

*ÉCOLE DOCTORALE MATHÉMATIQUES SCIENCES DE L'INFORMATION ET DE
L'INGÉNIEUR*

UMR CNRS 7357

THÈSE présentée par :

Julien CHAPELIN

soutenue le : **18 décembre 2024**

pour obtenir le grade de :

Docteur de l'université de Strasbourg

Discipline/ Spécialité : **Génie Industriel**

**Détection et diagnostic de dérives de
processus de production hétérogènes et
complexes : proposition d'une approche
générique d'intelligence artificielle basée sur
l'apprentissage continu**

THÈSE dirigée par :

M. Bertrand ROSE

M. Benoît IUNG

PR, université de Strasbourg

PR, université de Lorraine

RAPPORTEURS :

Mme. Aïcha SEKHARI SEKLOULI

M. François PERES

MCF-HDR, université Lumière Lyon 2

PR, université de Toulouse

AUTRES MEMBRES DU JURY :

M. Alexandre VOISIN

M. Jean-paul CHEMLA

M. Michel BASSET

M. Olivier JOTZ

PR, université de Lorraine

MCF, université de Tours

PR, université de Haute-Alsace

Responsable PPI, SEW USOCOME

Avertissement au lecteur / Warning to the reader

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition des membres de la communauté universitaire. Il est soumis à la propriété intellectuelle de l'auteur. Cela implique une obligation de citation et de référencement lors de l'utilisation de ce document. D'autre part, toute contrefaçon, plagiat, reproduction ou représentation illicite encourt une poursuite pénale.

This document is the result of a long process approved by the jury and made available to members of the university community. It is subject to the intellectual property rights of its author. This implies an obligation to quote and reference when using this document. Furthermore, any infringement, plagiarism, unlawful reproduction or representation will be prosecuted.

Code de la Propriété Intellectuelle

Article L122-4 :

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite. Il en est de même pour la traduction, l'adaptation ou la transformation, l'arrangement ou la reproduction par un art ou un procédé quelconque.

Any representation or reproduction in whole or in part without the consent of the author or his successors in title or assigns is unlawful. The same applies to translation, adaptation or transformation, arrangement or reproduction by any art or process whatsoever.

Articles L335-1 à L335-9 : Dispositions pénales / Penal provisions.

Licence attribuée par l'auteur / Licence attributed by the author



<https://creativecommons.org/licenses/?lang=fr-FR>

À mes parents,
Pour votre patience et vos sacrifices,
Pour avoir cru en moi,
Pour m'avoir appris la valeur du travail, de la résilience et de l'humilité,
Merci pour tout.

Remerciements

Cette thèse marque la fin d'un chapitre de ma vie riche en apprentissages, en découvertes, et en rencontres. Rien de tout cela n'aurait été possible sans le soutien et la confiance de nombreuses personnes que je tiens à remercier.

Tout d'abord, je souhaite exprimer ma profonde gratitude à **Bertrand Rose** et à **Benoît lung** d'avoir dirigé ma thèse avec beaucoup d'efforts et de patience. J'ai grandement apprécié votre approche pédagogique qui m'a permis de tirer profit de vos connaissances. Les conseils avisés et les critiques constructives que vous m'avez prodigués m'ont aidé à faire avancer mon travail. Votre bienveillance et votre expertise m'ont permis de donner le meilleur de moi-même. Je remercie également **Alexandre Voisin**, dont l'engagement a été déterminant lors des phases d'expérimentations. Grâce à son soutien, j'ai pu développer et valider la partie théorique de mon travail de manière plus approfondie. Ses conseils, sa prévenance et ses disponibilités témoignent de l'intérêt qu'il a porté à ce projet et de son rôle fondamental dans son aboutissement.

Je voudrais aussi remercier chaleureusement les membres du jury. Je remercie **François Peres** et **Aïcha Sekhari Seklouli** d'avoir accepté d'être rapporteurs de cette thèse. Vos remarques et commentaires m'ont permis de corriger certaines erreurs et d'améliorer ma thèse. Merci à **Michel Basset**, pour ses conseils avisés et l'expertise transmise durant les comités de suivi de thèse. Ils m'ont permis d'enrichir mon travail et d'explorer de nouvelles pistes. Merci à **Jean-Paul Chemla**, pour sa bienveillance durant mes années d'études d'ingénieur, et pour m'avoir fait l'honneur d'accepter d'être membre du jury.

Je souhaite également exprimer ma gratitude à **Olivier Jotz** pour m'avoir confié le tout premier travail de thèse de SEW USOCOME. Ta confiance en moi m'a fait prendre une direction cruciale, non seulement dans ma carrière, mais aussi dans mon développement personnel.

Je tiens aussi à remercier le **Service PPI** pour ces presque 9 années de collaboration. Elles ont joué un rôle essentiel dans mon évolution professionnelle et ont contribué à façonner ma maturité professionnelle.

Sans oublier mes nombreux collègues devenus amis, pour avoir transformé mes journées de travail en véritables moments de convivialité. **Mathieu, Marc, Ludovic, Romain, Pierrick, Christophe, Augustin, Lucas, Fabien¹, Joffrey, Thaïs**, et **Émilie**, votre présence, vos encouragements et votre bonne humeur ont rendu cette aventure mémorable.

Merci à mes frères spirituels **Fabien², Thibaud, Franck, Jérémy** ainsi qu'à **tous mes amis**. Entre les très (trop ?) nombreuses sorties restaurants, les concerts à n'en plus finir et nos voyages épiques, vous avez toujours su m'offrir des moments d'évasion qui m'ont permis de tenir le cap.

Un merci tout particulier à mes relecteurs **Fabien¹, Thaïs, Yoann**, et **Chloé**. Votre regard extérieur et vos observations ont été essentiels pour améliorer la qualité rédactionnelle de ma thèse.

Merci à **ma famille** pour son soutien constant et sa présence bienveillante.

Enfin, merci à ma compagne, **Lénie**. Les mots ne suffiront jamais pour exprimer ma reconnaissance. Ton soutien indéfectible, ta patience infinie, et ton amour m'ont accompagné à chaque étape de cette aventure.

Merci à chacun d'entre vous. Ce travail, c'est aussi un peu le vôtre.

Table des matières

Remerciements	4
Table des matières	5
Liste des tableaux.....	7
Liste des figures.....	8
Introduction générale	9
1 Contexte et ambition de la Recherche	9
2 Approche scientifique de la maintenance prévisionnelle.....	10
3 Problématiques scientifiques	11
4 Contributions aux problématiques scientifiques.....	12
5 Organisation du manuscrit	12
Chapitre 1 : Intégration de la maintenance prévisionnelle : un challenge pour les intégrateurs industriels.....	14
1.1 Introduction	14
1.2 Contexte industriel : Production de masse et personnalisation du produit exacerbées	14
1.2.1 L'entreprise SEW USOCOME	14
1.2.2 Le produit : motoréducteur électrique	15
1.2.3 Organisation de la production des usines	16
1.2.4 Les processus de production hétérogènes.....	17
1.2.5 Pôle Processus et Innovation : défis d'optimisation des moyens de production.....	19
1.2.6 Maintenance prévisionnelle : défis et enjeux industriels.....	20
1.2.7 Contraintes et verrous industriels pour une détection de dérive et un diagnostic pertinent	23
1.2.8 Résumé du contexte industriel	25
1.3 Contexte scientifique : PHM et Maintenance	25
1.3.1 Contexte scientifique général	25
1.3.2 Détection de dérive des processus de production hétérogènes.....	26
1.3.3 Diagnostic des dérives des processus de production hétérogènes.....	36
1.3.4 Intégration des étapes de détection de dérive et de diagnostic.....	41
1.4 Conclusion.....	42
Chapitre 2 Vers une approche méthodologique générique pour la détection de dérive et le diagnostic des processus de production hétérogènes.....	44
2.1 Introduction	44
2.2 Détection de dérive des processus de production hétérogènes.	44
2.2.1 Choix de la technique d'apprentissage	45
2.2.2 Proposition d'algorithmes adaptée à la détection de nouveautés	47
2.2.3 Méthodologie guidée pour l'optimisation du traitement des données et la sélection des algorithmes de détection de nouveautés.....	48
2.2.4 Synthèse de la contribution scientifique à la problématique d'adaptation des méthodes de détection des dérives aux spécificités des processus de production hétérogènes	50
2.3 Diagnostic de dérive des processus de production hétérogènes.	51
2.3.1 Choix d'une méthode de diagnostic basée sur les données pour l'identification précise des dérives	51
2.3.2 Complémentarité des algorithmes de classification binaire pour l'élaboration d'un classificateur ensembliste binaire de diagnostic	54
2.3.3 Méthodologie guidée pour le choix des algorithmes de classification binaire pour l'élaboration d'un classificateur ensembliste binaire de diagnostic.....	55
2.3.4 Choix d'une méthode de diagnostic basé sur les connaissances pour le diagnostic des dérives identifiées.....	57
2.3.5 Analyse critique des méthodes de KBM.....	58
2.3.6 Proposition d'une solution multimodèle de diagnostic basé sur des ensembles de classificateurs binaires et des arbres de décisions.	60
2.3.7 Synthèse de la contribution scientifique à la problématique de diagnostic des dérives de processus de production hétérogènes.	61

