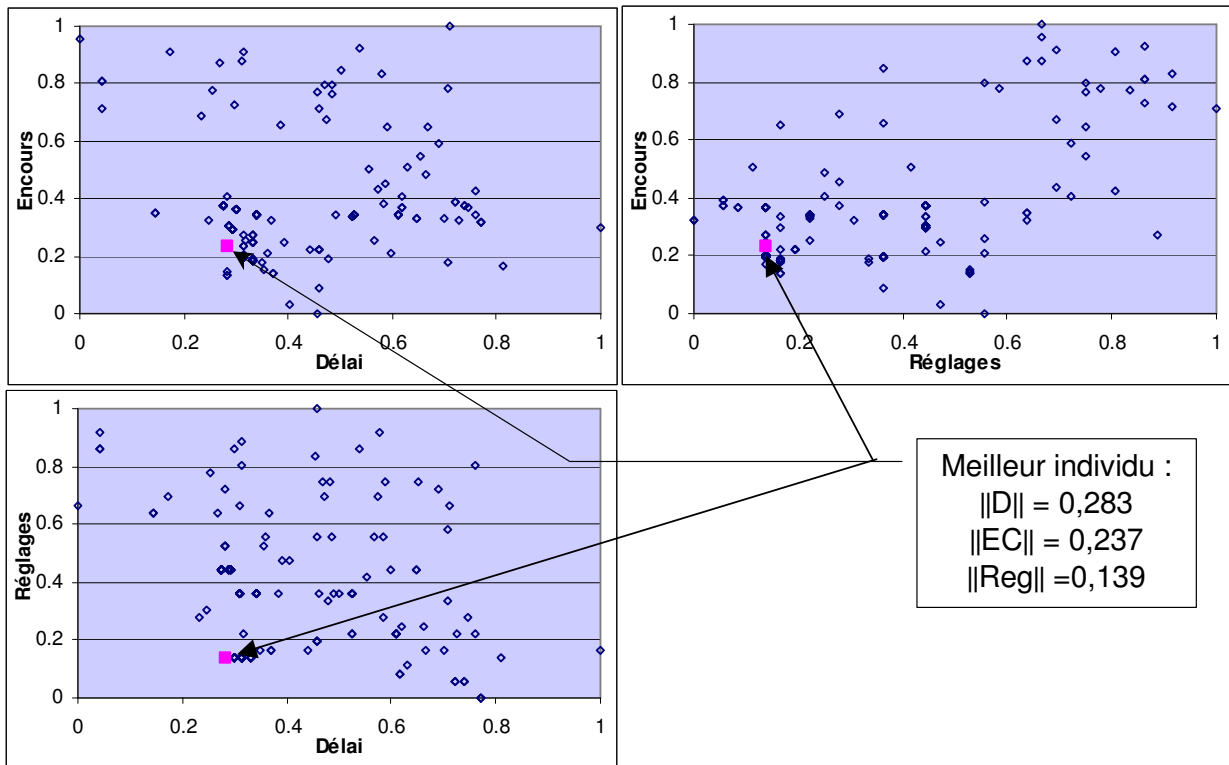


Figure 55 : Représentation des critères en 2 dimensions

La meilleure solution obtenue à l'issue des itérations donne les résultats suivants :

- Délai (D) = 573 secondes,
- Encours moyen (EC) = 13,4 palettes,
- Nombre total de réglages (Reg) = 21 réglages.

Le chromosome correspondant à cette solution est :

9-4-12-11-8 / 6-5-4-4-3-5 / 23

Le tableau de la figure 56 montre les résultats obtenus aux paliers du tracé minimal de la distance (moyenne de la somme des écarts pondérés), en relation avec le graphique de la figure 54.

Les colonnes "Chromosome" donne les valeurs des variables de pilotage (TiLa, StAm et Rp). Les colonnes "Valeurs brutes" indiquent les valeurs non normées des critères (Délai (D), Encours moyen (EC) et Réglages (Reg)). Les colonnes "Mini" donnent les valeurs minimales de chaque critère au cours des itérations. Les colonnes "Valeurs normées" correspondent au normage des critères en fonction des valeurs minimales et maximales

rencontrées depuis la première itération. Comme on l'a déjà vu, une autre possibilité de calcul consiste à prendre les valeurs minimales et maximales des critères de la génération courante uniquement. Comme pour la méthode N.P.S, les meilleurs résultats ont été obtenus avec les valeurs globales, depuis la première itération. Les colonnes "Dominants" correspondent au nombre de dominants de la solution suivant chaque critère. La colonne "Dist" est la distance moyenne (ou écart moyen) des écarts pondérés par rapport aux dominants. L'algorithme génétique sélectionne à chaque génération les individus qui possèdent la distance la plus faible.

Iter	Chromosome												Valeurs brutes			Mini			Valeurs normées			Dominants			Dist.
	TiLa 1 à TiLa 5					StAm 1 à StAm 6					Rp	D	EC	Reg	D	EC	Reg	D	EC	Reg	D	EC	Reg		
1	11	10	5	6	3	3	6	5	4	5	4	23	689	14.8	18	498	11.4	16	1.000	0.289	0.000	8	1	0	0.288
5	1	9	6	9	5	3	2	6	5	2	6	31	509	18.4	47	509	14.8	18	0.000	0.866	1.000	0	7	7	0.159
11	11	10	5	6	3	3	6	5	4	5	4	23	689	14.8	18	509	13	18	1.000	0.289	0.000	8	4	0	0.236
12	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	509	13	18	0.439	0.224	0.379	2	3	5	0.224
20	12	10	8	12	8	3	2	5	5	3	6	34	619	13.3	23	509	13	18	0.595	0.058	0.212	4	0	2	0.252
29	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	498	13	16	0.441	0.211	0.394	3	7	5	0.105
31	1	9	5	12	8	3	2	5	5	3	4	23	575	14	32	498	13	16	0.378	0.150	0.485	0	2	6	0.085
38	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	498	13	16	0.441	0.199	0.394	5	4	0	0.055
39	1	9	5	12	8	3	6	5	4	5	4	23	580	13.1	29	498	13	16	0.402	0.022	0.394	4	0	0	0.016
56	1	9	5	6	3	3	2	5	4	4	4	23	571	14.7	32	498	13	16	0.357	0.240	0.444	0	6	5	0.056
65	1	9	5	12	8	3	6	5	5	4	4	23	580	13.1	29	498	13	16	0.402	0.022	0.361	3	0	0	0.022
137	1	9	5	6	3	3	2	5	5	3	4	23	571	14.7	32	498	12.7	16	0.357	0.273	0.444	0	6	5	0.067
144	9	5	12	12	8	6	5	4	3	5	6	21	596	12.6	22	498	12.6	16	0.480	0.000	0.167	8	0	0	0.05
147	9	5	12	6	3	3	2	4	5	3	4	23	578	14.6	21	498	12.6	16	0.391	0.366	0.139	1	5	0	0.077
149	9	5	12	12	8	6	5	4	3	3	4	23	586	13	22	498	11.7	16	0.431	0.184	0.167	2	3	1	0.067
173	9	5	12	6	3	3	2	4	5	3	4	23	578	14.6	21	498	11.4	16	0.391	0.366	0.139	0	6	1	0.05
182	9	5	12	12	8	6	5	4	5	3	4	23	586	13	22	498	11.4	16	0.431	0.184	0.167	3	0	4	0.017
214	9	4	12	12	6	6	6	5	4	3	5	26	581	13.2	21	498	11.4	16	0.386	0.203	0.139	5	6	0	0.053
268	9	4	12	11	8	6	5	4	4	3	5	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	7	0	0.01

Figure 56 : Résultats aux paliers de la distance

Afin de vérifier et de comprendre le fonctionnement de la méthode N.D.P, les résultats des cinq premières génération de l'algorithme sont détaillés sur le tableau de la figure 57 et les deux dernières générations sur la figure 58. Les flèches indiquent les deux individus sélectionnés pour la génération suivante.

