

Contribution aux calculs des indices de dégradation pour la maintenance prévisionnelle des systèmes de transport

Houda SARIH¹, Mohamed SALLAK², Jérôme BARBAT¹, Walter Schön², Jean-François BARBET¹, Jérôme FENWICK³

¹ SECTOR GROUP, 12 Av. du Québec, 91140 Villebon-sur-Yvette, France. (houda.sarih, jerome.barbat, jean-françois.barbet)@sector-group.net

² Université de technologie de Compiègne, CNRS, Heudiasyc UMR CNRS 7253, 60203 Compiègne cedex, France. (mohamed.sallak, walter.schon)@hds.utc.fr

³ SYNOX, 836 Rue du Mas de Verchant, 34000 Montpellier, France. jfenwick@synox.io

Résumé — La maintenance est un facteur important de qualité, de sûreté de fonctionnement et de productivité dans les installations industrielles complexes. Le projet MAPSYD (Maintenance prévisionnelle des systèmes de transport en présence de données incomplètes et imprécises) est un projet de recherche collaborative qui propose une méthodologie complète de maintenance prévisionnelle partant du recueil des signaux fournis par les capteurs jusqu'à la définition d'une politique de maintenance prévisionnelle. L'objectif de ce papier est de présenter une partie des travaux proposés dans le projet MAPSYD. Notre objectif est d'appliquer les méthodes théoriques, proposées dans la littérature, dans le domaine de la maintenance prévisionnelle des systèmes de transport sur un cas d'étude réel en tenant en compte les contraintes technologiques et les types de données obtenues dans de telles applications.

Mots-clefs — *Maintenance prévisionnelle, systèmes de transport, capteurs, BIC, AIC.*

Abstract — Maintenance is an important factor of quality, operating safety and productivity in complex industrial installations. The MAPSYD project (Predictive Maintenance of Transportation Systems in the Presence of Incomplete and Imprecise Data) is a collaborative research project that proposes a complete predictive maintenance methodology from the collection of signals provided by sensors to the definition of a predictive maintenance policy. The objective of this paper is to present a part of the work proposed in the MAPSYD project. Our objective is to apply the theoretical methods, proposed in the literature, in the field of predictive maintenance of transportation systems on a real case study considering the technological constraints and the type of data obtained in such applications.

Keywords — *Predictive maintenance, transportation systems, sensors, BIC, AIC.*

I. INTRODUCTION

Dans le domaine de transport, les entreprises industrielles sont désormais soumises à de nombreuses contraintes de productivité lors de l'utilisation des équipements industriels, mais aussi aux enjeux sociaux d'impératifs de sécurité des

hommes et des matériels. Pour relever ces défis, la maintenance paraît aujourd'hui comme un élément essentiel permettant de garantir, tout au long de la phase d'exploitation, le maintien en conditions opérationnelles de ces équipements industriels. Même si les coûts des actions de maintenance ne sont pas négligeables, ceux liés aux arrêts de production ont un impact encore plus fort sur la production, les produits ou services proposés par l'entreprise et donc sur les clients. La maintenance industrielle s'est donc développée à partir d'une activité mineure pour devenir une préoccupation majeure et son rôle au sein de l'entreprise a été radicalement transformé. En effet, les objectifs de maintenance ne sont plus formulés afin de maintenir des dispositifs en cours de fonctionnement, mais surtout afin de garantir la flexibilité, la productivité, la qualité et la diminution des coûts. Le choix de stratégies de maintenance a ainsi des impacts essentiels sur la performance et l'efficacité de la stratégie commerciale globale de l'entreprise et son avenir.