2.4	Intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes.....	61
2.4.1	Justification de l'élaboration d'une approche méthodologique	62
2.4.2	Proposition d'une approche méthodologique en trois phases de vie	62
2.4.3	Présentation de l'approche méthodologique	63
2.4.4	Étapes de l'approche méthodologique	66
2.4.5	Phase d'expertise diagnostic.....	67
2.4.6	Phase d'apprentissage continu	70
2.4.7	Synthèse de la contribution scientifique à la problématique d'intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes	72
2.5	Conclusion.....	74
Chapitre 3 : Détection de dérive et diagnostic d'un processus de production hétérogène.....		76
3.1	Introduction	76
3.2	Présentation du cas d'usage industriel	76
3.2.1	Description du processus de taraudage.....	77
3.2.2	Données du cas d'usage.....	78
3.2.3	Outils de développement.....	79
3.3	Application de l'approche méthodologique proposée au cas d'usage industriel	79
3.3.1	P1 : Apprentissage de « nouveautés »	80
3.3.2	P2 : Expertise des diagnostics	83
3.3.3	P3 : Apprentissage continu	90
3.3.4	Résultats.....	92
3.4	Conclusion.....	93
Chapitre 4 : Conclusion et perspectives		94
4.1	Conclusion.....	94
4.2	Perspectives	95
Bibliographie de l'auteur.....		97
Bibliographie		98
Résumé.....		108
Abstract.....		108

Liste des tableaux

Tableau 1. Algorithmes d'apprentissage automatique appliqués à la détection de dérive de série temporelle utilisée dans le domaine du PHM	34
Tableau 2. Résumé des méthodes de diagnostics existantes	40
Tableau 3. Récapitulatif des algorithmes de classification adapté à la détection de nouveautés et de leurs spécificités vis-à-vis des données	47
Tableau 4. Récapitulatif des algorithmes de classification binaire adaptés au diagnostic de dérives et à leurs spécificités vis-à-vis des données.	53
Tableau 5. Tableau de complémentarité des algorithmes de classification binaire au regard des indicateurs définis.	55
Tableau 6. Description de l'architecture du fichier .csv de collecte des données.....	78
Tableau 7. Métriques d'évaluation des performances.	87
Tableau 8. Évaluation des performances des classificateurs binaires avec un jeu de données équilibré.....	88
Tableau 9. Évaluation des performances des classificateurs binaires avec un jeu de données déséquilibré.....	89

Liste des figures

Figure 1. Exemple de motoréducteur électrique fabriqué par SEW USOCOME	16
Figure 2. Représentation globale des flux interusines, de la matière première au client.....	17
Figure 3. Représentation du traitement des données et flux d'informations de la norme ISO 13374	22
Figure 4. Combinaisons potentielles des méthodes de diagnostics adaptés de (Montero Jimenez <i>et al.</i> , 2020). 36	
Figure 5. Logigramme de décision pour le traitement des données et le choix d'algorithme de détection de nouveautés associé.....	50
Figure 6. Méthodologie (Logigramme de décision) pour le traitement des données et le choix d'algorithme de diagnostic de dérive associé.	57
Figure 7. Principales phases de vie de l'approche méthodologique et fonctions associées.....	64
Figure 8. Diagramme de flux du cadre complet déployé (Px.y se réfère aux sous-phases dans la Figure 7). Les actions mises en évidence dans les zones (a) et (b) sont détaillées dans les Figure 12 et Figure 15	65
Figure 9. Illustration schématique à base des principes SADT de la méthodologie.....	65
Figure 10. Illustration schématique de la décomposition de la fonction initiale « Détecter/Diagnostiquer » structurant les phases P1, P2 et P3 de l'approche méthodologie.	66
Figure 11. Illustration schématique de l'interaction entre les étapes internes de la phase P1 de l'approche méthodologie.	67
Figure 12. Représentation détaillée sous forme de diagramme de flux du bloc d'action de la Figure 8a, indiquant le regroupement de la base de données "nouveautés" et les étapes d'expertise basée sur la connaissance.	68
Figure 13. Illustration schématique de l'interaction entre les étapes internes de la phase P2 de l'approche méthodologie.	69
Figure 14. Illustration schématique de l'interaction entre les étapes internes de la phase P3 de l'approche méthodologie	71
Figure 15. Représentation détaillée sous forme de diagramme de flux du bloc d'action de la Figure 8b, indiquant la gestion de la création du diagnostic dans l'approche méthodologie.	72
Figure 16. Taraudeuse multibroches et positionnement du flasque moteur sur le produit fini ; (a) taraudeuse multibroches, (b) moteur électrique sans réducteur.	77
Figure 17. Représentation graphique des historiques CSV ; (a) représente un fichier historique d'une heure d'acquisition de données, (b) représente les données acquises durant un cycle de taraudage.	79
Figure 18. Représentation de l'ACF des courants L1 à L3 pour un lag allant de 1 à 5 par pas de 1.	81
Figure 19. Représentation de l'ACF des courants L1 à L3 pour un lag allant de 10 à 50 par pas de 10.	82
Figure 20. Représentation de l'architecture d'exploitation du modèle de détection de nouveautés.	83
Figure 21. Présentation des données du comportement conforme et des profils de dérive identifiés ; (a) représentation des données du comportement conforme, (b) représentation des données de dérive représentant des taraud(s) très usé(s), (c) représentation des données de dérive inconnue, et (d) représentation des données de dérive de taraud(s) cassé(s).....	85
Figure 22. Processus d'isolation des dérives et des défauts ; (a) procédure d'isolation des dérives et des défauts avec DBSCAN, (b) étiquetage par l'expert des dérives des défauts associés aux clusters.....	85
Figure 23. Procédure d'identification experte des dérives du processus de taraudage ; (a) création des clusters, (b) étiquettes associées.	86
Figure 24. Représentation de l'architecture d'exploitation des modèles durant la phase 3 de l'approche méthodologie.	90

Introduction générale

1 Contexte et ambition de la Recherche

Cette thèse de doctorat est un projet collaboratif entre SEW USOCOME, le laboratoire ICube de Strasbourg (dirigeant cette thèse) et le laboratoire CRAN de Nancy. SEW USOCOME est la filiale française du groupe SEW EURODRIVE, leader mondial des systèmes d'entraînement. Les deux laboratoires ICube et CRAN sont des acteurs majeurs de la recherche scientifique en « Génie Industriel / Automatique », notamment sur les problématiques de l'industrie du futur et de l'usine intelligente. Cette thèse est le résultat d'un partage et d'une création de connaissances sur les problématiques de maintenance abordées par SEW USOCOME, ICube et le CRAN, autant d'un point de vue scientifique qu'industriel.

1.1 Contexte industriel

Dans le cadre de la transition vers l'Industrie 4.0, l'importance des données dans l'amélioration des processus de production est de plus en plus reconnue (Monostori *et al.*, 2016). Les données recueillies directement à partir de nouveaux dispositifs d'instrumentation (ex. capteurs, passerelles IO-Link intelligentes) permettent, après un traitement adéquat (ex. création d'indicateurs, suivis des tendances), d'intégrer une connaissance en temps réel de l'évolution des processus de production, offrant ainsi une alternative aux approches de contrôle commande conventionnelles (Zhang and Yan, 2021).

Cette utilisation accrue des données a conduit, entre autres, à une transformation des stratégies de maintenance, favorisant l'émergence de la maintenance prévisionnelle au détriment de la maintenance systématique. En effet, fondée sur la surveillance de l'état réel des systèmes industriels, la maintenance prévisionnelle permet de déclencher des actions de maintenance, « en juste à temps », c'est-à-dire en fonction de la dégradation des équipements analysée à travers des paramètres significatifs (Es-sakali *et al.*, 2022).

La maintenance prévisionnelle s'inscrit comme un levier essentiel pour réduire les temps d'arrêt et optimiser la disponibilité des équipements, un enjeu stratégique pour l'industrie manufacturière de masse, secteur où SEW USOCOME joue un rôle majeur. En effet, avec une production journalière de 5000 motoréducteurs, chaque pourcentage de capacité de production supplémentaire pour SEW USOCOME représente un gain économique significatif, tant en termes de volume produit que de réduction des coûts d'exploitation liés aux interruptions imprévues.

L'ambition de cette thèse, vis-à-vis du contexte industriel, est de faciliter l'implantation de la maintenance prévisionnelle, d'une part au sein de SEW USOCOME, et par généralisation des approches défendues, sein de toutes autres industries manufacturières. En développant des outils et des méthodologies adaptés, cette recherche vise plus précisément à contribuer à **l'intégration des activités de surveillance et de diagnostic de dérives de fonctionnement** des lignes de production afin d'anticiper leurs défaillances et donc les situations opérationnelles non maîtrisées. Dans le cadre de cette thèse, l'application au cas de SEW USOCOME sert à crédibiliser les problématiques initiales posées, mais aussi à valider les contributions scientifiques défendues. Cependant, pour ne pas répondre simplement à un problème d'ingénierie, ce cas est étendu et généralisé afin de justifier des problématiques, mais aussi des contributions qui sont crédibles à des classes d'applications (ex. processus de production manufacturiers), voire à des secteurs industriels complets.