Iter	Chromosome										Valeurs brutes			Mini			Valeurs normées			Dominants			Dist.		
	TiLa_1 à TiLa_5					StAm_1 à StAm_6					Rp	D	EC	Reg	D	EC	Reg	D	EC	Reg	D	EC		Reg	
1	11	10	5	6	3	3	6	5	4	5	4	23	689	14.8	18	689	14.8	18	1.000	0.289	0.000	8	1	0	0.288
2	4	12	1	9	12	5	6	4	3	2	1	30	685	18.2	37	685	14.8	18	0.976	0.827	0.655	6	5	3	0.438
3	3	7	4	12	1	3	4	4	4	3	6	35	626	18.3	36	626	14.8	18	0.651	0.845	0.621	4	6	2	0.338
4	3	3	5	4	4	4	5	2	4	6	3	33	581	19.3	45	581	14.8	18	0.401	1.000	0.931	2	8	6	0.296
5	1	9	6	9	5	3	2	6	5	2	6	31	509	18.4	47	509	14.8	18	0.000	0.866	1.000	0	7	7	0.159
6	4	2	9	7	4	3	6	6	3	6	5	21	685	13	28	509	13	18	0.979	0.000	0.345	7	0	1	0.309
7	4	8	12	5	2	1	1	3	4	4	2	31	626	18	43	509	13	18	0.649	0.801	0.862	3	4	5	0.403
8	2	12	11	7	7	5	1	5	5	1	6	31	577	17.7	47	509	13	18	0.378	0.751	1.000	1	3	7	0.421
9	7	5	5	2	10	6	1	6	3	3	3	23	650	15.2	41	509	13	18	0.782	0.350	0.793	5	2	4	0.332
10	1	9	6	9	5	3	2	6	5	2	6	31	509	18.4	47	509	13	18	0.000	0.866	1.000	0	8	8	0.278
11	11	10	5	6	3	3	6	5	4	5	4	23	689	14.8	18	509	13	18	1.000	0.289	0.000	8	4	0	0.236
12	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	509	13	18	0.439	0.224	0.379	2	3	5	0.224
13	11	10	6	9	5	3	2	6	5	2	6	31	673	15.6	25	509	13	18	0.914	0.422	0.241	6	5	3	0.299
14	1	9	6	9	5	3	2	6	4	5	4	23	669	14.3	32	509	13	18	0.892	0.213	0.483	4	1	6	0.403
15	11	10	4	6	3	3	6	5	5	2	6	31	560	17.3	26	509	13	18	0.281	0.697	0.276	1	7	4	0.28
16	1	9	6	9	5	3	2	6	4	5	4	23	669	14.3	32	509	13	18	0.892	0.213	0.483	4	1	6	0.403
17	11	10	5	6	3	3	6	5	5	2	6	31	674	17	22	509	13	18	0.919	0.646	0.138	7	6	1	0.316
18	12	10	8	12	8	3	2	5	5	3	6	34	619	13.3	23	509	13	18	0.613	0.061	0.172	3	0	2	0.227
19	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	509	13	18	0.441	0.211	0.394	3	3	4	0.257
20	12	10	8	12	8	3	2	5	5	3	6	34	619	13.3	23	509	13	18	0.595	0.058	0.212	4	0	2	0.252
21	1	9	5	6	3	3	6	5	4	5	6	34	498	19.6	40	498	13	18	0.000	1.000	0.727	0	8	5	0.265
22	12	10	8	12	8	3	2	5	5	3	4	23	619	13.3	23	498	13	18	0.595	0.058	0.212	4	0	2	0.252
23	1	10	8	11	8	3	2	5	5	3	6	34	509	17.6	49	498	13	18	0.055	0.691	1.000	1	5	8	0.283
24	12	9	5	6	3	3	6	5	4	5	4	23	693	14.6	18	498	13	18	0.959	0.252	0.061	7	4	1	0.364
25	1	9	5	12	8	3	2	5	5	3	6	34	565	18.1	44	498	13	18	0.328	0.772	0.849	2	6	7	0.398
26	12	10	8	6	3	3	6	5	4	5	4	23	702	14.2	16	498	13	16	1.000	0.184	0.000	8	2	0	0.334
27	1	6	9	1	2	4	6	3	3	6	5	31	622	18.3	43	498	13	16	0.610	0.796	0.818	6	7	6	0.381
28	12	10	8	12	8	3	2	5	5	3	4	23	619	13.3	23	498	13	16	0.595	0.058	0.212	4	0	3	0.147
29	1	9	5	6	3	3	6	5	4	5	4	23	588	14.4	29	498	13	16	0.441	0.211	0.394	3	7	5	0.105
30	12	10	8	6	3	3	6	5	4	5	4	23	702	14.2	16	498	13	16	1.000	0.184	0.000	7	4	0	0.243
31	1	9	5	12	8	3	2	5	5	3	4	23	575	14	32	498	13	16	0.378	0.150	0.485	0	2	6	0.085
32	12	9	5	6	3	3	6	5	4	5	4	23	693	14.6	18	498	13	16	0.959	0.252	0.061	6	8	2	0.275
33	1	10	8	12	8	3	2	5	5	3	4	23	581	13.8	48	498	13	16	0.408	0.119	0.970	2	1	8	0.186
34	12	10	8	6	3	3	6	5	4	5	4	23	702	14.2	16	498	13	16	1.000	0.184	0.000	7	4	0	0.243
35	1	9	5	12	8	3	2	5	5	3	4	23	575	14	32	498	13	16	0.378	0.150	0.485	0	2	6	0.085
36	7	8	12	10	1	2	6	5	4	3	4	33	690	14.2	24	498	13	16	0.943	0.191	0.242	5	6	4	0.325

Figure 57 : Résultats détaillés des cinq premières générations

Iter	Chromosome											Valeurs brutes			Mini			Valeurs normées			Dominants			Dist.	
	TiLa 1 à TiLa 5					StAm 1 à StAm 6						Rp	D	EC	Reg	D	EC	Reg	D	EC	Reg	D	EC		Reg
262	9	4	12	12	8	6	5	5	5	3	4	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
263	9	4	12	12	8	5	5	4	4	3	5	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
264	9	4	12	12	8	6	5	4	4	3	5	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
265	9	4	12	12	8	5	5	5	5	3	4	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
266	9	4	12	12	8	6	5	5	4	3	5	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
267	9	4	12	12	8	5	5	4	5	3	4	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
268	9	4	12	11	8	6	5	4	4	3	5	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	7	0	0.01
269	9	4	12	12	8	5	5	5	5	3	4	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	1	0	0	0.016
270	3	7	10	7	2	6	4	3	5	3	3	21	648	13.6	36	498	11.4	16	0.566	0.258	0.556	8	8	8	0.242
271	9	4	12	11	8	6	5	4	4	3	5	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	5	0	0.008
272	9	4	12	12	8	5	5	5	5	3	4	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	4	1	0	0.017
273	9	4	12	11	8	5	5	5	5	3	4	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	5	0	0.008
274	9	4	12	12	8	6	4	4	4	3	5	23	581	13.4	21	498	11.4	16	0.313	0.234	0.139	4	4	0	0.026
275	9	4	12	11	8	5	5	5	5	3	4	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	5	0	0.008
276	9	4	12	12	8	6	5	4	4	3	5	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	4	1	0	0.017
277	9	4	12	11	8	6	5	4	4	3	4	23	573	13.5	21	498	11.4	16	0.283	0.237	0.139	0	5	0	0.008
278	9	4	12	12	8	5	5	5	5	3	5	23	581	13.1	21	498	11.4	16	0.313	0.195	0.139	4	1	0	0.017
279	9	4	7	11	8	5	2	2	6	4	4	20	625	13.1	28	498	11.4	16	0.478	0.192	0.333	8	0	8	0.14

Figure 58 : Résultats détaillés des deux dernières générations

La synthèse des cinq expérimentations réalisées avec cette méthode est donnée sous forme d'un tableau (figure 59) où l'on trouve les meilleures solutions. Les valeurs optimales du délai (D), de l'encours moyen (EC) et des réglages (Reg) sont dans les colonnes "valeurs brutes". Ces résultats sont à comparer avec ceux obtenus avec la méthode N.P.S (voir figure 46, chapitre III.3-b).

Exp	Poids normés			Chromosome								Valeurs brutes			Mini			Valeurs normées			Dist.				
	W D	W EC	W Reg	TiLa 1 à TiLa 5					StAm 1 à StAm 6			Rp	D	EC	Reg	D	EC	Reg	D	EC		Reg			
1	0.73	0.18	0.09	8	4	1	6	10	3	1	3	6	1	4	32	503	19.11	36	503	12.32	17	0.00	0.91	0.49	0.027
2	0.62	0.24	0.14	1	1	6	10	9	3	6	2	5	4	5	26	537	13.95	35	487	12.66	18	0.18	0.19	0.40	0.025
3	0.57	0.27	0.16	6	11	10	2	9	3	4	3	5	5	5	21	583	13.64	23	487	12.16	18	0.35	0.20	0.13	0.003
4	0.55	0.24	0.21	9	4	12	11	8	6	5	4	4	3	5	23	573	13.45	21	498	11.41	16	0.28	0.24	0.14	0.008
5	0.33	0.33	0.33	11	10	9	11	8	2	4	1	5	1	4	35	602	13.09	16	509	12.95	14	0.33	0.05	0.05	0.017

Figure 59 : Synthèse des expérimentations réalisées avec la méthode N.D.P.

III.3-d) Conclusions sur les méthodes de sélection testées

Au vu des résultats, Il est très difficile de dire si une méthode est supérieure à l'autre. La méthode N.P.S donne systématiquement un délai inférieur à la méthode N.D.P., mais l'encours moyen et les réglages sont le plus souvent inférieurs avec la méthode N.D.P qu'avec la méthode N.P.S. On peut simplement conclure que la nouvelle méthode proposée semble être efficace puisqu'elle converge vers des solutions intéressantes. D'autres expérimentations et une analyse approfondie de cette méthode seraient utiles afin de la valider définitivement.

IV- Proposition d'intégration des outils informatiques en gestion de production

La méthodologie suivie dans ce travail conduit à une réflexion sur l'intégration des outils informatiques utilisés. C'est surtout la simulation de flux qui a été utilisée dans ce travail, sous un aspect encore peu répandu de couplage avec un autre outil informatique destiné à l'optimisation. Mais le gestionnaire chargé du pilotage d'une unité de production dispose aujourd'hui de nombreux autres outils informatiques qui sont censés lui faciliter la tâche. Chacun de ces outils utilise des données d'entrée et fournit à l'utilisateur des résultats en sortie. La principale difficulté réside dans l'intégration de ces outils, au sens de la communication et du processus d'utilisation. Dans ce chapitre, une analyse séparée des principaux outils informatiques va être présentée, sous l'aspect des données d'entrée utilisées et des données de sortie. Dans un second temps, nous présenterons un modèle d'intégration de ces outils destiné à améliorer la réactivité et la maîtrise de la production. L'idée générale est de définir un "schéma d'assemblage" de ces outils sans revenir sur les outils eux-mêmes. Chacun d'eux sera vu uniquement comme une boîte noire assurant une certaine fonction.

□ IV.1- Analyse fonctionnelle des principaux outils informatiques

IV.1-a) Gestion de Production Assistée par Ordinateur (G.P.A.O) :

La G.P.A.O a pour fonction de transformer les commandes fermes émanant des clients ou les prévisions en ordres de fabrication et d'approvisionnement. Les ordres de fabrication sont "bruts", c'est-à-dire qu'ils se présentent sous la forme d'un ensemble d'ordres non ordonnés à réaliser pour une date de livraison donnée et sont destinés à l'ordonnancement ou à l'optimisation. La G.P.A.O peut utiliser des informations provenant de l'outil de Suivi de Production afin de réactualiser sa base de données, notamment en

ce qui concerne les temps. Généralement, les outils de G.P.A.O travaillent à capacité infinie, c'est-à-dire qu'ils ne tiennent pas compte de la capacité réelle de l'atelier pour établir le planning de production. La figure 60 présente le bloc fonctionnel point de vue données d'entrée/sortie.