Par définition, la maintenance vise à entretenir un équipement afin de lui permettre d'assurer son rôle et de remplir ses objectifs. Cette définition générale nécessite des précisions afin de mieux la situer dans le contexte industriel qui nous concerne dans ce projet. Selon la norme NF-EN 13306, la maintenance est définie comme étant l'ensemble de toutes les actions techniques, administratives et de management durant le cycle de vie d'un bien, destinées à le maintenir ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise [1]. Les méthodes de maintenance peuvent être classées en deux catégories : la maintenance corrective et la maintenance préventive. La maintenance corrective est la maintenance qui est exécutée après la défaillance du système. La maintenance préventive est réalisée lorsque le système est encore en état de fonctionnement. Lorsque la maintenance préventive est exécutée en suivant les prévisions obtenues à partir de l'analyse et de l'évaluation de paramètres significatifs de la dégradation des systèmes, on parle alors de maintenance prévisionnelle (ou prédictive en référence au terme en anglais predictive maintenance) [1]. Auparavant centrée sur la remise en état d'un équipement, la maintenance a désormais pour

objectif de prévoir en temps réel les dysfonctionnements des équipements. En outre, les sous-tâches principales de surveillance des dégradations des composants induisent au niveau de la tâche principale de maintenance l'émergence d'une nouvelle génération de capteurs intelligents à bas coût bénéficiant des progrès croissants des nouvelles technologies de l'information et de communication. Cette émergence a contribué dernièrement à un regain d'intérêt aux méthodes et techniques de maintenance prévisionnelle.

Cependant, le développement d'un système de maintenance prévisionnelle réellement efficient par rapport aux objectifs d'un maintien en conditions opérationnelles requiert à ce jour de définir les fondements scientifiques de la maintenance prévisionnelle ainsi que les méthodes et technologies adéquates. Le premier objectif du projet MAPSYD est de développer un cadre formel de maintenance prévisionnelle des systèmes de transport complexes en présence de différents types d'incertitudes de données recueillies. L'approche proposée permettra de gérer tous les types d'incertitudes de données des capteurs et en particulier les cas où ces données sont parfois imprécises et/ou incomplètes afin de développer un système de capteurs embarqués dans une rame de train, un tramway ou un bus et localisés sur les pièces critiques. Ces capteurs doivent être capables de récupérer l'ensemble des données physico-chimiques nécessaires afin de juger de l'état effectif des composants par rapport à leur durée de vie prévisionnelle.

De manière globale, le projet vise le développement d'un système de prédiction des défaillances combinant de manière innovante l'analyse de SdF et les HUMS (Health and Usage Monitoring Systems : systèmes de suivi de la « santé » et des utilisations de systèmes) afin d'optimiser fortement les actions de maintenance. De manière plus détaillée, la contribution originale du projet MAPSYD dans ce papier peut se décliner selon le thématique suivant :

Modélisation prédictive en présence d'incertitudes : le modèle prédictif nécessaire qui transforme les données collectées en une instruction fiable pour modifier l'organisation des actions de maintenance. En effet, dans ce projet les modèles de SdF à base de la théorie des probabilités imprécises seront présentés et complétés par d'autres modèles liés à la dynamique de données provenant des capteurs, mais également de paramètres plus globaux tels que le nombre fluctuant d'usagers et leur comportement. Ces modèles prédictifs innovants et leurs cadres théoriques utilisant des modèles graphiques incertains doivent être développés.

II. APPROCHE SCIENTIFIQUE ET TECHNIQUE

Dès le démarrage du projet, un site web du projet a été créé (<https://mapsyd.hds.utc.fr>). Le site web contient une présentation du projet, des partenaires et la description de chaque tâche. Il contient aussi un espace interne et confidentiel où les membres du projet peuvent se connecter afin de faciliter la transmission d'informations, techniques et financières, entre les partenaires. L'ensemble des documents (compte-rendu de réunion, rapports d'études, etc.) qui sont nécessaires à l'accomplissement du projet sont ainsi stockés en ligne. Le site permet également d'organiser plus facilement les rencontres et réunions qui auront lieu lors du projet.