1.2 Problématiques industrielles

Les stratégies de maintenance des moyens de production de SEW USOCOME reposent principalement sur de la maintenance préventive systématique. En raison du caractère statique de cette stratégie et de la haute variabilité et hétérogénéité des processus de production (contraintes fortes du contexte SEW USOCOME) au sein de l'entreprise, des défaillances surviennent malgré tout. Les arrêts non planifiés qui en résultent affectent l'ensemble de la chaîne de production des motoréducteurs. En effet, la fabrication d'un motoréducteur suit un processus séquentiel, où chaque étape dépend de la précédente. Ainsi, lorsqu'une ligne de production est mise à l'arrêt, ce ne sont pas seulement les machines (en support des processus) directement impliquées qui sont impactées, mais aussi toutes les étapes suivantes, créant des retards cumulés sur l'ensemble de la chaîne de production. Par conséquent, un arrêt imprévu sur une seule machine ralentit le temps de réalisation du produit

fini, perturbant également le planning de production global. Dans cette optique, l'entreprise a mis en place une démarche à la fois technique, méthodologique, mais aussi économique d'intégration de la maintenance prévisionnelle pour améliorer ses stratégies de maintenance et augmenter le taux de disponibilité de ses moyens de production.

La maintenance prévisionnelle, avant les phases de pronostic et de décision, se concentre généralement sur l'étude et le suivi de la dégradation des composants physiques en exploitant des modèles de surveillance dédiés à chaque type d'équipement. Cependant, pour SEW USOCOME, cette approche axée sur les composants n'est pas envisageable en raison de la vaste échelle de production : l'entreprise exploite plusieurs centaines de lignes de production, elles-mêmes composées de centaines de composants aux caractéristiques diverses. La diversité et la quantité des composants rendent impossible leur suivi individuel (autre contrainte forte du contexte SEW USOCOME). C'est pourquoi SEW USOCOME cherche à mettre en place une maintenance prévisionnelle orientée sur la surveillance des processus de production à travers tout un ensemble d'indicateurs. L'objectif, au final, est d'obtenir une vue d'ensemble de l'état de santé des processus de production de chaque ligne de production et de détecter (par rapport aux variables, indicateurs retenus) les dérives à un niveau plus global qu'un composant physique unitaire.

Vis-à-vis de ce contexte industriel et des contraintes fortes qui s'y rapportent, plusieurs problématiques industrielles ont été identifiées et constituent les éléments de genèse de cette thèse vis-à-vis des activités de détection et de diagnostic des dérives de fonctionnement :

- **Contextualisation des données collectées** sur les processus de production hétérogènes pour identifier les profils de fonctionnement normal et de défaillance (base de la détection des dérives). En effet, les données brutes recueillies sur les lignes de production ne suffisent pas à anticiper efficacement les pannes. Il est nécessaire de les interpréter dans leur contexte d'utilisation spécifique.
- **Exploitation des données contextualisées pour créer des modèles de détection de dérives**, permettant de suivre en temps réel l'état de dégradation des processus de production (étape clé sur laquelle construire ensuite la prise de décision).
- **Développement de modèles de diagnostic** capables de lier les dérives des processus à des causes racines primaires. Plus les modèles sont pertinents, plus les groupes d'ambiguïté sont réduits et donc l'intervention de maintenance efficace et efficiente.
- **Passage à l'échelle et généralisation** des modèles de détection et de diagnostic de dérives pour une exploitation sur de multiples processus de production dans de multiples usines. L'objectif est de rendre, autant que possible, ces modèles génériques. C'est à dire répétables et applicables à grande échelle dans les usines de production, qu'importe le processus étudié, afin de standardiser les pratiques de maintenance prévisionnelle au sein de l'entreprise. Ou encore dans d'autres entreprises, par exemple de mêmes classes d'applications où la typologie des processus de production est analogue.

Cet ensemble de problématiques industrielles initiées par SEW USOCOME, mais tout à fait généralisables à d'autres secteurs industriels, se positionne en référence à des problématiques plus scientifiques de la maintenance prévisionnelle telles qu'elles sont abordées dans les travaux de la communauté internationale PHM (Prognostics and Health Management) (Kürs, *et al.*, 2021; Sahli, Evans and Manohar, 2021; Goswami and Nandan Rai, 2023; Nunes, Santos and Rocha, 2023).

2 Approche scientifique de la maintenance prévisionnelle

En effet, historiquement, la maintenance prévisionnelle est un réel sujet de recherche qui s'inscrit majoritairement dans les travaux de la communauté du PHM. En ce sens, la maintenance prévisionnelle est formalisée selon la norme ISO-13374 (ISO, 2003) comme l'accomplissement de 7 activités clés : acquisition de données, traitement des données, surveillance, diagnostic, pronostic, prise de décision et interface homme-machine. Chaque étape fournit des informations à l'étape suivante, et une mise en œuvre complète de la maintenance prévisionnelle nécessite de couvrir chacune de ces étapes.

Ce travail de thèse se concentrera sur les deux étapes de surveillance et de diagnostic, comme indiqué précédemment. Plus précisément, vis-à-vis du contexte posé, la surveillance consiste à détecter en temps réel les dérives dans le fonctionnement du système surveillé avec comme objectif de capter les signes de dégradation

dès leur apparition (Foube, 2021). Cette détection repose sur l'analyse des données du processus en temps réel ainsi que sur des modèles capables d'identifier les écarts par rapport aux conditions de fonctionnement normal. Elle conduit à proposer une première vision de l'état de santé du composant surveillé (ex. OK, en état dégradé, en état défaillant). Une fois les dérives détectées, l'étape de diagnostic vise à analyser et à lier ces anomalies détectées aux causes physiques sous-jacentes (Lei *et al.*, 2020).

Le diagnostic de l'état de santé fournit donc des informations essentielles pour comprendre les origines des dégradations et préparer l'étape de pronostic et de prise de décision. En ciblant précisément les composants affectés et les causes racines des dérives, la surveillance et le diagnostic offrent un premier outil dont les résultats sont nécessaires pour optimiser la disponibilité des systèmes par la réduction des arrêts imprévus.

Dans ce contexte scientifique PHM et pour répondre aux problématiques (généralisées) industrielles, ce travail de thèse se concentre ainsi sur le développement de modèles de surveillance et de diagnostic adaptés pour une maintenance prévisionnelle axée sur l'analyse des processus de production. Ces développements répondent à des questions de recherche justifiées par rapport au contexte PHM.

3 Problématiques scientifiques

En ce qui concerne le défi industriel consistant à développer des modèles de détection d'anomalies, l'un des principaux obstacles rencontrés en exploitation est lié à la rareté de données d'anomalies étiquetées. En effet, ce n'est pas toujours possible en raison de la rareté de ce type d'événements de défaillance (Chen *et al.*, 2023; C. Li *et al.*, 2024). Cet obstacle est d'autant plus vrai lorsque le processus de production est hétérogène et avec une grande variabilité. Comme dans le cas de SEW USOCOME qui peut aussi amener la génération de dérive initialement inconnue, car elles sont liées à la dynamique du processus (Van De Sand and Eng, 2020; Kamm, Jazdi and Weyrich, 2021). Ceci conduit pour la thèse à poser une première problématique scientifique : ***Comment développer une méthode de détection des dérives de processus de production hétérogènes ne nécessitant pas de données étiquetées et suffisamment évolutive pour s'adapter à toute dérive apparaissant tout au long de la vie du processus ?***

Les première et troisième problématiques industrielles soulignent quant à elles l'importance de contextualiser les données pour identifier les profils de fonctionnement et d'introduire des modèles de diagnostic capables de lier les dérives des processus à des causes racines physiques. Cependant, dans les processus de production hétérogènes, les dérives peuvent découler d'interactions complexes entre divers composants physiques impliqués dans les éléments surveillés. Cela rend le diagnostic très difficile. En ce sens, il n'existe pas réellement de travaux scientifiques sur des méthodes de diagnostic adaptées aux processus de production hétérogènes, capables de relier les dérives à des causes racines spécifiques (Swain, Khilar and Bhoi, 2017; Nunes, Santos and Rocha, 2023). Les interactions complexes entre éléments rendent cependant inefficace l'utilisation exclusive de modèles basés sur les données sans l'apport de connaissances expertes (Montero Jimenez *et al.*, 2020). Les connaissances expertes semblent par conséquent une voie pour pallier ces limitations en intégrant l'expertise humaine pour l'analyse des dérives, même en l'absence de grands ensembles de données étiquetées (Atoui and Cohen, 2021).

Ceci conduit pour la thèse à poser une deuxième problématique scientifique : ***Comment concevoir un système de diagnostic des dérives pour les processus de production hétérogènes, combinant des modèles basés sur les données (DDM) et des modèles basés sur la connaissance (KBM) ?***

Une réponse à cette question doit permettre de disposer d'un système permettant de relier les dérives observées à leurs causes racines spécifiques, en s'appuyant à la fois sur les données disponibles, et sur l'expertise des opérateurs et ingénieurs. Cela rendrait les interventions de maintenance plus précises et efficaces.

Enfin, la quatrième problématique industrielle posée impose que les modèles répondant aux deux questions de recherche précédentes puissent être une solution pour le cas SEW USOCOME, mais aussi une extension/généralisation de ce cas. Ceux-ci permettraient de résoudre l'implantation des activités de détection et de diagnostic, mais aussi leur intégration sur de multiples processus. Face à ce besoin d'implantation pour des cas multiples, très peu de travaux existent et/ou ne résolvent qu'une partie du problème (Potthoff Jean Pauland Tienken, 2023; Mahfoud Hassanaand Moutaoukil, 2024).