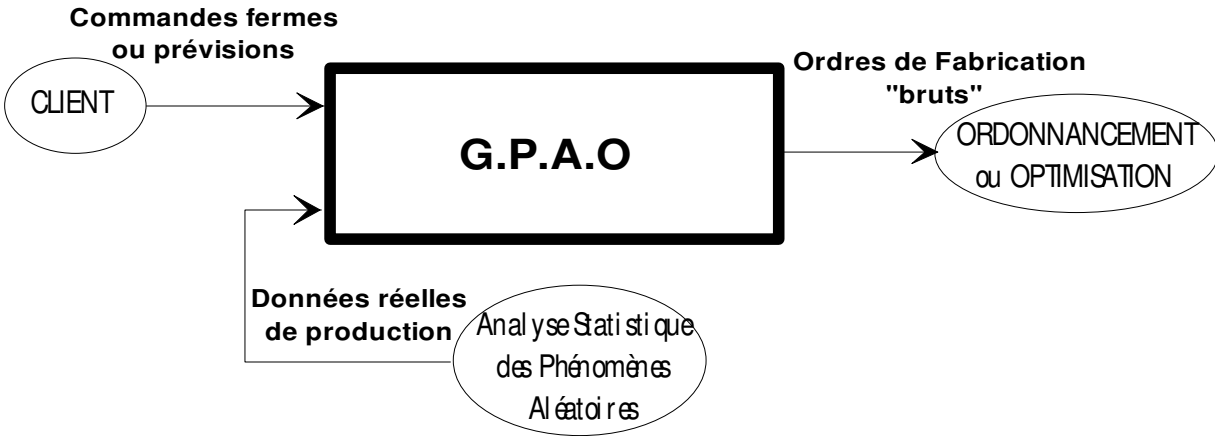


Figure 60 : Bloc fonctionnel de la G.P.A.O

IV.1-b) Ordonnancement

L'ordonnancement a pour fonction de traiter les ordres de fabrication "bruts" issus de la G.P.A.O pour définir de façon précise la séquence de lancement des ordres sur la ligne ou sur les postes de travail. L'ordonnancement peut utiliser la même base de données que la G.P.A.O, et à la différence de celle-ci, détermine si la capacité des ressources est suffisante pour absorber la charge demandée. Dans notre contexte, l'utilité de cet outil est fortement remise en question. En effet, l'utilisation de la simulation et de l'optimisation peut remplacer l'ordonnancement pour déterminer une séquence optimale des ordres de fabrication. Cette possibilité, encore peu exploitée, semble pourtant fort intéressante et permet en plus de visualiser les flux, ce que les outils d'ordonnancement ne proposent pas. Dans un contexte plus général et afin de ne pas débattre plus longtemps sur le bien fondé de cet outil, on l'utilisera en amont de la simulation. La figure 61 présente le bloc fonctionnel point de vue données d'entrée/sortie.



Figure 61 : Bloc fonctionnel de l'ordonnancement

IV.1-c) Simulation de flux

Cet outil a été présenté en détail dans la partie 4 de cette thèse et utilisé pour les expérimentations de la partie 6 et de la partie 8. Le modèle de l'atelier étant considéré comme stabilisé, les données d'entrée correspondent, comme on l'a vu, à des variables de pilotage de la ligne : capacité des stocks amont (StAm), Temps inter-Lancement (TiLa), Nombre de palettes à utiliser (Np), ou ce qui n'a pas été utilisé dans les parties 6 ou 8, séquençement des ordres de fabrication au lancement sur la ligne. Ces variables sont soit saisies manuellement dans un environnement non intégré, soit communiquées par l'outil d'optimisation si ces outils sont connectés. En sortie, la simulation fournit des critères tels que le délai, l'encours, la charge des postes de travail, etc... à l'outil d'optimisation, ou communique directement à la supervision les valeurs optimales des variables de pilotage. La simulation peut aussi utiliser des lois de probabilité provenant de l'outil d'Analyse Statistique des Phénomènes Aléatoires (A.S.P.A). La figure 62 présente le bloc fonctionnel point de vue données d'entrée/sortie.

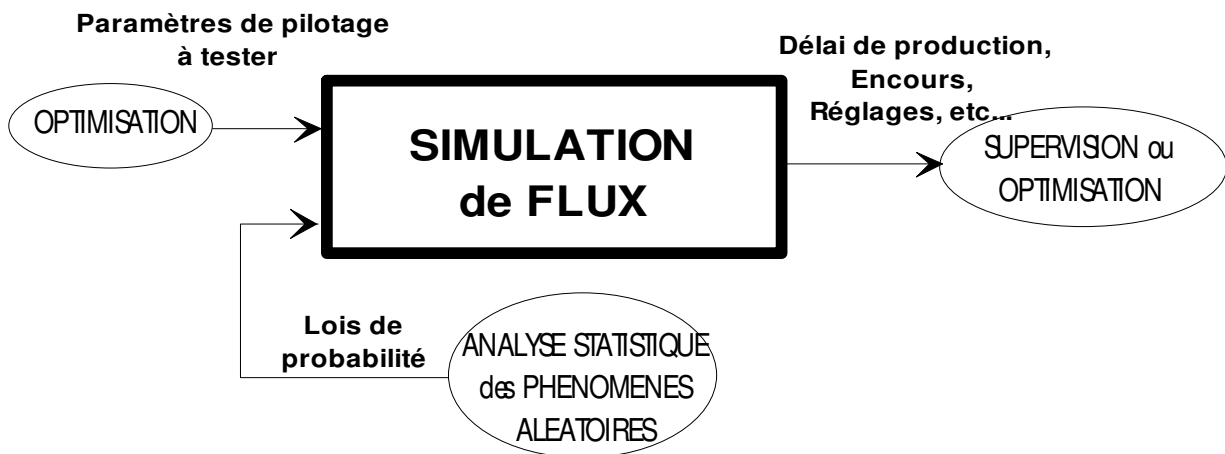


Figure 62 : Bloc fonctionnel de la simulation de flux

IV.1-d) Optimisation

Cet outil a été le thème principal de la partie 8. Comme on l'a vu, il est considéré comme étant connecté à la simulation de flux. Si la simulation n'utilise pas obligatoirement l'optimisation pour alimenter le modèle en données d'entrée, il n'en est pas de même pour l'optimisation qui ne peut fonctionner sans la simulation. On a démontré l'intérêt de cet outil dans les chapitres 1 à 3 de cette partie. Comme indiqué précédemment, couplé à la simulation il peut se substituer aussi à l'outil d'ordonnancement. Dans le schéma d'intégration proposé, cet outil sera donc placé en parallèle à la simulation. La figure 63 présente le bloc fonctionnel point de vue données d'entrée/sortie.

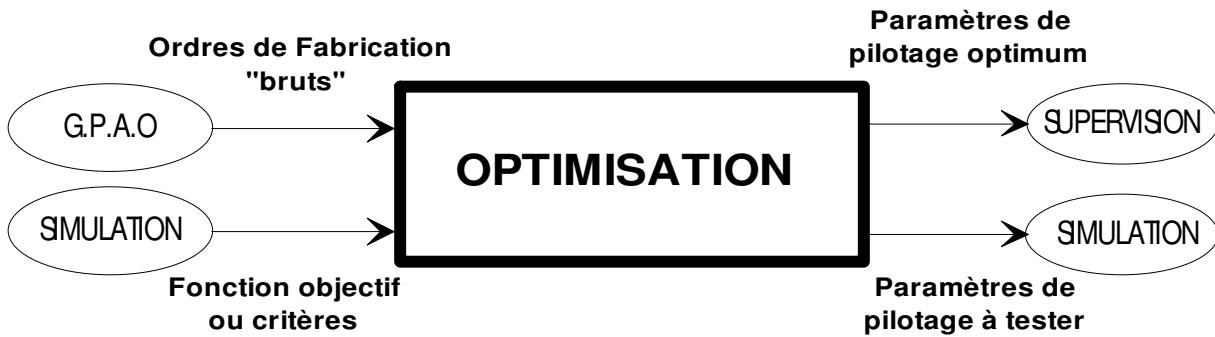


Figure 63 : Bloc fonctionnel de l'optimisation

IV.1-e) Supervision

La supervision est l'outil qui sert d'interface avec la ligne de production. Il est connecté directement aux automates programmables et autres parties commandes qui contrôlent les machines, robots et équipements automatisés de l'atelier. Sa fonction est de collecter les informations utiles et de les présenter à l'écran sous la forme de graphiques (courbes, histogrammes, etc...), d'indicateurs numériques (compteurs), de voyants (alarmes). L'interfaçage homme/machine assuré par cet outil est bidirectionnel. L'opérateur peut saisir des consignes ou des ordres qui seront émis vers l'atelier. Dans le cadre de l'intégration de cet outil, il recevra ces consignes directement de la simulation. La figure 64 présente le bloc fonctionnel point de vue données d'entrée/sortie.

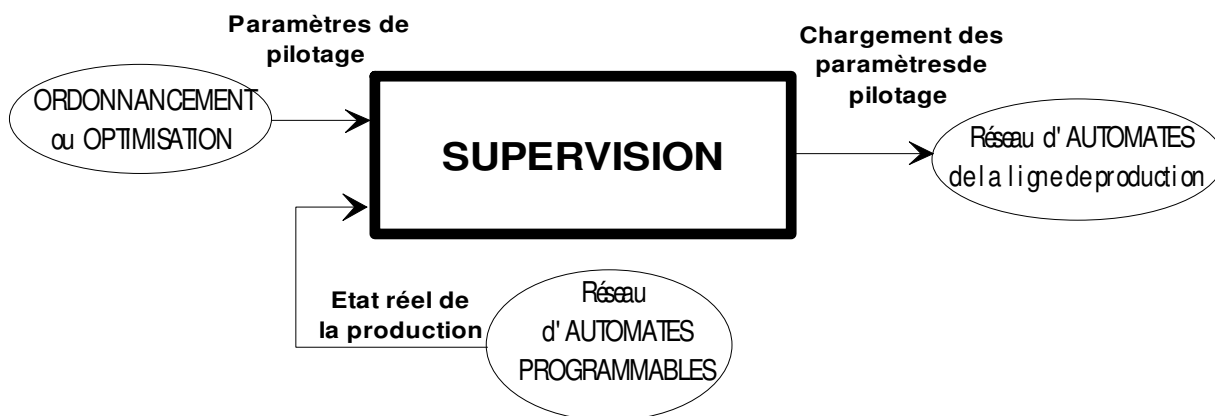


Figure 64 : Bloc fonctionnel de la supervision

IV.1-f) Analyse Statistique des Phénomènes Aléatoires (A.S.P.A)

Cet outil est assez peu connu et développé. Pourtant il est indispensable à une utilisation efficace de la simulation de flux, et peut s'avérer utile pour améliorer aussi la qualité des données de la G.P.A.O. Sa fonction consiste à transformer des données historiques issues d'une même variable (par exemple le temps opératoire réel d'un poste de travail) et de fournir des informations statistiques et probabilistes telles que moyenne, écart type, étendue et loi de distribution mathématique la mieux adaptée. En effet, le système de production réel est soumis à des perturbations aléatoires et le modèle de simulation doit en tenir compte. Il peut donc être risqué d'utiliser des données sous la forme de moyennes, alors que la distribution de ces données n'est ni symétrique, ni uniforme. Certains outils de simulation intègrent cet outil, mais étant donné qu'il peut être intéressant de l'utiliser aussi avec la G.P.A.O ou l'ordonnancement, un module indépendant est préférable. Cet outil traitera en entrée les variables issues de l'outil de suivi de production. La figure 65 présente le bloc fonctionnel point de vue données d'entrée/sortie.