Concernant le développement de la méthodologie et conception des modèles prédictifs, l'objectif de cette tâche est

d'utiliser les études de sûreté de fonctionnement afin de déterminer les organes d'un système à surveiller pour la mise en œuvre d'une maintenance prévisionnelle. Ces études sont obligatoires et permettent de mettre en évidence les composants critiques du système (critiques du fait de la sécurité, la disponibilité, la productivité ou des coûts, par exemple). Ainsi, les éléments identifiés seront surveillés pendant toute la durée de leur cycle de vie afin de déterminer leurs conditions réelles d'utilisation et leur état de santé. Il s'agit donc dans un premier temps de déterminer comment évaluer la santé ou la dégradation de chaque élément. De plus, il importe de pouvoir quantifier leur niveau de détérioration à chaque instant, et si possible d'estimer la vitesse de progression de cette détérioration. Par la suite, cela permet d'en évaluer l'impact sur le bon fonctionnement du système et de pronostiquer son état futur (par le biais de calculs de durées de vie restantes par exemple). Grâce à ces pronostics, il sera possible d'instaurer des plans de maintenance adaptatifs. En effet, l'objectif final consiste à développer une stratégie de maintenance dynamique dans laquelle le plan de maintenance évolue constamment en fonction de l'état du système.

L'étape clef de cette tâche est donc l'estimation d'une durée de vie résiduelle (aussi appelée Remaining Useful Life ou RUL en anglais) à l'aide de méthodes de pronostic. Cela soulève de nombreuses questions, notamment concernant la méthodologie à suivre pour pronostiquer la RUL d'un équipement. En effet, cette étude a pour objectif de proposer une méthodologie, la plus générale possible, d'estimation de la RUL d'un composant. Celle-ci s'appuie sur des modèles stochastiques, et plus précisément sur des modèles de Markov cachés (Hidden Markov Models ou HMM en anglais) [2]. Cette méthode est choisie puisque l'état du système n'est pas directement observable et en plus nous disposons de paramètres dont les valeurs seront exploitées pour définir l'état du système. Le principe est de construire un indicateur de dégradation à partir des données collectées sur le système, puis d'évaluer l'état de détérioration du composant, pour enfin estimer une valeur de RUL. L'indicateur est une valeur qui rend compte de l'état d'avancement d'une dégradation et qui permet donc de déterminer la santé d'un équipement à chaque instant. Il est alors possible, grâce au HMM [3, 4, 5], d'estimer le temps moyen restant avant que le composant entre dans un état défaillant.

Le pronostic est indissociable de la notion de confiance. En effet, une prévision implique une certaine incertitude dont il est important d'estimer l'ampleur. La valeur d'une RUL seule ne permet pas de prendre des décisions avisées et il est nécessaire d'y associer la confiance qu'on peut lui accorder. C'est pourquoi, ces travaux se sont intéressés à la gestion de plusieurs types d'incertitudes et à la définition d'intervalles bornant la RUL. En particulier, nous proposons des solutions pour traiter certaines incertitudes aléatoires et épistémiques.

Enfin, les résultats du pronostic dépendent fortement de la qualité des données collectées sur le système. Or l'algorithme de pronostic doit parfois traiter des données altérées. Par exemple, un capteur défaillant peut produire des mesures imprécises voire ne fournir aucune donnée. Une stratégie à adopter dans ce cas de figure est proposée.

III. RESULTATS OBTENUS

A. Résultat 1 : Estimation de la RUL et prise en compte des incertitudes

Nous avons proposé des solutions afin d'évaluer chaque type d'incertitude. Pour les incertitudes aléatoires, nous estimons l'écart-type des temps de fonctionnement avant d'atteindre l'état défaillant. Pour les incertitudes épistémiques nous traitons d'une part celles liées à la méthode en construisant deux nouveaux HMMs qui résultent en deux nouvelles valeurs de RUL qui bornent la valeur initiale. D'autre part, nous traitons les incertitudes liées aux données en proposant des stratégies à adopter en cas de données manquantes.