Cette limitation conduit pour la thèse à poser une troisième problématique scientifique : ***Comment élaborer une approche méthodologique unifiée pour la détection et le diagnostic des dérives dans les processus de production hétérogènes ?***

Une réponse à cette question doit permettre d'offrir une approche méthodologique construite sur une suite d'étapes organisées pour garantir une implémentation efficace et standardisée des solutions de maintenance prévisionnelle à grande échelle.

Des réponses scientifiques à chacune de ces trois problématiques de recherche sont nécessaires pour fonder les bases de la mise en œuvre d'une maintenance prévisionnelle basée sur la surveillance/diagnostic de la dégradation des processus de production hétérogènes.

4 Contributions aux problématiques scientifiques

En réponse aux 3 questions de recherche posées, cette thèse apporte trois contributions scientifiques majeures qui sont ensuite validées au cas de SEW USOCOME. Ces contributions s'inscrivent totalement dans le domaine du PHM et apportent des originalités majoritairement méthodologiques vis-à-vis de la détection/diagnostic de dérives de processus de production hétérogènes.

4.1 Contribution à la surveillance dans le PHM

La première contribution de cette thèse, en réponse à la première question de recherche, concerne la formalisation d'une méthodologie guidée pour le choix d'algorithmes de détection de dérives à base de techniques d'apprentissage automatique. Cette méthodologie repose sur le calcul d'indicateurs permettant de caractériser les données conformes (fonctionnement normal) et d'orienter en conséquence la sélection de l'algorithme le plus adapté pour détecter les dérives du processus avec un logigramme décisionnel. En analysant les caractéristiques spécifiques des données, cette approche vise à maximiser la précision et la pertinence de la détection de dérives en fournissant l'algorithme le plus efficace en fonction des caractéristiques des données. Cela permet d'ajuster l'étape de surveillance en fonction de la dynamique des processus surveillés, assurant ainsi une détection de dérive adaptée à chaque processus de production hétérogène.

4.2 Contribution au diagnostic dans le PHM

La deuxième contribution, en réponse à la deuxième question de recherche, porte sur le développement d'une méthodologie guidée pour le choix des algorithmes de diagnostic basés sur les données en fonction des caractéristiques de ces données, en s'appuyant une fois de plus sur le calcul d'indicateurs permettant de caractériser les données. Cette méthodologie facilite la sélection des algorithmes basée sur les données les plus appropriées via un logigramme décisionnel. En complément, une architecture de diagnostic combinant les modèles basés sur les données et les modèles basés sur les connaissances est proposée. Cette architecture hybride permet de surmonter les limitations des modèles basés sur les données seules, notamment dans des environnements où les interactions entre les composants de production sont nombreuses et complexes, et en tirant profit des connaissances expertes pour relier les dérives détectées à des causes racines physiques. Cette approche contribue ainsi à accroître la précision et l'efficacité des interventions de maintenance, en offrant un diagnostic complet qui combine données et expertises.

4.3 Contribution méthodologique unifiée pour le PHM

Enfin, la troisième contribution de cette thèse, en réponse à la troisième question de recherche, consiste en la proposition d'une approche méthodologique (une méthodologie) unifiée pour la détection et le diagnostic des dérives dans les processus de production hétérogènes. Cette approche méthodologique propose une suite d'étapes/activités cohérentes et intégrées pour le PHM, permettant une application systématique et évolutive des méthodes de détection et de diagnostic, quels que soient les processus étudiés. L'approche méthodologique permet aussi de contextualiser les données via une étape d'expertise des données de dérive. Cette approche méthodologique permet donc d'harmoniser les pratiques des activités de détection et de diagnostic des dérives en maintenance prévisionnelle et d'augmenter la connaissance des profils de dérive des données de processus.

5 Organisation du manuscrit

Afin de présenter et défendre ces trois contributions, le manuscrit est organisé comme suit :

Le chapitre 1 présente les caractéristiques essentielles des processus de production des motoréducteurs de SEW USOCOME, en identifiant les défis industriels qui motivent cette thèse et en les replaçant dans le contexte scientifique du PHM. Il est ainsi introduit le contexte industriel décrivant les spécifications techniques des motoréducteurs, et les contraintes associées aux processus de production, ainsi que les objectifs stratégiques de SEW USOCOME et leur influence sur les pratiques de maintenance. À partir de cette analyse, les principales problématiques industrielles (verrous industriels) rencontrées pour la mise en œuvre des étapes de surveillance et de diagnostic de dérives de fonctionnement sont dégagées. En complément de cette vision industrielle, le reste du chapitre expose le contexte scientifique, en présentant d'abord la communauté scientifique de la

maintenance prévisionnelle, puis en confrontant les problématiques industrielles aux travaux actuels du PHM. Cette analyse permet d'identifier les lacunes scientifiques qui empêchent de résoudre ces problématiques industrielles, ces lacunes étant transformées en trois questions de recherche justifiant le travail de thèse. Enfin, ce chapitre se termine en résumant les points clés et en introduisant les chapitres suivants.

Le chapitre 2 répond aux trois questions de recherche soulevées au chapitre précédent en proposant des solutions méthodologiques dans le domaine de la surveillance et du diagnostic. Chaque section du chapitre 2 est dédiée à l'une des questions. La première section traite de l'adaptation des méthodes de détection de dérives aux spécificités des processus de production hétérogènes en proposant une solution basée sur la détection de nouveautés, permettant d'identifier des dérives sans connaissance préalable des modes de dérives des processus. La deuxième section aborde le diagnostic des dérives dans des processus industriels hétérogènes, en présentant une solution combinant l'expertise humaine avec des techniques d'IA pour associer les dérives à leurs causes racines physiques. La troisième section présente une approche méthodologique unifiée intégrant de manière complémentaire les activités de détection et de diagnostic des dérives, avec une boucle d'apprentissage continu qui enrichit le système de nouvelles données et connaissances pour optimiser la détection et le diagnostic des dérives. Le chapitre se conclut par un résumé des contributions méthodologiques de la thèse.

Le chapitre 3 a pour objet de proposer une première validation (« restreinte ») des contributions méthodologiques en les appliquant à un cas d'usage industriel proposé par SEW USOCOME, démontrant ainsi leur pertinence dans un contexte de production réelle. La première section présente le cas d'usage industriel, en détaillant les caractéristiques techniques et les défis spécifiques rencontrés dans la surveillance de ce processus de production. La seconde section évalue l'applicabilité et l'efficacité des contributions méthodologiques, notamment les méthodologies de sélection d'algorithmes et l'approche méthodologique unifiée pour la détection et le diagnostic des dérives. Le chapitre se termine par une discussion des résultats, en examinant les possibilités d'amélioration et d'extension des approches proposées à d'autres environnements industriels.

Enfin, **le chapitre 4** conclut le manuscrit par un récapitulatif des objectifs initiaux de la thèse, en reprenant les réponses méthodologiques apportées aux problématiques industrielles et scientifiques, et propose également 6 perspectives pour de futurs travaux dans ce domaine.

Chapitre 1 :

Intégration de la maintenance prévisionnelle : un challenge pour les intégrateurs industriels

1.1 Introduction

L'objectif de ce chapitre est de présenter les caractéristiques essentielles des processus de production des motoréducteurs SEW USOCOME. Il s'agit d'identifier les verrous industriels rencontrés dans le cadre de cette thèse de doctorat et de les replacer dans le contexte scientifique de la communauté du « Prognostics and Health Management » (PHM).

En ce sens, la Section 1.2 présente le contexte industriel dans lequel s'inscrit le travail de thèse. Pour ce faire les spécifications techniques et les caractéristiques des motoréducteurs produits par SEW USOCOME sont introduites. Ensuite, les contraintes imposées aux processus de production par les spécifications et caractéristiques produites sont discutées. Afin de compléter cette vision très axée sur le terrain, cette présentation se concentrera sur les objectifs stratégiques de SEW USOCOME et leur impact sur les pratiques de maintenance en se concentrant sur la détection et le diagnostic des dérives de processus. À partir des contraintes et des objectifs identifiés, les principaux verrous industriels rencontrés par SEW USOCOME pour la mise en œuvre des étapes de surveillance (détection) et de diagnostic de la maintenance prévisionnelle sont isolés.

La Section 1.3 présente le contexte scientifique du travail de recherche qui, vis-à-vis du positionnement en maintenance prévisionnelle, se réfère à la communauté PHM. Dans un premier temps la communauté scientifique associée à la thématique de la maintenance prévisionnelle est présentée. Dans un second temps les verrous industriels sont replacés dans un contexte de recherche scientifique plus global en les confrontant aux travaux actuels du PHM. De cette analyse résulte un état de l'art concernant les activités de détection de dérive et de diagnostic du PHM. À partir de cet état de l'art, des problématiques scientifiques non résolues en regard de verrous industriels sont isolées, justifiant ainsi le travail de thèse. La méthodologie de recherche globale appliquée à la thèse est également expliquée et illustrée dans cette section.

Enfin, le chapitre se conclut en résumant les points clés des sections précédentes du chapitre et en introduisant les chapitres suivants, où les verrous industriels et scientifiques sont adressés à travers la proposition d'une approche méthodologique générique pour la détection de dérive et le diagnostic des processus de production.