Figure 65 : Bloc fonctionnel de l'analyse statistique des phénomènes aléatoires

IV.1-g) Suivi de production

La fonction de cet outil est de mettre en forme les données sous leurs formes multiples provenant du système de production afin qu'elles soient utilisables par l'outil d'analyse statistique des phénomènes aléatoires. La figure 66 présente le bloc fonctionnel point de vue données d'entrée/sortie.



Figure 66 : Bloc fonctionnel du suivi de production

□ IV.2- Modèle d'intégration des outils

La figure 67 décrit un modèle d'intégration des outils informatiques présentés précédemment.

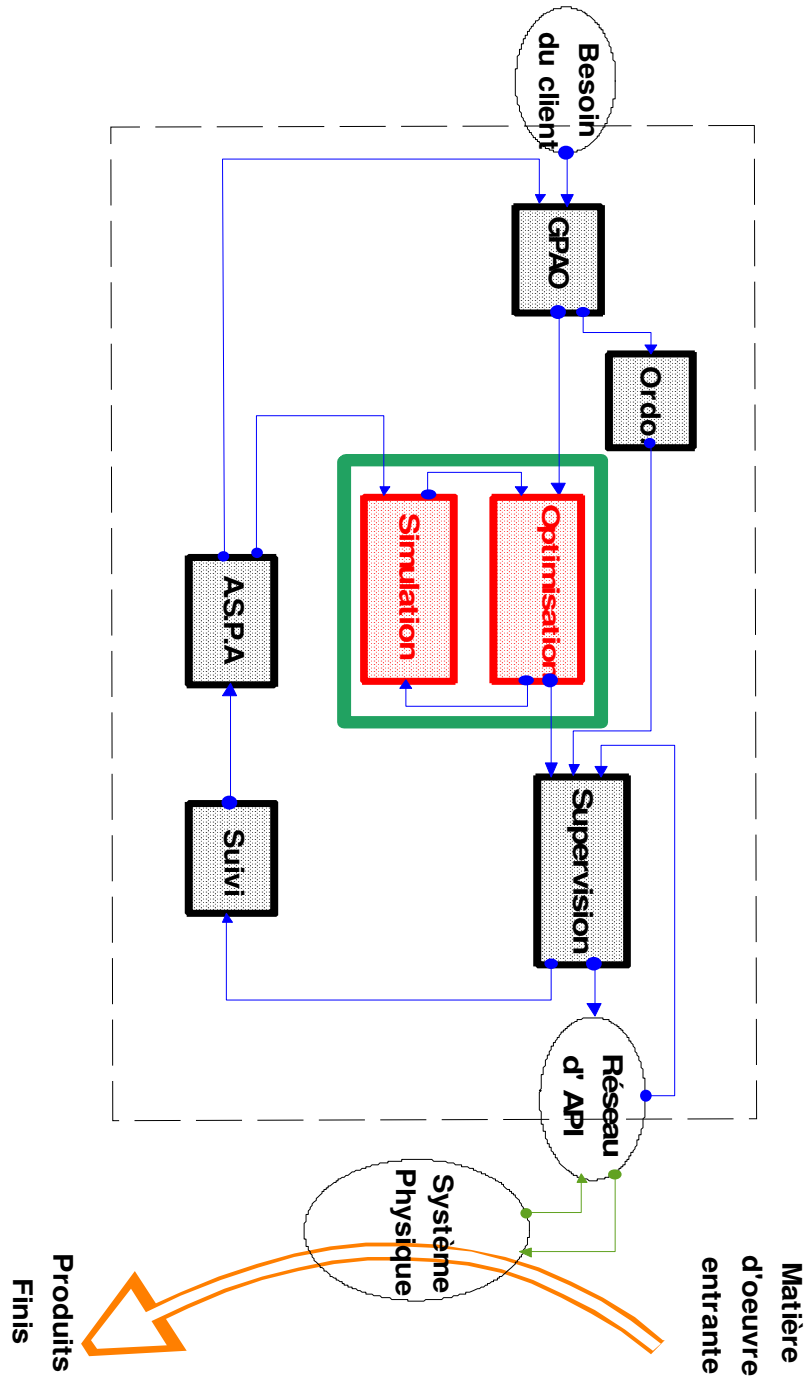


Figure 67 : Modèle d'intégration des outils informatiques

Le processus d'utilisation de ces outils est le suivant :

- A partir de la réception d'une commande client, la G.P.A.O détermine, sous la forme d'Ordres de Fabrication (O.F) et d'Ordres d'Achat (O.A), les composants, sous-ensembles et produits à fabriquer pour une date butoir, sans tenir compte des capacités réelles des ressources. L'ensemble des O.F. est appelé "Campagne de production".
- Les O.F "bruts" de la campagne de production, provenant de la G.P.A.O sont soit acheminés vers un outil d'ordonnancement, soit vers l'outil d'optimisation. L'ordonnancement va se charger de définir une séquence de lancement ou de passage sur les postes de travail en tenant compte de leur capacité réelle. L'optimisation, couplée à l'outil de simulation, peut effectuer non seulement le séquençage optimal du lancement des O.F mais en plus déterminer l'ensemble des paramètres de pilotage en fonction de critères définis par l'utilisateur. Il est important de préciser que le processus d'optimisation/simulation peut se dérouler "off line", pendant qu'une production est en train d'être réalisée dans l'atelier. On peut donc faire "tourner" l'optimisation tant que la campagne de production précédente n'est pas terminée.
- L'ensemble des paramètres de pilotage, optimum ou pas, est communiqué à la supervision qui sert d'interface avec le système physique. Quand la campagne de production précédente est terminée, la supervision transmet aux automates programmables, machines et robots de l'atelier, les paramètres de pilotage et programmes de fabrication de la campagne de production suivante.
- En cours de production, le réseau d'automates renvoie en temps réel au superviseur des informations sur le déroulement de la campagne. Ces informations, sous des formes différentes, sont traitées par l'outil de suivi qui les archive sous un format exploitable par d'autres outils. Cet outil permet aussi d'assurer la traçabilité de la production.
- L'outil d'analyse statistique des phénomènes aléatoires utilise les données mises en forme par l'outil de suivi, les analyse et les transforme en lois de distribution mathématiques. Les données concernées sont les temps opératoires, les temps de panne, les temps de fonctionnement entre pannes, les temps de réglage, etc... Les lois de distribution ainsi déterminées en rapport avec des données réelles peuvent être utilisées à la fois par le modèle de simulation et par la G.P.A.O afin de fournir des résultats plus précis et plus réalistes.

Une telle intégration permet donc d'assurer une utilisation des données en boucle fermée, c'est-à-dire en tenant compte de ce qui se passe réellement dans l'atelier. La G.P.A.O ainsi que la simulation et l'optimisation réactualisent en permanence leurs données afin de fournir des résultats fiables.

Conclusion

Plusieurs thèmes ont été traités dans cette partie avec, pour support, une utilisation de certains outils informatiques encore peu répandue dans l'industrie. Tout d'abord, le problème de l'optimisation par algorithme a été abordé afin de fournir au décideur un outil capable de l'assister pour déterminer, dans des situations complexes, les paramètres optimum de pilotage d'un système de production. Nous avons donc couplé le modèle de simulation, développé dans la partie 4 et déjà exploité "en manuel" dans la partie 6, avec un outil d'optimisation par algorithme, développé pour permettre des échanges de données avec d'autres applications dans l'environnement WINDOWS 95™, en particulier la simulation. Tout d'abord, plusieurs algorithmes ont été testés sur un même problème : algorithme d'exploration aléatoire, algorithme de descente stochastique, algorithme de recuit simulé et enfin algorithme génétique. Il est très difficile d'établir une comparaison objective de chacun de ces algorithmes. La littérature sur ce sujet démontre qu'en fonction des applications, mais aussi des paramètres intrinsèques à ces algorithmes, les conclusions peuvent être différentes. L'objectif de ce test était donc principalement d'étudier le comportement de plusieurs algorithmes couplés au modèle de simulation et de justifier le choix d'un algorithme génétique. Les résultats de ce test montrent la supériorité de l'algorithme génétique, mais les différences ne sont pas significatives, en particulier par rapport à un algorithme d'exploration purement aléatoire. Même si c'est un algorithme génétique qui a été retenu pour la suite du travail, il faut surtout retenir qu'un algorithme de recherche globale à base d'heuristique (Recuit Simulé ou Algorithme Génétique) semble être plus sûr ou plus fiable pour trouver une bonne solution qu'un algorithme de descente ou qu'un algorithme d'exploration aléatoire. Pour résoudre un problème très complexe, l'algorithme d'exploration aléatoire a une chance presque aussi grande qu'un algorithme à base d'heuristique de "tomber" sur une bonne solution. La différence provient du fait que l'algorithme d'exploration aléatoire ne cherchera pas à améliorer une solution trouvée comme le font les algorithmes à base d'heuristique.

La structure de l'algorithme génétique utilise le principe de l'élitisme : seules les deux meilleures solutions d'une génération sont sélectionnées pour le croisement. Il comporte aussi, au vu des résultats fournis par l'algorithme d'exploration aléatoire, un individu par génération totalement issu du hasard (le mutant).

Les expérimentations lancées avec le couplage simulation / optimisation montrent l'efficacité de cette solution. Dans un premier temps, l'optimisation avait pour but de minimiser le délai de production, comme cela avait été fait dans la partie 6. Par rapport aux expérimentations réalisées dans la partie 6, les contraintes sur les variables de pilotage ont été relâchées : les stocks amont (StAm) peuvent varier d'un poste à l'autre, les Temps inter-Lancement (TiLa) peuvent être variables d'une palette à l'autre dans le cas d'une production monoproduit, ou d'un O.F à l'autre dans le cas d'une production multiproduit. Pour la production multiproduit en horizon fini, le délai obtenu avec

l'optimisation est même meilleur que le délai théorique déterminé par la relation 1. Pour la production en horizon infini, le délai minimal est nettement inférieur à celui obtenu dans la partie 6. L'optimisation par algorithme à base d'heuristique couplée à la simulation est donc concluante, à la fois par rapport aux résultats, mais aussi en ce qui concerne le temps nécessaire pour fournir une solution utilisable. On a toujours limité le temps de calcul (proportionnel au nombre d'itérations) afin que celui-ci soit dans les limites de l'acceptable en production. Avec l'augmentation incessante de la puissance des ordinateurs, on peut espérer, à très court terme, pouvoir réaliser plus d'itérations dans un même temps, ce qui favorisera encore la possibilité d'améliorer les résultats.