Cette méthodologie a été évaluée sur un cas d'application créé justement pour l'évaluation de modèles de pronostics : le PHM challenge de 2008. Cela a permis de valider la méthode de construction du HMM et d'estimation des RULs, puis de tester les approches proposées pour la gestion d'incertitudes. L'évaluation des incertitudes aléatoires permet de déterminer un intervalle qui encadre la valeur de RUL obtenue. Plus le processus de dégradation du composant est avancé plus l'intervalle est petit.

L'application de la méthode proposée pour traiter les incertitudes liées à la méthode a montré des limites préoccupantes. En effet, sur certains réacteurs du jeu de données test, les résultats sont concluants : les 2 nouveaux HMMs permettent de définir des bornes à la valeur de RUL obtenue. Toutefois, dans un nombre de cas non négligeable, la RUL initialement estimée n'est pas comprise dans l'intervalle donné par la méthode. Enfin, le traitement des données manquantes permet de proposer une RUL dans les cas où certaines données seraient altérées ou très dégradées. Mais cette méthode présente aussi des limites, en particulier si le nombre de données concernées est important. Elle peut toutefois s'avérer efficace si le nombre de données manquantes est faible et peut constituer une solution de repli temporaire intéressante.

B. Résultat 2 : Développement d'un outil logiciel d'estimation de la RUL

a. Présentation de l'outil

Avant de présenter l'outil, il faut savoir les étapes de l'application (cf. Figure 1). La première étape consiste à construire un HMM à X états et Y observations afin de simuler les données d'apprentissage. Ceci dit il faut définir d'abord, la matrice de transition $\mathbf{A} = [a_{ij}]$ qui donne les probabilités de transition d'un état i vers un état $i+1$ ainsi que la probabilité de rester à l'état i . Ensuite la matrice d'émission de symboles $\mathbf{B} = [b_j(\mathbf{k})]$, qui donne la probabilité d'émission du symbole j étant à l'état i . Finalement le vecteur de distributions initiales \mathbf{P}_i qui donne la probabilité que l'état i soit l'état initial.

L'application a pour but d'exécuter le programme d'implémentation des HMM à partir d'une interface (le code initial étant développé sous Matlab mais l'utilisation de l'interface n'a pas besoin d'utiliser Matlab), ce qui faciliterait la tâche des utilisateurs. Ainsi ils pourraient estimer la RUL en faisant entrer le nombre de composants ainsi que le modèle

HMM $(\lambda, \mathbf{A}, \mathbf{B}, \mathbf{P}_i)$ décrivant leur système comme le montre la figure 2.

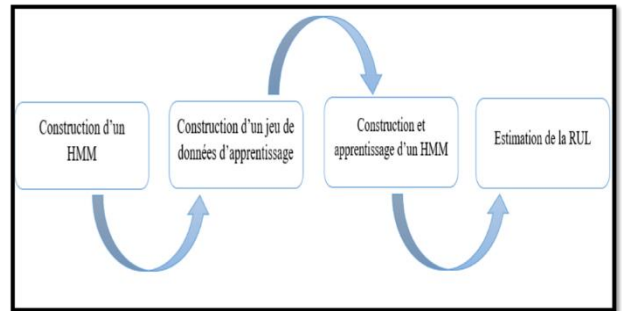


Figure 1. Les étapes de la méthode HMM

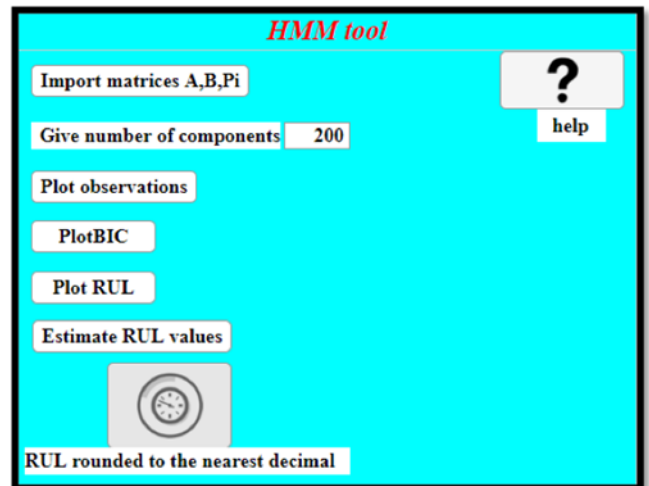


Figure 2. Interface de l'outil RUL

L'interface est constituée des éléments suivants :