1.2 Contexte industriel : Production de masse et personnalisation du produit exacerbées

1.2.1 L'entreprise SEW USOCOME

La société SEW USOCOME est la filiale française du groupe SEW EURODRIVE, leader mondial dans le secteur des motoréducteurs électriques (Figure 1). Avec une présence établie sur le territoire Français depuis 1959, SEW USOCOME s'est forgé une réputation en offrant des produits et des services de haute qualité, tout en assurant des délais de livraison et de service après-vente parmi les plus compétitifs du marché. Avec sa stratégie axée sur la qualité, la personnalisation et la satisfaction client, SEW USOCOME se place aujourd'hui comme un

acteur majeur de l'industrie française et s'efforce de répondre à une grande diversité de défis industriels pour atteindre ses objectifs stratégiques, à savoir :

- **Maintenir une capacité de production élevée** : SEW USOCOME vise à produire jusqu'à 5000 ensembles motoréducteurs par jour.
- **Offrir une personnalisation du produit étendue** : L'entreprise propose jusqu'à 7 millions de références uniques au travers de son architecture produit modulable, démontrant son engagement envers la personnalisation du produit.
- **Assurer des produits et des services de haute qualité** : SEW USOCOME met l'accent sur la qualité des produits et des services pour répondre aux attentes élevées de ses clients.
- **Garantir un temps de fabrication des commandes de 5 jours** : L'objectif de l'entreprise est de garantir un temps d'attente minimal entre la commande et la livraison pour répondre aux besoins des clients dans les plus brefs délais.
- **Assurer des délais de dépannage (service après-vente) parmi les plus compétitifs du marché** : SEW USOCOME s'engage à fournir des délais de dépannage de ses produits compétitifs pour garantir la satisfaction des clients.

1.2.2 Le produit : motoréducteur électrique

Concernant le produit fabriqué par SEW USOCOME, le motoréducteur, il est la combinaison d'un moteur et d'un réducteur. Dans les systèmes d'entraînement fabriqués par SEW USOCOME, le moteur est toujours électrique, principalement des moteurs asynchrones triphasés. Les motoréducteurs, ou les systèmes d'entraînement électrique au sens large, convertissent l'énergie électrique d'un réseau ou d'une batterie en énergie mécanique. Le moteur/motoréducteur va ensuite permettre d'utiliser la force mécanique qui découle de cette conversion pour générer des mouvements et fournir une puissance ajustée avec précision aux besoins de l'application des clients.

Les possibilités d'applications des motoréducteurs sont très variées, comme dans les domaines de la pharmaceutique, l'automobile, le BTP, la métallurgie... pour mettre en mouvement des : barrières, ponts, ascenseur, convoyeur, manège de parc d'attractions...

Pour comprendre comment s'opère la transformation de l'énergie électrique en énergie mécanique, il faut d'abord s'intéresser au fonctionnement du moteur asynchrone triphasé. L'architecture du moteur électrique repose sur deux éléments principaux, le stator et le rotor. Le stator est la partie fixe du moteur et le rotor est la partie mobile du moteur. Le stator supporte 3 enroulements (aussi appelé bobines ou phases), chacun décalé de 120° par rapport aux autres. Une fois alimentées, ces bobines produisent un champ magnétique tournant (aussi appelé champ tournant). Ce champ tournant va induire des courants dans le rotor. Le rotor va alors tourner dans le même sens que le champ tournant. En complément de ces deux éléments essentiels, le moteur électrique est constitué d'autres composants. L'arbre moteur, emboîté dans le rotor, permet de transmettre l'énergie mécanique générée par la rotation du rotor. Des roulements soutiennent l'arbre moteur à l'intérieur du corps du moteur électrique, permettant une rotation fluide en réduisant le frottement. Le corps du moteur électrique est composé de la carcasse et de ses flasques d'extrémité, qui encadre et protège le stator, le rotor, l'arbre moteur et les roulements. Des vis d'assemblage solidarisent le corps, assurant la robustesse et la compacité de l'ensemble.

L'arbre moteur dépasse de part et d'autre du corps du moteur électrique. D'un côté, un ventilateur est directement couplé à l'arbre et protégé par un capot, assurant le refroidissement en automatique ventilant le système pendant son fonctionnement. De l'autre côté, l'arbre moteur peut être couplé mécaniquement via un système de clavette à un système externe ou à un réducteur pour adapter la vitesse et la puissance selon les besoins. La boîte à bornes permettant l'alimentation du stator se greffe directement sur la carcasse. Des joints divers ainsi qu'un frein optionnel complètent la structure du moteur électrique.

Les réducteurs, quant à eux, sont composés d'une carcasse, d'un arbre de transmission, et d'un ensemble de pignons dimensionné selon le rapport de réduction et le couple souhaité. Cet ensemble de pignons permet de modifier la vitesse et la puissance transmises par le moteur, adaptant ainsi le mouvement à l'application spécifique.

Chez SEW USOCOME, toutes les pièces des moteurs et réducteurs, à l'exception des roulements et des joints, sont fabriquées en interne, depuis le coulage jusqu'à l'usinage.



Figure 1. Exemple de motoréducteur électrique fabriqué par SEW USOCOME

1.2.3 Organisation de la production des usines

Sur le territoire Français, SEW USOCOME possède principalement deux usines de production et une usine logistique. Les usines de productions ont des rôles spécifiques dans le processus de fabrication des éléments des moteurs électriques et réducteurs. L'usine de Forbach est principalement dédiée à la fonderie et à l'usinage des différents éléments de châssis des moteurs et réducteurs, ainsi qu'au coulage des paquets rotor. L'usine de Haguenau, elle, se concentre sur l'usinage des arbres moteurs/réducteurs et des roues dentées, en plus d'avoir la charge du bobinage des paquets stator. Les usines de production assurent ainsi les étapes initiales de fabrication et sont les principales garantes de la qualité du produit final.

Quant à l'usine logistique, située à Brumath, elle centralise les actions d'assemblage des motoréducteurs ainsi que le conditionnement et l'envoi des commandes aux clients. En effet, une fois que les composants sont produits dans les usines de Forbach et de Haguenau, ils sont envoyés à l'usine de Brumath et stockés dans un magasin central de très grande capacité. Là, les composants sont assemblés spécifiquement selon les commandes des clients avant d'être expédiés. L'usine logistique assure ainsi la coordination fluide de toutes les étapes finales de production, et est garante de la livraison efficace et ponctuelle des produits aux clients.

En complément de ses usines de production et de l'usine logistique, des centres de montages rapides et de service après-vente sont dispersés sur le territoire Français. Ils agissent comme point de proximité pour les actions de réparations et de maintenances aux plus proches des clients. Ces centres baptisés « Drive technology center » par le groupe SEW EURODRIVE sont placés de manière stratégique en France afin de couvrir le plus de territoire possible tout en gardant une réactivité optimale.

Du point de vue de leur organisation interne, chacune des usines de fabrication et l'usine logistique sont divisées en mini-usines. Chaque mini-usine est une unité autonome et spécialisée, dédiée à une partie spécifique du processus de fabrication ou à la production de certains éléments. Chaque mini-usine fonctionne de manière semi-indépendante, avec sa propre gestion des opérations, ses ressources dédiées et son plan de production. L'objectif de cette organisation est d'améliorer l'efficacité, la flexibilité et la réactivité de la production tout en maintenant une coordination efficace entre les différentes unités.

De plus, l'usine de Brumath se distingue des usines de production par sa certification 'usine vitrine de l'industrie 4.0'. Dans cette usine logistique, le lean manufacturing est intégré de manière plus poussée et innovante que dans les usines de production, car beaucoup plus récente. Dans son flux interne, toutes les pièces nécessaires à l'assemblage des motoréducteurs d'une commande sont transmises automatiquement à des opérateurs via des magasins automatisés en fonction des ordres de fabrication. Ces opérateurs placent ensuite les pièces fournies automatiquement par le magasin sur des plateaux, en fonction des ordres de fabrication des commandes. Ces plateaux sont ensuite transmis en fonction d'un planning d'ordonnement aux lignes d'assemblage des différentes mini-usines de manière fluide et totalement automatisée par des véhicules autoguidés (AGV), optimisant ainsi les flux, les processus et réduisant les temps d'attente ainsi que les erreurs de nature humaine. Dans un contraste frappant avec cette approche modernisée, les usines de production adoptent des flux plus classiques. Dans les usines de production, la gestion et la mise en œuvre des ordres de fabrication et des ordonnancements sont réalisées de manière conventionnelle, avec des équipes logistiques et

des flux de chariots pilotés par des hommes. Contrairement à la haute automatisation et à l'efficacité des AGV, les opérateurs interviennent activement dans le déplacement et la manipulation des pièces. Les processus, bien que fonctionnels, peuvent être plus sujets aux retards et aux erreurs humaines. La distinction entre ces deux approches souligne l'écart entre les méthodes de production traditionnelles et les avancées de l'industrie 4.0 adoptées par des usines comme celle de Brumath.

L'orchestration du flux entre les usines de production et l'usine logistique de SEW USOCOME représentée par la Figure 2 est également essentielle pour assurer une production efficace et une livraison ponctuelle des produits aux clients. Chaque usine joue un rôle spécifique dans le processus de fabrication global, et leur coordination, notamment grâce aux avancées de l'industrie 4.0 et du lean manufacturing. Ces avancées permettent d'améliorer la capacité à répondre aux demandes variées des clients tout en maintenant les objectifs stratégiques en terme qualité, de délais et de satisfaction client. La défaillance d'une de ces usines aurait un impact important sur la capacité de SEW USOCOME à fournir ses clients en temps et en heure. Il est donc primordial pour SEW USOCOME d'avoir des flux optimisés et des moyens de production robustes pour ne pas mettre en péril ses activités et permettre d'atteindre/maintenir ses objectifs stratégiques.