Dans le chapitre 3, on a comparé des méthodes permettant de sélectionner les individus de l'algorithme génétique dans le cas où plusieurs critères sont à optimiser (optimisation multicritère du délai, de l'encours moyen et des réglages). La méthode d'Optimisation par Somme des critères (O.S.), la plus élémentaire, s'avère inadaptée à des cas où les critères ont des étendues de variation différentes. La méthode de Normage, Pondération et Somme (N.P.S) commence par adapter tous les critères pour qu'ils évoluent entre 0 et 1 (Normage). De cette façon, quelle que soit l'étendue de variation des critères, ils sont tous limités à 0 (valeur minimale) et à 1 (valeur maximale), ce qui évite que les critères ayant une grande étendue de variation faussent les résultats. La deuxième étape de cette méthode est la pondération des critères en fonction de leur importance. C'est une étape délicate, surtout quand il y a plus de deux critères. Pour la réaliser, on utilise une matrice de pondération qui permet de calculer les poids normés des critères en les comparant deux à deux. De cette façon, lors de la somme des critères normés et pondérés, la fonction objectif oscillera entre 0 et 1. Une autre méthode a été développée : l'optimisation par Non-Domination et Pondération (N.D.P). Elle s'inspire du principe de l'optimalité de Pareto mais donne en plus la possibilité de pondérer les critères. Pour comparer ces méthodes d'optimisation multicritères, on a réalisé une série d'expérimentations pour une production multiproduit à horizon fini, avec plusieurs valeurs de pondération pour chaque critère. Les résultats montrent que la méthode N.D.P semble tout aussi efficace que la méthode N.P.S, même si dans un tel cas d'optimisation, il est difficile de dire si une solution est meilleure qu'une autre. L'une ou l'autre de ces méthodes, utilisées conjointement avec l'algorithme génétique, ont prouvé leur efficacité pour l'optimisation multicritère.

Enfin, en s'inspirant de la méthodologie suivie dans cette thèse, le dernier chapitre propose un modèle d'intégration des différents outils informatiques utiles au pilotage d'une ligne de production. Evidemment, la simulation et l'optimisation y ont une place stratégique, au milieu d'autres outils tels que la G.P.A.O ou l'ordonnancement. Le modèle proposé permet de mettre en évidence les échanges de données entre chaque outil pour passer d'une commande d'un client à la production, en insistant sur la réactualisation des données provenant de l'atelier. Dans ce modèle, la simulation et l'optimisation ont une place encore peu répandue dans l'industrie, puisqu'elles sont utilisées en phase d'exploitation d'un système, alors que la plupart du temps, leur champ d'application se limite à la conception (ou à la reconception) du système.

Bibliographie de la partie 8 : "Intégration d'un outil d'optimisation"

[GOLDBERG 94] : "Algorithmes Génétiques"

Editions Addison-Wesley 1994 (traduction française de l'ouvrage anglais paru en 1991
"Genetic Algorithms"

D. E. Goldberg

[RENDERS 95] : "Algorithmes génétiques et réseaux de neurones"

Editions Hermes 1995

J.M. Renders

[VINCENT 98] : "Algorithmes génétiques pour l'optimisation de lignes de production"

Rapport de DEA, E.N.S Cachan, 1998

A. Vincent

[FONTANILI 97] : "Modeling an assembly line with bypass workstations and seeking its optimum management parameters"

Conférence IEPM'97, pages 530 à 540

F. Fontanili, A. Vincent, T. Soriano, R. Ponsonnet

[FONTANILI 98a] : "Comment optimiser un atelier à l'aide de la simulation de flux ?"

Revue Française de Gestion Industrielle, Vol. 16, N°3, 1997

F. Fontanili, A. Vincent

[FONTANILI 98b] : "Flow simulation and genetic algorithms as optimization tools"

International Journal of Production Economics, à paraître fin 99

F. Fontanili, A. Vincent, R. Ponsonnet

[SHINGO 90] : "Le système Shingo"

Editions d'organisation, 1990

S. Shingo

[MOLET 97] : "Une nouvelle gestion industrielle"

Editions Hermes, 1997

H. Mollet

[COHON 78] : "Multiobjectives programming and planning"

Academic Press, 1978

J.L. Cohon

[BENTLEY 98] : "Finding acceptable solutions in the Pareto-Optimal range using multiobjective genetic algorithms"

Cranfield Bath Bombay East Lansing Nagoya, 1998

(on line ed. : http://biole.nuee.nagoya-uc.ac.jp/wsc2/ind_paper/p_bent2.htm)

P.J. Bentley, J.P. Wakefield

[SHAW 96] : "Initial study of multi-objective genetic algorithms for scheduling the production of chilled ready meals"

2nd International Mendel Conference on Genetic Algorithms, 1996

K.J. Shaw, P.J. Fleming

[MARCO 99] : "Multi-objective optimization in CFD by genetic algorithms"

Rapport de recherche INRIA N°3686, 1999

N. Marco, J. A. Désidéri, S. Lanteri

***Bilan et perspectives
de la thèse***

En huit parties, cette thèse est progressivement passée d'une présentation générale de la production, où la variété des situations et des problèmes rencontrés a été mise en évidence, à une proposition d'intégration d'outils informatiques permettant d'optimiser le pilotage d'une ligne de production. ***L'un des aspects importants de ce travail est son application volontairement pratique et concrète, qui a pour origine à la fois l'observation de cas industriels et aussi l'expérience acquise sur une ligne expérimentale en vraie grandeur servant de laboratoire d'essais. Le fil conducteur a été la recherche de solutions permettant à un gestionnaire de production de mieux piloter son atelier, en s'aidant d'outils de simulation et d'optimisation.***

Le système de production expérimental, dont l'architecture est à base de transfert libre asynchrone avec postes de travail en dérivation et construction modulaire, représente une solution qui semble répondre au besoin de disposer d'un système flexible et évolutif, notamment dans le cas des processus d'assemblage. ***L'analyse fonctionnelle qui a été faite sur la base d'un tel système permet au concepteur de maîtriser la réalisation et l'intégration de chaque fonction.***

Sa modularité et sa structure, tant matérielle qu'informationnelle, laissent aussi imaginer des pistes de recherche intéressantes. Autour de l'anneau central (la ligne principale), de multiples configurations de production sont possibles, simplement en jouant sur le nombre de postes en dérivation, sur la longueur des convoyeurs, ou sur le séquençement des opérations. ***Il vient à l'idée que les différents modules fonctionnels peuvent constituer des modules génériques pouvant servir de briques à la conception automatique d'une ligne d'assemblage, en partant de la gamme d'assemblage de chaque produit à assembler et des quantités à réaliser par période.***

La modélisation et la simulation de flux, introduites dans la partie 4, sont rendues nécessaires par les temps de réponse du système réel. Entre le lancement et la livraison d'une production, le temps écoulé peut dépasser plusieurs dizaines de minutes, en fonction de la quantité de produits à réaliser. Grâce à la simulation, on peut connaître ce temps en quelques secondes tout au plus. ***La simulation a donc été utilisée dans cette thèse comme outil de prédiction ou d'anticipation du pilotage d'un système de production en exploitation, permettant de savoir à l'avance quels seront les résultats d'une production en fonction de paramètres de pilotage.***

Pour compléter la piste de recherche abordée précédemment, la conception automatique mais statique d'une ligne d'assemblage pourrait avantageusement être améliorée par la génération, automatique aussi, d'un modèle de simulation de flux afin d'avoir en plus une image dynamique de la future ligne. ***L'idée serait donc d'avoir un générateur automatique de modèle de simulation, utilisant par exemple la base de données d'un logiciel de C.A.O sur lequel on aurait dessiné le futur atelier en s'aidant du générateur automatique de ligne évoqué précédemment.*** Aujourd'hui, le concepteur d'une ligne de production part des produits à fabriquer et de leurs gammes, à partir desquels il dessine l'architecture générale. S'il dispose d'un outil de simulation de flux, ce

qui n'est pas encore systématique, il valide sa conception en construisant un modèle de simulation sur lequel il visualisera la dynamique des flux. Les étapes de conception et de simulation peuvent bien sûr être répétitives, la simulation mettant en évidence des dysfonctionnements que le concepteur corrigera en modifiant ou en ajoutant certaines solutions technologiques. La piste de recherche évoquée consisterait donc à intégrer toutes ces étapes de la conception, en offrant au concepteur des outils d'assistance et de simulation lui permettant d'améliorer sa productivité. En résumé, le processus de conception serait celui-ci : à partir des futurs produits à fabriquer, on commence par élaborer leurs gammes de fabrication, éventuellement en s'aidant de la Technologie de Groupe. Puis, à partir de la gamme et des quantités à fabriquer, le générateur automatique d'atelier propose au concepteur une architecture, qu'il peut valider ou modifier avec son outil de C.A.O. Enfin, automatiquement, en utilisant à la fois la base de données C.A.O et les données de production (gammes et quantités), le générateur de modèle de simulation fournit au concepteur une visualisation dynamique du futur atelier.

Pour en revenir à la simulation de flux, telle qu'elle a été utilisée dans ce travail, on propose de réaliser une étude de simulation en suivant quatre grandes macro-étapes. Cette démarche a pour but de structurer le travail afin de ne rien oublier et de ne rien laisser au hasard avant la réalisation et l'exploitation du modèle. Chacune des macro-étapes est décomposée en étapes plus élémentaires, parmi lesquelles la réalisation du modèle proprement dit qui est la plus délicate. Il faut passer du système réel, existant ou à concevoir, à un modèle informatique. Pour y arriver, **on propose de réaliser un modèle intermédiaire, appelé Modèle Graphique de Connaissances, dont le but est de formaliser le passage du réel au modèle informatique afin de servir de support de communication et d'analyse à tous les acteurs d'une étude de simulation.** Après plusieurs exemples généraux, un Modèle Graphique de Connaissances de la ligne expérimentale a été construit (plus précisément le modèle d'un poste en dérivation, élément le plus représentatif et le plus complexe de cette ligne).