- Un élément « import matrices A,B,Pi » pour le HMM décrivant le système.
- Un élément « give number of components » pour donner le nombre des composants.
- Un élément « plot observations » pour exécuter le programme et visualiser les observations qui sont prises comme indice de dégradation.
- Un élément « plot BIC » pour tracer le critère BIC et connaître ainsi le nombre d'états optimal.
- Un élément « plot RUL » pour tracer la RUL.
- Un élément « estimate RUL » pour estimer et stocker les valeurs de la RUL dans un fichier

b. Résultats obtenus par l'outil

Séquences d'observations des états :

Cette étape consiste en la création d'un élément « plot observations » permettant le traçage des observations. Ces observations sont en fait approximées par des polynômes de degré 1 et sont considérées comme étant des indices de dégradation (cf. Figure 3).

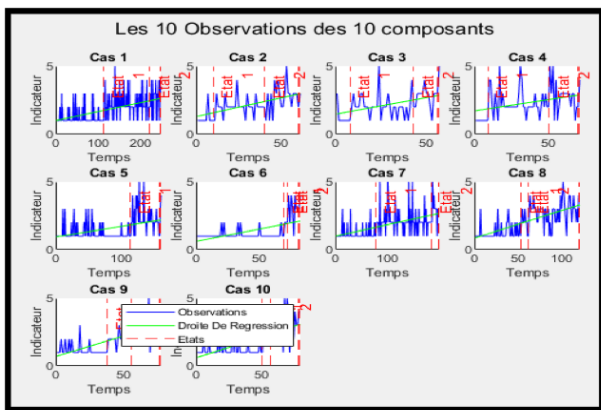


Figure 3. Indices de dégradation des composants

Critère BIC :

Le critère BIC (Bayesian Information criterion) est le critère utilisé pour déterminer le nombre d'états optimal afin de créer les données d'apprentissage qui sont utilisées pour estimer statistiquement les matrices A, B et Pi. Ceux-ci sont les données d'entrée de l'algorithme d'apprentissage. Le BIC est défini par :

$$BIC = -2 \log(\hat{L}) + k \log(n)$$

Où :

- \hat{L} est la valeur maximale de la fonction vraisemblance,
- k est le nombre de paramètres libres du modèle, c'est-à-dire le nombre de paramètres qu'il faut calibrer lorsqu'on construit le modèle. Dans le cas d'un HMM, il s'agit du nombre de paramètres non nuls dans les matrices A, B et Pi,
- n est le nombre d'observations, c'est-à-dire la taille de l'échantillon.

Le critère BIC vise la sélection de variables statistiquement significative dans le modèle pour déterminer le nombre d'états qu'il contient. En effet le critère BIC pénalise plus le nombre de variables présentes dans le modèle ainsi le plus optimal sera celui qui minimise la valeur BIC.

L'exécution du critère BIC montre que le nombre d'états optimal est 4 dans notre étude, par exemple, car c'est celui qui minimise la valeur du BIC (cf. Figure 4).

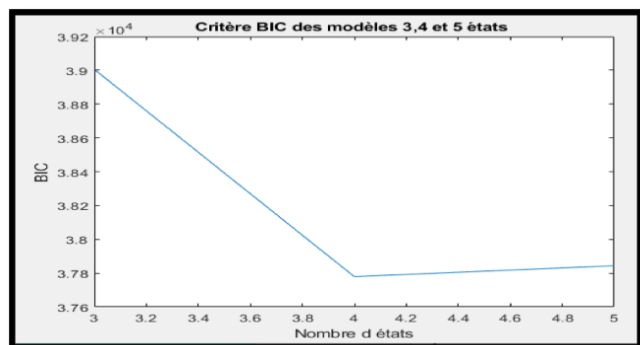


Figure 4. Critère BIC

Traçage de la RUL

Pour permettre le traçage de la RUL, qui est l'objectif principal du programme, on a créé un élément « plot RUL » qui permet le tracé de la RUL (cf. Tableau 1).