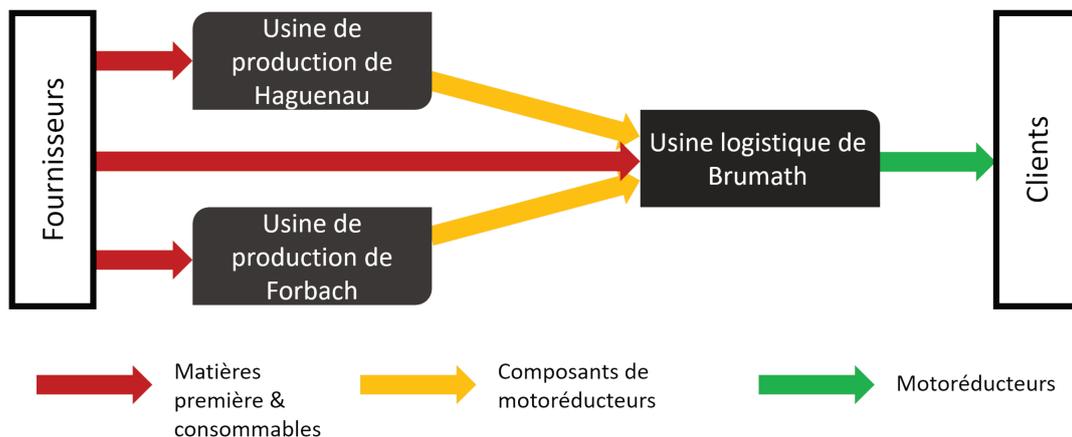


Figure 2. Représentation globale des flux interusines, de la matière première au client.

1.2.4 Les processus de production hétérogènes

À ce jour, cette organisation permet à SEW USOCOME d'atteindre ses objectifs de production, avec une capacité de production de 5000 ensembles motoréducteurs par jour, sans compromis sur la personnalisation du produit ou les délais de livraison. Face à un marché en pleine croissance, SEW USOCOME doit bientôt décupler sa capacité de production pour répondre à la demande accrue. Cela implique une maîtrise encore plus rigoureuse des processus de production afin d'assurer l'efficacité et la qualité malgré cette augmentation substantielle de volume. Cependant, la mise en œuvre des opérations de production chez SEW USOCOME révèle une variabilité significative au niveau des processus de production. En effet toutes les pièces des moteurs électriques et réducteurs sont fabriquées en interne, du coulage à l'assemblage final, en passant par l'usinage et le test des composants. Cette intégration complète génère une grande variété de processus de production. De plus, le produit est entièrement personnalisable par le client, ce qui entraîne une multitude de variations des processus de production dédiés à l'assemblage des différents composants des motoréducteurs. Pour un même élément, il existe plusieurs variations où certaines caractéristiques changent, telles que les dimensions, la matière, la couleur et l'usinage. Cela ajoute encore de la variabilité aux processus de production.

Afin de mieux comprendre cette variabilité, il est essentiel de distinguer les différents types de processus qui composent une ligne de production industrielle à SEW USOCOME. On distingue deux types de processus du point de vue des lignes de production : les processus de support et les processus opérationnels.

- Les processus de support assurent le flux interne des matériaux et la logistique dans une ligne de production sans modifier la forme du produit, en assurant des tâches telles que le transport, le stockage intermédiaire, et la gestion des équipements (Exemple, un processus de convoyage d'une pièce d'un point A vers un point B).
- Les processus opérationnels transforment directement le produit ou la matière première pour lui donner ses caractéristiques finales. Ils agissent sur le produit de manière concrète, en modifiant sa forme, sa composition ou ses propriétés (Exemple, un processus de perçage d'un flasque moteur).

Chez SEW USOCOME, la personnalisation poussée des produits rend les processus de production naturellement **hétérogènes**. Cette hétérogénéité se voit surtout dans les processus opérationnels. Prenons l'exemple d'un processus de taraudage, une opération qui permet de créer un filetage dans un matériau. Selon la matière transformée, le comportement du processus change. Par exemple, en taraudant de l'acier, le processus de taraudage demande plus d'énergie et prend plus de temps qu'en taraudant de l'aluminium, plus facile à travailler. Cette différence dans la consommation d'énergie et la durée de l'opération est un exemple de ce qui rend le processus hétérogène. Un même processus agit de manière différente en fonction de son contexte d'application.

L'hétérogénéité est également caractérisée par l'état des éléments physiques impliqués dans le processus, tels que les capteurs, les actionneurs et les pièces mécaniques. Au fil du temps, l'usure ou la dégradation de ces composants peut entraîner des changements dans le comportement du processus. Par exemple, un capteur de température qui perd en précision avec une déviation dans son étalonnage ou un actionneur qui devient moins réactif au fil du temps modifie le comportement du processus. Ces variations liées aux éléments physiques ajoutent une dimension supplémentaire à l'hétérogénéité : elles introduisent des variations qui ne dépendent pas seulement des spécificités du produit ou du matériau traité, mais aussi de l'état de santé des équipements utilisés.

Le caractère hétérogène d'un processus de production résulte donc des conditions opérationnelles et des caractéristiques évolutives des éléments physiques qui le composent.

Définition : **Un processus de production hétérogène** est un processus dont les caractéristiques et le comportement varient en fonction des spécificités des produits, des matériaux, et des paramètres de production. Cette variabilité peut être due à des différences dans les matières premières, aux ajustements de configuration, ou à l'usure des équipements qui le composent.

Aujourd'hui, chez SEW USOCOME, les processus de production hétérogènes sont représentés par des données numériques collectées via des automates industriels. Ces données incluent des variables automates représentatives de l'état des processus (état de fonctionnement, temps d'exécution, temps d'attente...), des valeurs issues de capteurs (mesurant température, vibration, courant, etc.), des valeurs issues d'actionneurs (paramètres de fonctionnement de l'actionneur, états...), des valeurs de capteurs/actionneurs et des variables contextuelles (mode de marche, références produites...). Ces données, numériques, peuvent se caractériser sous un des types de données de la **norme IEC 61131-3**. Cette norme qui définit les types de variables en automatisation inclut des valeurs logiques (vrai/faux), des nombres entiers et décimaux, des chaînes de texte, des dates et heures, des tableaux de valeurs similaires, des structures regroupant différents types de données, des listes d'options prédéfinies, et des sous-ensembles de valeurs pour des utilisations spécifiques. Ces informations sont organisées en séries temporelles qui capturent l'évolution des variables au fil du temps, avec une granularité pouvant aller jusqu'à 5 ms (200 Hz), la limite des automates utilisés.

Cette structuration en séries temporelles permet de représenter informatiquement les processus de production et d'observer leur hétérogénéité. L'évolution des données ne reflète pas seulement le déroulement du processus, mais également l'impact potentiel de l'usure des éléments physiques, comme les capteurs et actionneurs. Les dégradations ou défauts de ces composants influencent directement les données, affectant la performance ou l'efficacité du processus. Une dérive dans un processus de production peut donc provenir de l'usure des composants, d'un mauvais paramétrage, d'un changement de matériau, ou d'un défaut qualité initial, chacun pouvant entraîner des écarts de qualité, des défauts sur le produit final, ou des arrêts non planifiés de production.

Ainsi, ces séries temporelles offrent un potentiel de surveillance et d'analyse des processus, bien qu'aucune analyse directe de ces données temporelles ne soit actuellement mise en œuvre chez SEW USOCOME.

L'hétérogénéité des processus de production pose de nombreux défis pour améliorer et optimiser les lignes de production. En raison de la diversité des procédés et de la personnalisation des produits, la standardisation et l'automatisation des tâches sont limitées, rendant la formation des opérateurs coûteuse et exigeante en termes de compétences. La gestion des stocks et la planification de la production sont également compliquées par la variabilité des commandes, qui impose des ajustements constants.

Cette hétérogénéité complique également la détection des dérives des processus de production hétérogène. Cette difficulté à détecter les dérives affecte directement la capacité à maintenir des performances optimales et constantes des lignes de production.

Définition : **Une dérive de processus** est un écart significatif entre la performance attendue d'un processus et celle effectivement constatée, pouvant entraîner des défauts de qualité, des inefficacités ou des interruptions dans la production.

En outre, le contrôle de la qualité, la traçabilité des composants et la gestion logistique des flux deviennent d'autant plus complexes. Dans ce contexte, réduire les temps de cycle, optimiser les coûts et mettre en œuvre des méthodologies d'amélioration continue, comme le Lean Manufacturing ou Six Sigma deviennent des tâches ardues. L'adaptation aux variations de la demande, l'introduction de nouvelles technologies et la recherche d'efficacité énergétique sont également entravées par la diversité des processus et des équipements.

Pour relever ces défis et répondre aux besoins d'amélioration de la chaîne de production, SEW USOCOME s'appuie sur l'expertise de son service interne dédié à l'intégration des moyens de production : le Pôle Processus et Innovation.