Pour réaliser la suite du travail, on a utilisé exclusivement le modèle de simulation, celui-ci étant validé par comparaison avec le système réel. On a ainsi la garantie de disposer d'un modèle de simulation dont le comportement est très proche de la réalité, afin de l'utiliser comme outil de pilotage. Dans la partie 5, l'objectif était de définir des expérimentations sur le modèle de simulation pour analyser son comportement dynamique en fonction de paramètres de pilotage. **On a mis rapidement en évidence la complexité du système au niveau de la multitude de combinaisons d'hypothèses et de variables de pilotage.** Afin de réduire l'étendue du problème, on a limité volontairement le nombre de variables et leurs niveaux ainsi que les hypothèses associées à ceux qui semblent être intuitivement les plus influents et les plus représentatifs. Des choix difficiles ont dû être faits dans cette partie, ce qui ouvre encore les possibilités de recherche dans d'autres directions, en faisant d'autres choix. Par exemple, en ce qui concerne la règle de priorité des palettes en sortie de poste en dérivation, plusieurs possibilités ont été présentées, mais une seule a été retenue pour les expérimentations. Il semble pourtant intéressant de rechercher l'influence

de cette priorité sur le comportement global du flux. D'autant plus que ce problème de priorité est transposable à d'autres applications (on pense aux réseaux de chariots filoguidés ou autoguidés, mais aussi aux réseaux routiers où la simulation de flux devrait ne pas tarder à percer).

En ne retenant que deux jeux d'expérimentations (un pour des lancements monoproduits, l'autre pour des lancements multiproduits), avec uniquement certaines hypothèses, et en faisant varier uniquement deux variables de pilotage en les contraignant fortement (le Temps inter-Lancement (TiLa) et la capacité des stocks amont (StAm)), on a montré qu'il faut réaliser 120 simulations à chaque expérimentation pour tester toutes les combinaisons. La démonstration de la complexité du problème et de la multitude du nombre de situations est indiscutable et permet là encore d'imaginer d'autres directions possibles de recherche.

Dans la partie 6, on a donc réalisé les expérimentations définies précédemment, en simulant de façon systématique et exhaustive toutes les combinaisons de valeurs des deux variables de pilotage. L'objectif de cette partie était d'analyser le comportement dynamique de la ligne en fonction des deux variables de pilotage choisies. On s'est intéressé en particulier à la détermination des valeurs à fixer aux variables de pilotage pour minimiser une des variables de sortie : le délai (D) de réalisation d'une campagne de production. **Les résultats mettent en évidence l'importance des stocks amont à chaque poste de travail. Leur rôle de régulateur ou d'amortisseur du flux est très net.** Quand la capacité des stocks amont des postes est à sa valeur minimale, le délai est soumis à des variations importantes en fonction du Temps inter-Lancement, alors que ces variations s'atténuent au fur et à mesure que la capacité augmente. Parmi les 120 combinaisons testées, plusieurs permettent de satisfaire la minimisation du délai (D). Le problème est alors de choisir une de ces combinaisons. **On a défini alors un critère complémentaire, appelé critère de stabilité, qui permet de sélectionner les solutions qui garantissent l'obtention du délai minimal, même en cas de légère variation des variables de pilotage.** En prenant en compte encore un autre critère à minimiser, l'encours moyen de palettes sur la ligne (EC), on a réalisé que les valeurs des variables de pilotage sont différentes des valeurs permettant de minimiser les délais (D) tout en respectant le critère de stabilité. **Ce constat a permis de mettre en évidence que le choix des variables de pilotage est issu d'un compromis qui ne permet pas de satisfaire tous les critères simultanément.** Une solution a été proposée en dégradant un peu chaque critère. **Une analyse plus fine des résultats a permis d'établir un certain nombre de relations pour les productions de type monoproduit.** Pour commencer, avec la théorie des contraintes on détermine le Temps inter-Lancement (TiLa) optimal, fonction du temps opératoire du poste goulet. A partir de là, on peut aussi mettre en équation le délai de production minimal d'une campagne. Cette relation sert surtout à fixer un objectif à l'un des critères que l'on tente de minimiser, car comme on l'a vu, les variables de pilotage doivent assurer un compromis et n'ont donc pas automatiquement la valeur théorique issue de cette relation. Ensuite, une autre relation basée sur la minimisation de l'autre critère (l'encours moyen (EC)) permet de déterminer le nombre minimal de palettes dont il faut disposer pour assurer une production sans attente du poste goulet et avec un encours minimal.

En réalisant d'autres expérimentations avec des temps de cycle différents sur les postes de travail, on a déterminé une relation permettant de dimensionner la capacité minimale du stock amont du poste goulet. Avec cette relation et la précédente, le concepteur peut rapidement dimensionner deux paramètres importants de la ligne : le nombre minimal de palettes dont il faut disposer et la longueur des tronçons situés en amont des postes de travail.

Pour les expérimentations réalisées en horizon infini, on constate que le délai nécessaire à la réalisation de 100 produits en régime nominal est très peu influencé ni par la capacité des stocks amont (StAm), ni par le Temps inter-Lancement (TiLa) si celui-ci est inférieur à sa valeur optimale. **En analysant la dynamique du flux, on remarque qu'il se produit un phénomène d'autorégulation de l'encours.** En effet, le nombre de palettes en réserve étant fini, pour pouvoir lancer une palette en production à partir de la ligne de chargement, il faut attendre qu'une palette soit revenue sur la ligne de déchargement. On obtient donc un fonctionnement identique au Kanban, mais sur une boucle de retour entre le dernier et le premier poste de travail. C'est bien sûr le poste goulet qui détermine la cadence de retour des palettes sur la ligne de déchargement : le Temps inter-Lancement (TiLa) est égal au Temps inter-Livraison (TiLi) qui est lui même égal au temps opératoire du poste goulet (Tg) additionné du temps mort (Tm). C'est la raison pour laquelle en régime nominal, le Temps inter-Lancement (TiLa) n'est plus une variable sur laquelle on peut agir. Un réglage optimal conditionne simplement la période de montée en régime. **On a ainsi mis en évidence sur ce type de ligne le principe du ConWip (Constant Work In Process)** défini par Spearman (voir bibliographie de la partie 6). On utilise la relation précédente pour déterminer le nombre optimal de palettes à utiliser pour obtenir cette autorégulation tout en minimisant le ratio de fluidité (Rf) et l'encours moyen (EC). Ce principe, beaucoup moins connu que le Kanban (à ce jour, très peu de chercheurs se sont intéressés aux travaux de Spearman), mérite d'être approfondi et développé à la fois par les chercheurs et par les industriels. Il est applicable à des ateliers de tous types, en particulier ceux qui ne comportent pas de système de transfert asynchrone puisque dans un tel cas, la régulation est automatique, hormis la quantité de produit à lancer qui reste à déterminer (on peut alors utiliser la même relation que celle qui permet de déterminer le nombre minimal de palettes à utiliser). Peut être moins efficace que le Kanban, en ce qui concerne le niveau des stocks entre les postes de travail, il est moins contraignant d'utilisation et plus simple à mettre en œuvre. **Une recherche plus détaillée, éventuellement à base de simulations de cas variés et par comparaison avec un fonctionnement en Kanban ou en flux poussé, devrait permettre de promouvoir le ConWip auprès des industriels.**

Les expérimentations de type multiproduits du jeu d'essai N°2 sont destinées à utiliser la ligne de production comme une ligne capable de traiter simultanément plusieurs Ordres de Fabrication (O.F) différents, avec des gammes différentes et des tailles de lots différentes. Toutes les expérimentations ont été réalisées avec un séquençement du lancement des O.F qui était imposé et figé. **L'influence de ce séquençement n'a donc pas été étudiée et constitue une autre piste à explorer, proche des problèmes d'ordonnancement.**

La première relation développée dans ce cas multiproduit est destinée à déterminer le délai de production d'une campagne comportant plusieurs O.F. On a commencé par proposer une relation issue de celle utilisée dans le cas monoproduit, puis une autre relation basée sur la charge du poste goulet (le poste ayant la charge la plus élevée). Ni l'une ni l'autre de ces relations n'a pu être vérifiée en pratique, ce qui montre la complexité d'un tel cas. Cela étant, elles permettent de fixer un objectif à atteindre pour minimiser le délai. En particulier, la deuxième relation fournit un résultat que l'on est certain de ne pas pouvoir améliorer en pratique. Quand on cherche à optimiser un système, il est important de savoir où l'on se situe par rapport à un objectif absolu, même théorique. De la même façon, la relation permettant de déterminer le nombre minimal de palettes à utiliser n'a pu être vérifiée dans ce cas. Seule la réalisation d'expérimentations avec différentes valeurs permet de trouver une solution optimale. ***Cela a donc amené à rechercher des méthodes permettant d'optimiser automatiquement les paramètres de pilotage de la ligne, en s'appuyant toujours sur le modèle de simulation.*** Si l'on ne contraint plus les variables à ne prendre que certaines valeurs, le nombre de combinaisons devient rapidement très grand, ce qui oblige d'écartier toutes les méthodes de recherche exhaustive, comme celle utilisée dans la partie 6.

Dans la partie 7, on a réalisé un état de l'art des principales méthodes d'optimisation algorithmique permettant de trouver une solution au problème. Compte-tenu de la complexité du problème, ces méthodes n'ont pas pour vocation de trouver l'optimum global, mais plutôt de proposer une bonne solution, ce qui fait d'elles des méthodes approchées ou heuristiques. Parmi elles, on a retenu particulièrement les algorithmes génétiques qui sont choisis pour leur efficacité à résoudre les problèmes complexes.

Le principe du couplage entre la simulation et un algorithme d'optimisation est développé dans la partie 8. ***L'idée consiste à utiliser le modèle de simulation comme une boîte noire recevant en entrée des valeurs des variables de pilotage envoyées par l'algorithme d'optimisation et, après exécution accélérée de la simulation avec ces valeurs, retourne, en sortie vers l'algorithme, la valeur de la fonction objectif.*** En fonction de cette valeur, l'algorithme choisi progresse vers une autre solution qui sera évaluée de la même façon.

