

Sujet de thèse MOdélisation de Dégradation de SysTème de production d'Énergie: Fonction de Répartition empirique et tests d'Adéquation Pour Processus stochastiquE

Acronyme : MODESTE FRAPPE

Directeurs : FOULADIRAD Mitra, TOMASSI Diego
mitra.fouladirad@utt.fr, diego_rodolpho.tomassi@utt.fr

Résumé vulgarisé

L'extension de la durée de vie des systèmes industriels surtout dans le domaine de la technologie de pointe passe par leur surveillance régulière, par la prédiction de leur durée de vie et par la mise en place des actions de maintenance préventive. Nous considérons le cas des systèmes de production d'énergie. Ces systèmes et plus particulièrement ceux qui concernent l'énergie renouvelable sont soumis à des conditions environnementales et opérationnelles très changeantes (le vent, la température, l'intensité du rayon solaire...). De plus, la demande d'énergie fluctue et pour maintenir une production régulière et sans perte, le mode de fonctionnement du système est aussi très variable. Dans ces conditions, il est important d'être en mesure d'estimer la durée de vie restante du système à tout instant et tout au long de sa durée de fonctionnement. Une estimation précise de la durée de vie restante nécessite une analyse complète, efficace et rapide des données de surveillance et repose sur la modélisation du comportement du système dans le temps. Lors de la surveillance, grâce aux données disponibles, des indicateurs de santé du système sont extraits. La modélisation de l'évolution de ces indicateurs dans le temps, par exemple l'évolution de la longueur d'une fissure apparue sur une pale d'éolienne, permet d'être en mesure de prédire en moyenne la date de défaillance du système. Plus le modèle proposé est fidèle au comportement réel du système, meilleur sera le résultat de prédiction. Si la prédiction prévoit une défaillance bien avant la défaillance réelle, les actions de maintenance préventive qui sont mises en place peuvent engendrer des coûts inutiles. Si la prédiction est optimiste à tort, la maintenance préventive sera planifiée après la défaillance et peut avoir des conséquences désastreuses. Une bonne prédiction de la durée de vie d'un système passe par une bonne modélisation du phénomène de dégradation ou de l'indicateur de santé du système. Afin d'être en mesure de prendre en compte l'aspect aléatoire et dynamique de l'évolution de l'indicateur de santé ainsi que l'influence des conditions environnementales et opérationnelles, les modèles probabilistes sont des bons candidats à la modélisation. Avec les connaissances que nous disposons sur le système et sur son environnement un ensemble de candidats de modèle est proposé. Afin de choisir le modèle le plus adapté dans cet ensemble, un outil statistique nommé « test d'adéquation » est indispensable. Ils existent de nombreux outils capables de valider ou rejeter un modèle. Les outils sont souvent applicables sous des hypothèses très contraignantes de régularité des données. Par leur nature, les données de surveillance sont très hétérogènes. Elles sont obtenues par de nombreux et différents capteurs par des inspections, les inspections peuvent être périodes ou non périodiques. Ainsi, l'hypothèse de la régularité des données est rarement satisfaite. Pour construire un outil adapté à la modélisation des indicateurs de santé des systèmes production d'énergie, un long travail scientifique de qualité est requis.

Cette thèse cherche à investiguer la mise en place d'un outil d'aide à la décision pour la modélisation de l'état de santé d'un système de production d'énergie sujet à un environnement changeant. L'ensemble de tous les modèles de dégradation existants dans la littérature sera considéré. Un test d'adéquation flexible, efficace et simple à mettre en place sera proposé. A long terme, grâce à cet outil statistique, la mise en place d'une plateforme de pronostic (prédiction de durée de vie de système) est envisageable. Cette plateforme fournira un jeu

numérique d'un système de production d'énergie sujet à dégradation et permettra la mise en place de tout le processus de pronostic/maintenance en commençant par la collecte des données jusqu'à la prise de décision de maintenance.

Contexte et Problématique

La prolongation de la durée de vie des systèmes de production d'énergie tels que les éoliennes, les centrales nucléaires, les panneaux photovoltaïques, est indispensable pour être en mesure d'assurer une bonne production d'énergie en termes de quantité, de qualité et de durabilité. Ces dernières années, la durée de vie des systèmes de production d'énergie a été prolongée et l'objectif est de maintenir cette tendance. Cette prolongation est le fruit de nouvelles méthodes de surveillance et de la mise en place des actions de maintenance efficaces. Ces dernières ont vu le jour grâce aux nouvelles technologies. La surveillance et les actions de maintenance ne sont efficaces qu'à travers une utilisation intelligente des données disponibles et des indicateurs de santé des systèmes surveillés.

Le sujet de la thèse concerne la modélisation d'indicateur de santé d'un système de production d'énergie à partir des données de surveillance afin de prédire la future défaillance ou panne. Ce problème est au cœur des préoccupations de la région et rentre dans les axes prioritaires de la politique européenne.

La modélisation de la dégradation ou d'un indicateur de santé d'un système vieillissant peut être effectuée à l'aide des processus stochastiques (Lévy, diffusion, etc.) qui représenteraient l'évolution aléatoire, temporel du processus de dégradation. Cette modélisation repose d'une part sur les informations disponibles sur le système telles que : une bonne compréhension des modes de dégradation, de leurs causes et des modèles physiques associées à ces phénomènes et d'autre part sur une analyse des données disponibles. L'intégration des connaissances dans l'analyse des données pour la construction d'un modèle stochastique permettra d'obtenir une meilleure précision dans l'estimation de l'évolution future du système. L'évaluation de la durée de vie restante d'un système est au cœur des études de fiabilité et de maintenance et ces dernières années elle a suscité beaucoup d'intérêt sous le nom de pronostic. L'efficacité des résultats de pronostic repose essentiellement sur le choix des modèles de dégradation et les conditions environnementales et opérationnelles du système.