1.2.5 Pôle Processus et Innovation : défis d'optimisation des moyens de production

Le Pôle Processus et Innovation occupe une place centrale au sein de SEW USOCOME, agissant comme intégrateur industriel et acteur majeur de l'innovation au sein de l'entreprise. Plus précisément, le Pôle Processus et Innovation est une entité pluridisciplinaire chargée de définir les flux des nouvelles usines, d'étudier et de concevoir les futures lignes de production, de les construire et de les intégrer harmonieusement dans le cadre opérationnel de l'entreprise pour s'aligner avec ses objectifs. L'amélioration continue des moyens de production et des flux est également une des missions du Pôle Processus et Innovation.

Ainsi, le pôle Processus et Innovation possède ses propres objectifs pour améliorer son efficacité et ainsi supporter les objectifs stratégiques de SEW USOCOME. En collaboration étroite avec les services méthodes, le premier objectif du Pôle Processus et Innovation est d'optimiser les processus de production afin d'assurer une efficacité opérationnelle maximale et une qualité de produit irréprochable. Le second objectif vise à garantir la continuité numérique entre les différentes strates de la pyramide ISA 95, permettant ainsi une intégration harmonieuse des systèmes de production, une gestion transparente de l'information et l'interopérabilité des systèmes de production. Le troisième objectif est celui de l'innovation, le Pôle Processus et Innovation s'attelle à explorer et intégrer les dernières avancées technologiques pertinentes, afin d'atteindre ses objectifs et de consolider la position de SEW USOCOME en tant que vitrine de l'industrie 4.0.

Dans le développement et la mise en place des solutions pour atteindre ses objectifs, le Pôle Processus et Innovation doit s'accorder avec certaines contraintes industrielles. Ces contraintes industrielles découlent directement des objectifs stratégiques et du contexte industriel de SEW USOCOME :

- **Contrainte d'hétérogénéité due à la variabilité des références et des volumes de production importants** : Comme énoncé dans la Section précédente, la nature variée des références et des volumes de production induit une hétérogénéité des processus de production.
- **Contrainte d'efficacité due à l'objectif de zéro retour client** : L'objectif ambitieux de zéro retour client concernant la qualité met une pression constante sur l'efficacité opérationnelle des processus de production.
- **Contraintes liées à la capacité de production** : L'objectif de 5 jours par commande met une pression sur la capacité de production et son efficacité pouvant être altérée par des manques de matières premières, des arrêts imprévus ou encore des perturbations de flux.

Les contraintes industrielles couplées aux objectifs d'innovation du Pôle Processus et Innovation ont permis de définir des axes d'innovation pour améliorer les lignes de production durant toutes les phases de leurs cycles de vie. Ces axes ont été formalisés sous la forme d'une feuille de route (roadmap) d'innovation centrée sur la standardisation de l'automatisme, les jumeaux numériques, la maintenance prévisionnelle et les systèmes de production reconfigurable.

La première étape de cette roadmap consiste à standardiser l'automatisme. L'objectif est d'optimiser l'efficacité de la phase d'automatisation, de la conception à la maintenance, en standardisant les composants utilisés, la structure générale des codes automates et la construction du nom des variables. Cette standardisation vise à accroître l'efficacité opérationnelle tout en facilitant la compréhension et la maintenance des lignes de production, réduisant ainsi les risques de dérives processus et de pannes de machines.

La deuxième étape de la roadmap concerne la mise en place de jumeaux numériques (Glaessgen and Stargel, 2012). Cette technologie permet d'obtenir des représentations virtuelles temps réel des lignes de production, facilitant ainsi la télémaintenance, la téléformation et la simulation hors ligne. Complémentaire à la standardisation de l'automatisme, les jumeaux numériques offrent des outils puissants pour anticiper et résoudre les problèmes potentiels, réduisant ainsi les perturbations de flux intra et inter mini-usines.

La troisième étape de la roadmap est dédiée à la maintenance prévisionnelle (Sahli, Evans and Manohar, 2021; Nunes, Santos and Rocha, 2023). Cette approche de maintenance permet une surveillance continue des équipements de production, anticipant ainsi les dérives et pannes potentielles. De plus, la maintenance

prévisionnelle autorise la construction de nouveaux indicateurs des lignes de production représentant leur santé et leur capacité de production en temps réel. Ceci doit renforcer la fiabilité opérationnelle et contribuer à maintenir les objectifs de qualité et de délais de SEW USOCOME.

Enfin, la quatrième et dernière étape de la roadmap s'appuie sur les trois premières pour proposer des systèmes de production reconfigurables (Morgan *et al.*, 2021). Les systèmes de production reconfigurables offrent une adaptabilité accrue des lignes de production en s'adaptant en temps réel aux besoins changeants selon le contexte de production. Par exemple, en cas de panne ou d'anticipation de dégradation d'une ligne de production, le système peut reconfigurer certaines lignes de production pour prioriser la réalisation de commandes et maintenir les objectifs de production. Cela implique notamment à SEW USOCOME de réordonnancer en temps réel les ordres de fabrication et de rediriger dynamiquement les véhicules autoguidés (AGV) pour optimiser les flux de production et minimiser les perturbations, en plus de modifier dynamiquement les programmes automatés des lignes de production concernés. Cette flexibilité renforcée doit permettre à SEW USOCOME de maintenir sa position compétitive en s'adaptant aux aléas en temps réel pour atteindre ses objectifs de temps de passage et de livraison.

Ainsi, cette roadmap de l'innovation offre une approche cohérente pour résoudre les problématiques industrielles liées à l'amélioration des lignes de production identifiées et rencontrées par SEW USOCOME.

Des recherches antérieures ont déjà abouti à une standardisation de l'automatisation grâce au développement d'une solution logicielle permettant de générer automatiquement du code automate à partir de la description formelle d'une ligne de production. Cette solution, basée sur une structure générique prédéfinie, a permis de standardiser les codes automatés ainsi que les noms des variables associées. Cette standardisation a été essentielle pour améliorer l'efficacité des équipes chargées de l'automatisation, faciliter la maintenance des lignes de production et aider les opérateurs à mieux comprendre leurs outils de travail grâce à des interfaces homme-machine et des types de données uniformisées au niveau des lignes de production, et ce à l'échelle de toutes les usines françaises.

Sur la base de ces travaux de standardisation, un système de jumeaux numériques a été développé et intégré à l'approche projet du Pôle Processus et Innovation (Chapelin, 2021). En adéquation avec les objectifs du Pôle Processus et Innovation (Chapelin *et al.*, 2022), ces travaux de recherches et développement ont été axés sur l'interopérabilité des systèmes, la continuité numérique et la réduction des redondances.

Par conséquent, la standardisation de l'automatisme et les jumeaux numériques ayant été introduits dans la démarche projet du Pôle Processus et Innovation et ayant atteint le niveau de maturité souhaité, cette thèse marque le début des travaux de recherche sur la troisième phase de la roadmap, à savoir l'intégration de la maintenance prévisionnelle. En effet, les outils de collecte de données et les nouvelles bonnes pratiques mises en place lors des deux premières phases de la roadmap forment un socle technologique et organisationnel solide pour cette troisième tâche.

1.2.6 Maintenance prévisionnelle : défis et enjeux industriels

La société SEW USOCOME se trouve actuellement à un carrefour crucial vis-à-vis de la maintenance prévisionnelle. Bien que la maintenance systématique soit déjà en place sur les lignes de production, elle s'avère insuffisante pour anticiper efficacement les pannes de certains processus. Un exemple illustratif lié aux processus d'usinage est le taraudage, une opération courante qui consiste à créer des pas de vis dans des pièces métalliques. Ce processus est particulièrement sensible aux variations et aux dégradations imprévues. Par exemple, l'usure progressive de l'outil de taraudage peut entraîner des défauts de filetage, affectant la qualité des pièces produites. De plus, des variations dans les matériaux ou les conditions de coupe peuvent accélérer l'usure ou provoquer des casses soudaines de l'outil. Ces incidents sont difficiles à prévoir avec une maintenance systématique, car les profils de dégradation peuvent varier considérablement en fonction des spécificités de chaque opération de taraudage. En ce sens (Bagheri, Yang, *et al.*, 2015), ont montré que la maintenance systématique basée sur des calendriers fixes est souvent sous-optimale, surtout pour les processus critiques comme l'usinage.

De plus (Monostori *et al.*, 2016) mettent en évidence que par une maintenance systématique, les outils sont remplacés selon un planning fixe basé sur les jours d'utilisation, sans prendre en compte leur utilisation réelle (nombre de pièces usinées, types de références usinées, conditions de coupe, matériaux travaillés, etc.). Ce planning, souvent établi avec un facteur de sécurité très élevé pour éviter les non-qualités, ne reflète pas précisément les conditions d'utilisation réelles. Cela conduit à une surconsommation d'outils de coupe et va à l'encontre de la politique de responsabilité sociétale de l'entreprise (RSE). En intégrant la maintenance prévisionnelle, il serait possible de maximiser la durée de vie des outils, réduisant ainsi le gaspillage et alignant mieux les pratiques de maintenance avec les objectifs RSE de l'entreprise. De plus, la maintenance prévisionnelle,

en surveillant en temps réel les paramètres critiques du processus de taraudage, doit permettre d'anticiper plus efficacement les pannes, assurant ainsi une meilleure continuité et qualité de production (l'intégration de modèles prévisionnels basés sur des données en temps réel permet de prolonger la durée de vie des équipements tout en alignant les pratiques industrielles avec les objectifs de durabilité (Bender and Sextro, 2021)).