Pour commencer, on a comparé les résultats de quatre algorithmes : descente stochastique, exploration aléatoire, recuit simulé et algorithme génétique. Cette comparaison a permis de vérifier l'efficacité de l'algorithme génétique. On retiendra de cette comparaison les résultats assez bons obtenus par l'exploration aléatoire. Cela est probablement dû à la très grande complexité du problème qui ne permet pas à un algorithme "intelligent" de se démarquer nettement par rapport à une exploration purement aléatoire. Cela explique aussi les meilleurs résultats obtenus par l'algorithme génétique qui explore l'espace de recherche en plusieurs endroits en parallèle. L'algorithme génétique utilisé a été détaillé pour mettre en évidence ses particularités : codage réel des variables, sélection élitiste, consanguinité des parents. ***Le choix des paramètres de l'algorithme et les multiples adaptations possibles laissent envisager encore de nombreuses recherches sur ce thème.*** Les résultats obtenus par l'algorithme sont systématiquement

meilleurs que ceux obtenus dans la partie 6, surtout dans le cas de la production multiproduit. **L'algorithme permet d'obtenir rapidement (en moins de 5 minutes) une bonne solution, vérifiée par la simulation. Le gestionnaire de production peut donc programmer les valeurs des variables de pilotage en toute confiance.**

Pour être encore plus proche des problèmes réels, **la suite de la recherche a porté sur l'optimisation multicritère en utilisant toujours l'algorithme génétique et la simulation.** La structure de l'algorithme reste la même que celle utilisée pour l'optimisation monocritère, mais la principale difficulté dans un tel cas est de sélectionner les individus satisfaisant plusieurs critères simultanément. On a choisi trois critères pour illustrer ce problème : le délai de production (D) à minimiser, l'encours moyen (EC) à minimiser et le nombre total de réglages sur l'ensemble des postes de travail (Reg) à minimiser. Deux méthodes de sélection ont été utilisées : la méthode de Normage, Pondération et Sommation (N.P.S) et la méthode de Non-Domination et Pondération (N.D.P), développée à partir du principe de l'optimalité de Pareto. Les résultats obtenus avec l'une et l'autre de ces méthodes sont très proches. Toutefois, **la méthode N.D.P, mise au point expérimentalement pour ce travail mérite d'être approfondie et validée sur d'autres cas.**

Pour conclure cette thèse, un **modèle d'intégration des outils informatiques** à disposition du gestionnaire de production est présenté. **Celui-ci a pour but de définir une structure de communication entre chacun d'eux en ce qui concerne les échanges de données. Il permet aussi d'analyser le processus d'utilisation de ces outils.** Evidemment, la simulation et l'optimisation ont une place importante au sein de ce modèle. Un travail plus poussé mérite d'être fait sur ce thème, d'autant plus que l'arrivée de progiciels d'E.R.P (Enterprise Resources Planning), de S.C.M (Supply Chain Management) et autres E.D.I (Echange de données informatiques) semble être le fait marquant de la fin des années 90.

Table des matières

Sommaire général	1
Problématique de la thèse et présentation générale	3
Partie 1 : Introduction sur la production industrielle	13
Introduction	15
I- Définitions, techniques, règles et principes généraux d'organisation et de gestion de la production industrielle	17
1.1- Contexte général	17
1.2- Définitions de base	17
1.3- Définition de la production	20
1.4- Définition et objectif de la gestion de production	20
II- Evolution de la production et nouvelles règles	22
II.1- Les principales périodes	22
II.2- Les nouvelles règles de la production moderne	23
III- Typologie de la production	24
III.1- Structure d'un produit	24
III.2- Modes de production	25
III.3- Circulation des produits dans l'atelier	26
III.4- Relations avec le client	28
IV- Synthèse des différents modes de pilotage des flux de production	32
IV.1- Pilotage centralisé	32
IV.2- Pilotage par l'aval et décentralisé	33
IV.3- Pilotage par l'amont et décentralisé	34
IV.4- Pilotage synchronisé entre flux physique et flux informationnel	34
IV.5- Pilotage par les contraintes ou méthode OPT	34
V- Typologie des flux physiques et implantation des lignes de production	37
V.1- Flux physique avec opérations et manutentions (transferts) synchronisés (ou liés)	37
V.2- Flux physique avec opérations et manutentions (transferts) asynchrones (ou lires)	38
V.2.a) Manutention non automatisée	39
V.2.b) Manutention automatisée	41
V.3- Autre typologie	41
V.3.a) Les lignes transfert	41
V.3.b) Les cellules flexibles	42
Conclusion	43
Bibliographie	44

Partie 2 : Caractéristiques des processus d'assemblage.....	47
Introduction	49
I - Typologie des processus d'assemblage	50
<i>I.1- Assemblage par « kits ».....</i>	<i>50</i>
<i>I.2- Assemblage par « Rendez-vous ».....</i>	<i>51</i>
<i>I.3- Assemblage par « Pièce Maîtresse ».....</i>	<i>51</i>
II- Analyse fonctionnelle et description technologique des lignes d'assemblage multiproduits à transfert asynchrone automatisé	53
<i>II.1- Fonctions générales</i>	<i>53</i>
<i>II.2- Fonction TRANSFERER.....</i>	<i>53</i>
<i>II.3- Fonction ASSEMBLER</i>	<i>60</i>
II.3.a) Description de la sous fonction STOCKER.....	61
II.3.b) Description de la sous fonction ALIMENTER.....	63
II.3.c) Description de la sous fonction MANIPULER.....	64
<i>II.4- Fonction CHARGER/DECHARGER</i>	<i>65</i>
III- Aspect dynamique du flux de palettes	66
Conclusion.....	69
Bibliographie.....	70
Partie 3 : Présentation du système expérimental	71
Introduction	73
I- Architecture et décomposition fonctionnelle du système expérimental	74
II- Analyse du système informationnel	77
<i>II.1- Typologie du séquençement des opérations sur une ligne transfert</i>	<i>77</i>
II.1.b) Séquençement de type "Transfert libre / Gamme liée / Phases non redondantes"	77
II.1.c) Séquençement de type "Transfert libre / Gamme liée / Phases redondantes"	78
II.1.d) Séquençement de type "Transfert libre / Gamme libre / Phases non redondantes".....	78
<i>II.2- Données de pilotage d'une ligne transfert avec postes en dérivation.....</i>	<i>79</i>
II.2.a) Paramètres intrinsèques à la palette (produit)	80
II.2.b) Paramètres intrinsèques au poste en dérivation	82
<i>II.3- Organisation de la structure de commande</i>	<i>82</i>
Conclusion.....	85

Partie 4 : Modélisation et simulation de flux de production.....	87
Introduction	90
I- Définitions	92
1.1- Définition de la simulation de flux.....	92
1.2- Classification des modèles de simulation.....	92
1.3- Terminologie utilisée en simulation à événements discrets.....	96
II- Limites et objectifs de la simulation.....	98
II.1- Ce que la simulation peut faire.....	98
II.2- Ce que la simulation ne peut pas faire (seule).....	99
III- Résumé de quelques travaux de recherche utilisant la simulation.....	101
IV- Méthodologie de conduite d'une simulation.....	102
IV.1- Analyse du problème.....	103
IV.2- Modélisation et programmation	104
IV.3- Expérimentations sur le modèle.....	106
IV.4- Rapport et conclusions	108
V- Approfondissement de la démarche de modélisation	110
V.1- Objectif d'un Modèle Graphique de Connaissances.....	111
V.2- Etat de l'art et pratiques industrielles sur les modèles graphiques	112
V.2.a) Analyse de déroulement	112
V.2.b) Travaux de recherche	115
V.3- Proposition d'un Modèle Graphique de Connaissances.....	118
VI- Implémentation du modèle du système expérimental.....	126
VI.1- Modèle géométrique du système physique.....	126
VI.2- Modèle Graphique de Connaissances	127
VI.3- Modèle d'action.....	129
Conclusion.....	133
Bibliographie.....	134
Partie 5 : Elaboration du plan expérimental sur le modèle de simulation.....	139
Introduction	142
I- Hypothèses générales	144
II- Hypothèses propres	145
II.1- 1 ^{ère} hypothèse : variété des produits lancés	145
II.1.a) Lancement de type "Monoproduit".....	145
II.1.b) Lancement de type "Multiproduits".....	145

II.2- 2 ^{ème} hypothèse : Horizon de fabrication	146
II.2.a) Horizon de fabrication fini.....	146
II.2.b) Horizon de fabrication infini	146
II.3- 3 ^{ème} hypothèse : Séquencement du lancement.....	147
II.3.a) Séquencement imposé sans découpage des Ordres de Fabrication	147
II.3.b) Séquencement libre sans découpage des Ordres de Fabrication	148
II.3.c) Séquencement avec découpage des Ordres de Fabrication	148
II.4- 4 ^{ème} hypothèse : Séquencement des opérations de fabrication.....	149
II.4.a) Opérations dans un ordre imposé sans redondance.....	149
II.4.b) Opérations dans un ordre imposé avec redondance	149
II.4.c) Opérations dans un ordre quelconque sans redondance	149
II.4.d) Opérations dans un ordre quelconque avec redondance	149
II.5- 5 ^{ème} hypothèse : Nombre de palettes disponibles.....	149
II.5.a) Réserve de palettes supérieure ou égale à la quantité à lancer	149
II.5.b) Réserve de palettes inférieure à la quantité à lancer	150
III- Elaboration du plan expérimental complet.....	151
III.1- Définition du plan expérimental.....	151
III.2- Récapitulation de la typologie des hypothèses propres et estimation du niveau de contrainte.....	151
III.3- Structuration du plan et nombre d'expérimentations.....	152
IV- Variables de pilotage de la ligne	153
IV.1- 1 ^{ère} variable : Cadence de lancement	154
IV.1.a) Cadence de lancement constante entre chaque produit et pour tous les ordres de fabrication.....	154
IV.1.b) Cadence de lancement constante sur chaque ordre de fabrication , mais différente d'un produit à l'autre	154
IV.1.c) Cadence de lancement variable entre chaque produit.....	155
IV.2- 2 ^{ème} variable : Capacité des stocks statiques Amonts	155
IV.2.a) Capacité identique sur tous les postes	156
IV.2.b) Capacité différente d'un poste à l'autre.....	156
IV.3- 3 ^{ème} variable : Capacité des stocks statiques Avals	156
IV.3.a) Capacité identique sur tous les postes	156
IV.3.b) Capacité différente d'un poste à l'autre.....	156
IV.4- 4 ^{ème} variable : Règle de priorité en sortie de poste.....	157
IV.4.a) Règle 1.1 : Prise en compte du nombre de palettes en attente	157
IV.4.b) Règle 1.2 : évolution de la règle 1.1.....	157
IV.4.c) Règle 2.1 : Priorité à la 1 ^{ère} palette à l'intersection et priorité au poste en dérivation	158
IV.4.d) Règle 2.2 : Priorité à la 1 ^{ère} palette à l'intersection et priorité à la ligne principale	158
IV.4.e) Règle 3 : Priorité aux palettes en retard.....	159
IV.4.f) Règle 4 : Prise en compte d'un numéro d'ordre de priorité	159
IV.5- 5 ^{ème} variable : Nombre de palettes à utiliser	160
IV.5.a) $N_p = R_p$: Toutes les palettes en réserve sont utilisées.....	160
IV.5.b) $N_p < R_p$	160
IV.6- Niveau d'influence estimé sur les variables de sortie et valeurs.....	160
V- Sélection des expérimentations et des variables d'entrée	161
V.1- Expérimentations retenues.....	161
V.1.a) Jeu N°1 (Lancement Monoproduit).....	162
V.1.b) Jeu N°2 (Lancement Multiproduits).....	163
V.2- Variables de pilotage testées à chaque expérimentation	166
VI- Mesures des variables de sortie (ou indicateurs de performance) réalisées à chaque simulation	167
VI.1- Temps de défilement de chaque palette : T_d	167
VI.2- Délai de fabrication d'une campagne : D	168
VI.3- Nombre de réglages par poste de travail : $N_r(i)$	168
VI.4- Ratio de fluidité de chaque palette : R_f	168
VI.5- Encours de palettes sur la ligne : E_c	168
Conclusion.....	170