Nombre de composants	Durée de Vie Résiduelle (RUL)
1	670
2	398
3	300
4	670
5	675
6	398
7	320
8	680
9	650
10	680

Tableau 1. RUL des composants

C. Résultat 3 : Analyse et classification des signaux prélevés sur les portes du busway

Une application python a été créée qui permet d'indiquer le degré de dégradation des portes (nombre d'états des portes) à partir des données des paramètres environnementaux prélevés (température, pression, ...). Cette application évolue en utilisant un ensemble de données de mesure qui ont été prises sur la porte précédemment décrite et compare une mesure récente par rapport à l'état actuel des portes.

Bien que le principal signe de dégradation de la porte semble être le nombre de chocs subis par les portes tous les jours, nous avons tenu compte des différents paramètres prélevés : pression, température, etc. Les données collectées sont hébergées sur un site externe, dans lequel elles sont représentées dans divers graphiques. A partir de ce site, nous pouvons exporter des fichiers de valeurs numériques qui sont meilleurs pour un traitement automatisé. Nous avons commencé par importer les trois paramètres et un paramètre physique à la fois. Une fois les données nettoyées, le script peut commencer à organiser les données et à travailler avec. La première étape consiste à créer des couples de valeurs. Pour chaque valeur physique V, et les paramètres environnementaux (température), P (pression) et H (humidité), trois couples se forment. On a alors l'équivalent d'une fonction : la valeur physique en fonction des paramètres environnementaux.

Ensuite, l'étape de clustering a été réalisée. Le clustering est un ensemble de techniques utilisées pour partitionner les données en groupes, appelés clusters. Ainsi un cluster est un ensemble d'objets de données qui peuvent être considérés comme similaires, ou au moins plus similaires entre eux qu'avec les objets de données dans l'autre groupes. L'algorithme de clustering que nous avons choisi d'utiliser est l'algorithme K-means :

```

Randomly initialize k centroids
repeat
  Assign each point to its closest centroid
  Compute the new centroid (mean) of each cluster
until
  The centroid position does not change

```

Cependant, nous avons utilisé l'algorithme K-means ++ qui est une variation de l'algorithme des K-moyennes. La différence est avec l'initialisation aléatoire des centroïdes : bien que le premier soit initialisé aléatoirement, les autres ne le sont pas. Ils sont choisis sur tous les points restants mais avec une probabilité qui est proportionnelle au carré de la distance entre le point et la plus proche grappe. Cet algorithme augmente la probabilité d'obtenir la solution optimale.

Le résultat de l'algorithme de clustering (nombre d'état des portes) est montré sur la figure 5. Chaque couleur représente un cluster, et nous avons fixé le nombre de clusters à 6. Nous pouvons voir que chaque cluster n'a pas nécessairement le même nombre de valeurs. Par exemple le jaune au bas de la figure n'a que 2 points. Les valeurs sont normalisées : les 0 sont la moyenne de la série, les 1 sont l'écart type. Une fois toutes les valeurs regroupées, la première étape de notre script est terminée. Nous pouvons alors importer les valeurs des paramètres environnementaux que nous voulons tester dans le script. Une fois les clusters formés, on veut assimiler les valeurs récentes aux grappes. Pour cela, le script forme les mêmes couples, entre la température, la pression, l'humidité et la valeur physique. Ensuite, le script calcule la valeur moyenne $[X, Y]$ de chaque cluster et associe le nouveau couple de valeurs à un cluster avec la méthode `kmeans.predict()`. Cette méthode calcule la distance entre le point formé par le nouveau couple et le centre de chaque cluster. Le cluster dont le centre est le plus proche du nouveau point est celui auquel il est associé. Par exemple, on peut voir sur la figure 5 que le nouveau point, le violet, sera associé au cluster rose.