Dans ce contexte l'existence d'un outil statistique efficace permettant de choisir et valider un modèle mathématique en l'occurrence un processus stochastique parmi un ensemble de candidats est d'une grande utilité. Dans le cadre général des données respectant certaine homogénéité, il existe de nombreux outils statistiques connus et puissants qui permettent de calibrer des modèles stochastiques. Dans le cadre des données dépendantes ou distribuées de manière hétérogène, les outils statistiques classiques ne sont pas utilisables. Dans de nombreux cas spécifiques et sous des hypothèses très contraignantes ou difficilement atteignables quelques outils efficaces sont proposés dans la littérature. Mais dans le cadre des modèles de dégradation, ces outils ne sont pas très répandus ou utilisables et à ce jour, avoir à disposition un tel outil est un grand besoin aussi bien pour les industriels que pour les académiques.

Le sujet proposé rentre dans le cœur des activités de LM2S et fait intervenir les thématiques de recherche de ses deux axes. La modélisation de dégradation et pronostic de durée de vie des systèmes avec des processus stochastiques est un des sujets phares de l'axe fiabilité et maintenance de LM2S. Le traitement des données de capteurs, les tests non-paramétriques et les tests d'adéquation font partie de la large palette des spécialités de recherche de l'axe 1 de LM2S. Le co-encadrement demandé fait apparaître un membre de chaque axe et se fera aussi en collaboration avec d'autres membres l'équipe. Ainsi la complémentarité des compétences

des encadrants permet de traiter un sujet de modélisation et sûreté des systèmes dans sa globalité et enrichir la liste des résultats scientifiques de l'équipe.

Objectifs scientifiques

Les données de dégradation collectées lors de la surveillance sont très hétérogènes. Une des grandes causes de cette hétérogénéité est la manière dont les données sont obtenues. Elles peuvent être par exemple issues de capteurs ou des inspections périodiques ou non-périodique. Ainsi, les données disponibles ne sont pas nécessairement identiquement distribuées. Afin de calibrer le modèle de dégradation, la mise en place d'un test d'adéquation est indispensable. Les tests d'adéquation classiques nécessitent très souvent des données indépendantes et identiquement distribuées. Les tests existants qui concernent les données indépendantes mais non identiquement distribuées ont très souvent une efficacité asymptotique et en pratique exigent des échantillons de très grandes tailles. L'hypothèse de l'indépendance est aussi une grande contrainte, sachant qu'un grand nombre de processus candidat à la modélisation de la dégradation n'ont pas des incréments indépendants mais vérifient seulement l'hypothèse markovienne. Les tests appliqués aux données conditionnellement indépendantes (processus markovien) ne sont pas toujours très efficaces. Dans le cas d'un processus markovien lorsque l'hypothèse « identiquement distribuées » n'est pas satisfaite, il est très difficile de trouver un test efficace. Une autre difficulté majeure pour la modélisation de dégradation est l'aspect temporel des données. Les tests de calibration et validation sur les processus stochastiques très souvent concernent les processus Lévy ou stationnaires. Dans le cadre des processus non-stationnaire ou de diffusion, ils existent des tests cas par cas mais il est difficile de se baser sur la littérature pour construire un outil générique. La majorité des tests d'adéquation se base sur la fonction de répartition empirique ou les statistiques d'ordre. Le problème des variables aléatoires non identiquement distribuées est contourner par une transformation qui ramènerait le problème à un test de loi uniforme. La dépendance et l'aspect temporel des données ne permettent pas d'utiliser directement cette transformation pour effectuer le test. La piste d'une transformation pourrait être explorée pour être en mesure d'exploiter les données non i.i.d. dans un espace plus maniable.

A ce jour, dans le cadre de la modélisation de dégradation, concernant les processus stochastiques markoviens, il y a un grand besoin d'un test d'adéquation simple à mettre en place par les utilisateurs (industriels) et facile à interpréter. La proposition d'un outil générique est un objectif ambitieux et cette thèse va essayer d'apporter quelques éléments de réponse à ce besoin et d'ajouter quelques outils de validation de modèle à la grande bibliothèque des tests disponibles.

Au contraire de la statistique classique, l'efficacité du test proposé doit être aussi validée suivant les critères de pronostic. En d'autres termes, un grand intérêt sera porté sur l'analyse de l'impact de l'erreur de première espèce ou la puissance du test sur la précision de l'estimation de la durée de vie restante du système. Au final, les propriétés de la loi du temps d'atteinte du processus seront des facteurs importants de prise de décision.

Méthodologie de la recherche

Afin de répondre aux objectifs scientifiques du projet, la thèse suivra les étapes suivantes.

- Recherche bibliographique sur les tests d'adéquation appliqués aux données dépendantes ou non identiquement distribuées ou les deux à la fois ;
- Recherche bibliographique sur les tests d'adéquation appliqués aux processus stochastiques markoviens ;

- Construction d'un domaine de décision pour tester l'adéquation au processus markovien et non stationnaire ;
- Validation théorique et numérique du test d'adéquation proposé dans le cadre d'un ou plusieurs processus de diffusion et Gamma non homogène ;
- Analyse de sensibilité du test sur des données simulées ;
- Analyse de l'impact de l'erreur du test sur les résultats de pronostic, plus précisément sur la densité du temps d'atteinte du processus ;
- Analyse de sensibilité des intervalles de confiance de pronostic vis à vis des performances du test ;
- Proposition d'un critère de sélection avec le test proposé tout en optimisant les performances du pronostic ;
- Application des résultats sur des données réelles
- Réflexion sur une perspective d'application à la maintenance