Partie 6 : Analyse des résultats des simulations.....	171
Introduction	173
I- Analyse des résultats du jeu d'essai N°1.....	174
1.1- Rappel des conditions expérimentales générales	174
1.2- Expérimentation N°1	174
1.2-a) Rappel des conditions propres à cette expérimentation	174
1.2-b) Résultats	176
1.2-c) Analyse des données récoltées et recherche d'une combinaison optimale	177
1.2-d) Relations établies à partir de l'expérimentation N°1	184
1.2-e) Expérimentations avec temps de cycle différents	188
1.2-f) Expérimentation avec temps de cycle équilibrés	195
1.3- Expérimentation N°2.....	196
1.3-a) Rappel des conditions propres à cette expérimentation	196
1.3-b) Résultats	196
II- Analyse des résultats du jeu d'essai N°2	202
II.1- Rappel des conditions expérimentales générales.....	202
II.2- Expérimentation N°1.....	202
II.2-a) Rappel des conditions propres à cette expérimentation	202
II.2-b) Résultats	203
II.2-c) Analyse des données récoltées et comparaison au cas monoproduit	205
II.2-d) Expérimentation avec $N_p > 25$ palettes.....	211
II.3- Expérimentation N°2.....	214
II.3-a) Rappel des conditions propres à cette expérimentation	214
II.3-b) Résultats.....	215
II.3-c) Analyse des données récoltées	217
II.3-d) Expérimentation avec $N_p = 30$ palettes.....	220
Conclusion.....	223
Bibliographie.....	224
Partie 7 : Optimisation de la production	225
Introduction	227
I- Modélisation pour l'optimisation	229
1.1- Optimisation sur le système lui-même.....	229
1.2- Optimisation sur un modèle du système.....	230
1.2-a) Formulation (ou formalisation) mathématique ou analytique	230
1.2-b) Utilisation d'un modèle de simulation de flux	231
II- Mise en évidence de la complexité du problème	233
II.1- Objectifs de l'optimisation.....	233
II.2- Calcul de la combinatoire des solutions possibles	234
II.3- Démarche d'optimisation suivie	236
III- Typologie des méthodes d'optimisation	237
III.1- Méthodes exactes ou de recherche opérationnelle	237

III.2- Méthodes approchées ou heuristiques.....	238
III.2.a) Formulation du problème d'optimisation appliquée à la ligne expérimentale.....	239
III.2.b) Principe du voisinage d'une solution	239
III.2.c) Méthodes de recherche locale définies par un voisinage.....	240
III.2.d) Méthodes de recherche globales définies par un voisinage.....	243
III.2.e) Méthodes de recherche aléatoire	247
III.2.f) Algorithmes Génétiques.....	248
Conclusion.....	258
Bibliographie de la partie 7	259
Partie 8 : Choix et intégration d'un outil d'optimisation	265
Introduction	268
I- Comparaison et validation des algorithmes d'optimisation	269
I.1- Rappel des conditions expérimentales choisies.....	269
I.2- Optimisation par algorithme d'exploration aléatoire	271
I.3- Optimisation par algorithme de descente	272
I.4- Optimisation par algorithme du recuit simulé	274
I.5- Optimisation par algorithme génétique	275
I.5-a) Structure de l'algorithme.....	277
I.5-b) Analyse des solutions générées	286
I.5-c) Améliorations possibles de l'algorithme génétique au vu des résultats.....	288
I.6- Synthèse et conclusion.....	288
II- Optimisation du délai de production (cas multiproduit).....	290
II.1- Choix des variables et codage du problème	290
II.2- Optimisation en horizon fini	291
II.2-a) Rappel des résultats obtenus sans algorithme d'optimisation	291
II.2-b) Résultats obtenus par l'algorithme d'optimisation	291
II.3- Optimisation en horizon infini	294
II.3-a) Rappel des résultats obtenus sans algorithme d'optimisation	294
II.3-b) Résultats obtenus par l'algorithme d'optimisation	294
III- Optimisation multicritère	296
III.1- Choix des critères.....	297
III.2- Problématique de l'optimisation multicritère.....	298
III.3- Développement expérimental d'une méthode d'optimisation multicritère.....	300
III.3-a) Méthode 1 : Optimisation par Sommatation (O.S).....	300
III.3-b) Méthode 2 : Optimisation par Normage, Pondération et Sommatation (N.P.S).....	304
III.3-c) Méthode 3 : Optimisation par Non-Domination et Pondération (N.D.P)	313
III.3-d) Conclusions sur les méthodes de sélection testées.....	323
IV- Proposition d'intégration des outils informatiques en gestion de production.....	323
IV.1- Analyse fonctionnelle des principaux outils.....	323
IV.1-a) Gestion de Production Assistée par Ordinateur (G.P.A.O)	323
IV.1-b) Ordonnancement	324
IV.1-c) Simulation de Flux	325
IV.1-d) Optimisation	325
IV.1-e) Supervision	326

Table des matières

IV.1-f) Analyse Statistique des Phénomènes Aléatoires (A.S.P.A)	327
IV.1-g) Suivi de production.....	327
<i>IV.2- Modèle d'intégration des outils.....</i>	<i>328</i>
Conclusion.....	330
Bibliographie.....	332
<i>Bilan et perspectives de la thèse.....</i>	<i>335</i>
<i>Table des matières.....</i>	<i>343</i>

Intégration d'outils de simulation et d'optimisation pour le pilotage d'une ligne d'assemblage multiproduit à transfert asynchrone

Résumé : La faible durée de vie des produits et la multitude des variantes proposées aux clients obligent les entreprises à disposer de systèmes de production flexibles et modulaires. Dans ce travail de recherche, les systèmes d'assemblage, particulièrement représentatifs puisqu'ils constituent souvent l'une des dernières étapes du processus de production où se fait la personnalisation, sont utilisés comme support. On se focalise plus précisément sur l'aspect pilotage du flux dans le cas d'une ligne flexible d'assemblage à transfert asynchrone. En effet, outre les difficultés inhérentes à la conception d'un système flexible d'assemblage, la gestion du flux en cours d'exploitation nécessite, pour chaque nouvelle production, la détermination optimale des paramètres de pilotage afin de minimiser les délais, les encours, les réglages, etc... L'objectif de ce travail est de proposer une architecture d'intégration d'outils informatiques permettant d'optimiser les paramètres de pilotage d'une ligne en exploitation.

La simulation de flux est le premier outil utilisé. Son temps de réponse, très nettement inférieur à celui du système réel, permet de tester un grand nombre de combinaisons des paramètres. Mais si le nombre de combinaisons est très important, la simulation exhaustive de chacune d'elles s'avère longue et fastidieuse.

On propose donc de coupler la simulation à un autre outil : un algorithme d'optimisation de type heuristique. Les expérimentations réalisées avec plusieurs algorithmes et en particulier avec un algorithme génétique, montrent l'efficacité de ce couplage. Pour satisfaire simultanément plusieurs objectifs, l'algorithme est ensuite complété par une méthode de sélection multicritère.

Mots-clés : Systèmes flexibles d'assemblage, Ligne transfert asynchrone, Paramètres de pilotage du flux, Modélisation et simulation de flux, Algorithmes d'optimisation, Algorithme génétique, Optimisation multicritère, Architecture d'intégration des outils.

Integration of simulation and optimization tools for the control of a multiproduct assembly line with asynchronous transfer

Abstract : The weak lifespan of the products and the multitudes of alternatives suggested to customers oblige companies to have some flexible and modular production systems. In this research task, assembly systems, particularly representative since they often constitute one of the last production process stages where personalization is done, are used as support. One focuses oneself more precisely on the flow control in the case of a flexible assembly line with asynchronous transfer. Indeed, in addition to the inherent difficulties in the design of a flexible assembly system, the flow management during exploitation requires, for each new production, the optimal determination of the control parameters in order to minimize deadlines, works in process, setups, etc...

The objective of this work is to propose an architecture of data-processing tools integration making it possible to optimize the control parameters of a line in exploitation.

The flow simulation is the first tool used. Its response delay, very definitely lower than that of the real system, makes it possible to test a great number of parameters combinations. But if the number of combinations is very important, the exhaustive simulation of each one of them proves to be long and tiresome.

One thus proposes to couple simulation with another tool : a heuristic optimization algorithm. The experiments carried out with several algorithms and in particular with a genetic algorithm, show this coupling effectiveness. To satisfy several objectives simultaneously, the algorithm is then supplemented by a multicriterion selection method.

Key-words : Flexible assembly systems, Asynchronous transfer line, Flow control parameters, Modeling and flow simulation, Optimization algorithms, Genetic Algorithm, Multicriterion optimization, Architecture of tools integration.