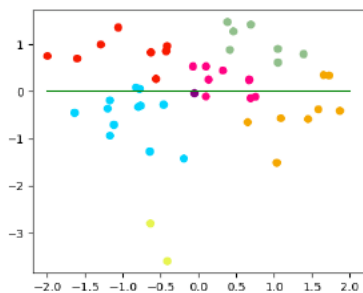


Figure 5. Clusters des différentes valeurs des paramètres environnementaux

IV. DISCUSSION

Le projet MAPSYD a réussi à atteindre l'essentiel de ses objectifs en matière de production de connaissances, d'outils et de solutions de maintenance prédictive dans les systèmes de transport. Ceci fût possible grâce à la pluridisciplinarité et la complémentarité des partenaires issus de la recherche, et de l'industrie mais grâce aussi au personnel de la société Semitan de Nantes.

Malheureusement, nous avons dû annuler de la 2ème campagne d'essais (chez la SEMITAN) destinés à la validation de la partie prédiction suite à la crise sanitaire et nous avons choisi de concentrer nos efforts sur développement de l'outil logiciel de prédiction et l'assemblage des différents modules pour l'obtention d'un module unique de maintenance prédictive.

Le projet a également permis d'identifier des nouvelles pistes de recherche. Nous avons développé et appliqué des méthodes originales pour proposer des solutions de maintenances prédictives. Il reste maintenant à utiliser les méthodes dans des applications différentes et dans d'autres domaines tels que les systèmes de santé.

V. CONCLUSION

Le projet MAPSYD a démontré qu'il est possible de proposer une solution complète allant de la détection des signaux prélevés sur les composants jusqu'à l'analyse et la classification des signaux de dégradation. Beaucoup de conclusions sur la pertinence des méthodes utilisées n'ont cependant qu'une valeur relative, ce qui démontre la nécessité d'appliquer les méthodologies proposées sur différents types de systèmes afin de définir les règles flexibles d'utilisation d'une telle politique de maintenance mais aussi le rôle essentiel des modèles pour établir ces règles, en tenant compte des spécificités et des contraintes liées à chaque système.

Le cas d'étude réelle était le fil rouge de ce projet. En synthétisant les valeurs des paramètres de signaux recueillis étudiant les processus de dégradation des systèmes de portes sous une forme utilisable par les acteurs pour évaluer et concevoir des politiques de maintenance prédictive, ce projet a permis de combiner des disciplines, d'associer des partenaires divers et de raisonner à différentes échelles de travail. Les différents modèles et outils développés sont complémentaires, assistant la définition du nombre d'états des systèmes à la fois pour les systèmes des portes et par la suite la prédiction de leurs durées de vie et permettant une validation des performances des algorithmes de prédiction proposés.

- [1] AFNOR, NF-EN-13306-X-60-319 : Maintenance terminology. Norm, "Association Française de Normalisation", 2018.
- [2] Giantomassi, A, Ferracuti, F, Benini, A, Ippoliti, G, Longhi, S, & Petrucci, A. "Hidden Markov Model for Health Estimation and Prognosis of Turbofan Engines." Proceedings of the ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 3: 2011 ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications, Parts A and B. Washington, DC, USA. August 28-31, 2011. pp. 681-689. ASME.
- [3] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni and G. Tripot, "A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models," in *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 491-503, June 2012, doi: 10.1109/TR.2012.2194177.
- [4] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni and G. Tripot, "A mixture of Gaussians Hidden Markov Model for failure diagnostic and prognostic," *2010 IEEE International Conference on Automation Science and Engineering*, 2010, pp. 338-343, doi: 10.1109/COASE.2010.5584759.
- [5] Yu, J., Liang, S., Tang, D. *et al.* A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction. *Int J Adv Manuf Technol* **91**, 201-211 (2017).