Une étape préalable à la maintenance prévisionnelle peut être la maintenance conditionnelle (Teixeira, Lopes and Braga, 2020a), qui repose sur la surveillance des paramètres d'état des machines pour déclencher des interventions. Mais cette forme de maintenance présente également des limites, car, bien qu'elle soit plus réactive que la maintenance systématique, elle ne permet pas toujours de prévoir les pannes suffisamment à l'avance. En effet, la maintenance conditionnelle dépend de seuils préétablis et peut parfois ne détecter des problèmes qu'à un stade avancé de dégradation, réduisant ainsi la marge de manœuvre pour planifier et effectuer les réparations sans perturber la production.

En ce sens, les tentatives de mise en place de maintenance conditionnelle pour la maintenance des outils de coupe se sont toute soldée par un échec au sein de SEW USOCOME corroborant les propos de (Zhao *et al.*, 2019) expliquant que la maintenance conditionnelle, bien qu'efficace dans certains contextes, manque souvent de flexibilité pour s'adapter à des environnements de production très dynamiques, où les paramètres de dégradation varient fortement.

Plus globalement, cet exemple sur la maintenance des outils de coupe met en évidence toute la difficulté pour le Pôle Processus et Innovation de mettre en œuvre une maintenance prévisionnelle efficace et efficiente. En effet, généralement, la maintenance prévisionnelle est principalement déployée à l'échelle de composants, de flotte de composants ou de systèmes (Medina-Oliva *et al.*, 2014; Jimenez, Bouhmala and Gausdal, 2020; Kammerer *et al.*, 2020; Qi, Gryllias and Mauricio, 2021). Dans le contexte industriel de SEW, avec des centaines de lignes de production comprenant chacune des centaines de composants et de processus, implanter cette approche à grande échelle devient complexe sans automatisation.

L'ambition du Pôle Processus et Innovation est donc de proposer une solution « systématique » de maintenance prévisionnelle des lignes de production orientée processus, où la surveillance des processus et de leurs dérives permet d'anticiper les dysfonctionnements, les dérives et les dégradations de composants impliqués dans les processus surveillés. Cette vision de la maintenance prévisionnelle doit donc se référer à la norme ISO 13374 représenté par la Figure 3, et qui oblige à suivre les sept activités consécutives pour mettre en œuvre efficacement cette maintenance.

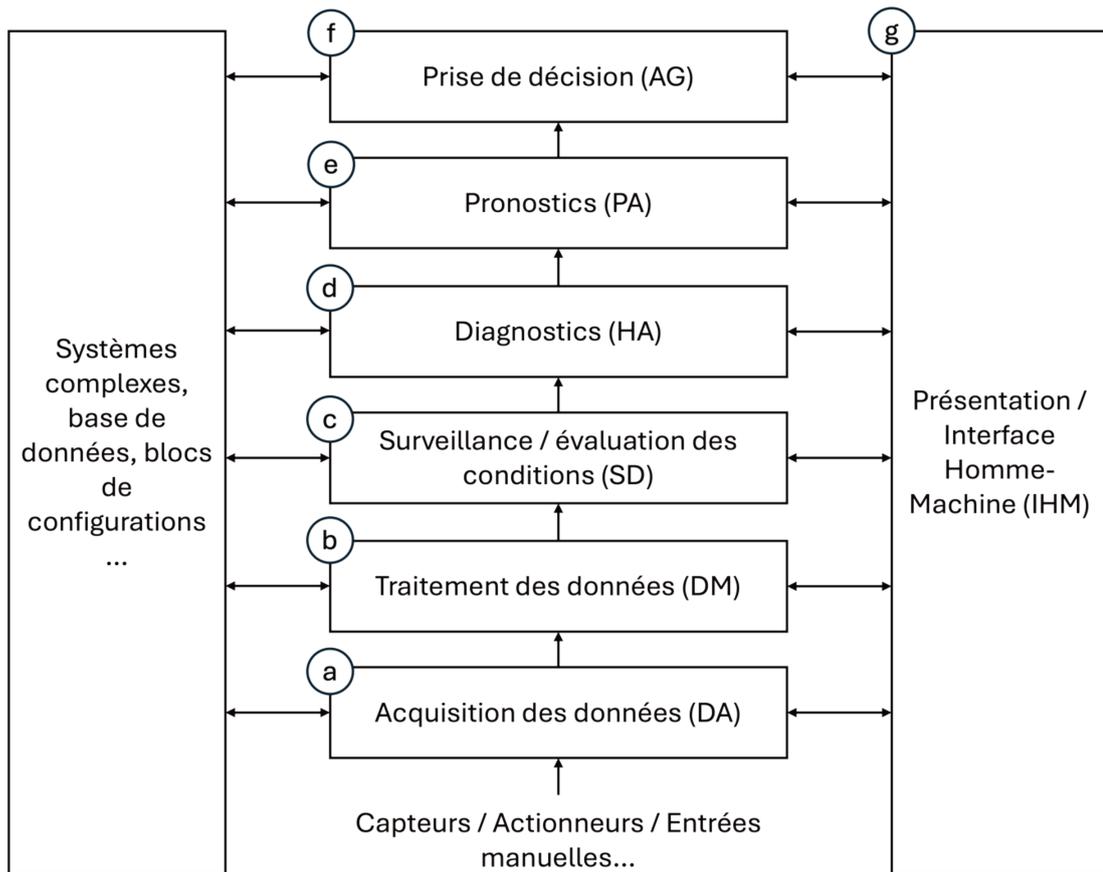


Figure 3. Représentation du traitement des données et flux d'informations de la norme ISO 13374

Ces étapes/activités, allant de l'acquisition des données à la prise de décision, fournissent une approche méthodologique robuste et générique pour orienter les efforts de maintenance prévisionnelle dans divers contextes industriels. En détail, ces sept étapes de la maintenance prévisionnelle sont :

1. **Acquisition des données (DA)** (Figure 3a) : Cette étape consiste à collecter les données pertinentes relatives au système ou aux composants à surveiller. L'objectif est d'obtenir des informations précises et complètes sur l'état actuel du système, notamment en ce qui concerne les paramètres opérationnels, les performances et le contexte.
2. **Traitement des données (DM)** (Figure 3b) : Une fois les données collectées, cette étape implique leur traitement et leur préparation pour l'analyse ultérieure. Cela peut inclure le nettoyage des données, la normalisation et la transformation pour assurer leur cohérence et leur qualité, facilitant ainsi les étapes suivantes de l'analyse.
3. **Surveillance / évaluation des conditions (SD)** (Figure 3c) : Cette étape vise à analyser en temps réel les données traitées afin d'identifier les signes précurseurs de dégradations ou de dysfonctionnements potentiels. L'objectif est de détecter les anomalies (Schmidl, Wenig and Papenbrock, 2022) ou les tendances anormales dans les données qui pourraient indiquer un risque accru de défaillance du système.
4. **Diagnostics (HA)** (Figure 3d) : Basé sur les résultats de l'évaluation des conditions, le diagnostic (Qiao and Brian A Weiss, 2018; Lei *et al.*, 2020) consiste à identifier la nature et la cause des problèmes détectés. Cette étape implique une analyse approfondie des données et des performances du système pour déterminer les causes sous-jacentes des dégradations ou des anomalies observées.
5. **Pronostics (PA)** (Figure 3e) : Les pronostics (Bender and Sextro, 2021) consistent à prédire l'évolution future de l'état du système en se basant sur les données historiques, les tendances observées et les modèles prédictifs. L'objectif est d'estimer la durée de vie restante (RUL – Remaining Useful Life) des

composants ou du système dans son ensemble, afin de planifier les actions de maintenance de manière anticipée.

6. **Prise de décision (AG) (Figure 3f) :** Sur la base des diagnostics et des pronostics, la prise de décision implique de déterminer les actions de maintenance appropriées à entreprendre. Cela peut inclure la planification proactive des remplacements de composants avant qu'ils n'atteignent un niveau d'usure critique, l'ajustement des paramètres de fonctionnement pour prolonger la durée de vie des équipements, ou encore la mise en œuvre de contrôles périodiques ciblés pour détecter les signes avant-coureurs de défaillances futures, par exemple.
7. **Présentation / Interface Homme-Machine (IHM) (Figure 3g) :** Enfin, la présentation ou l'IHM vise à communiquer les informations de chaque étape de l'ISO 13374 aux parties prenantes concernées. L'objectif est de fournir des informations claires et exploitables afin de soutenir la prise de décision et faciliter la gestion efficace de la maintenance prévisionnelle.

Cette norme, en plus de structurer les activités de la maintenance prévisionnelle, est un des éléments de référence de la communauté du PHM. En effet le PHM est un « cadre de référence » qui englobe l'ensemble des méthodes et technologies utilisées pour diagnostiquer l'état de santé d'un système, anticiper son évolution, et optimiser les actions de maintenance (Guillén Antonio J. and González-Prida, 2016). Le PHM permet donc de répondre aux défis posés par la maintenance prévisionnelle, en fournissant des solutions basées, entre autres, sur la collecte et l'analyse de données, les modèles prévisionnels, et les algorithmes de pronostics.

Ainsi, pour mettre en place la maintenance prévisionnelle chez SEW USOCOME, et plus précisément les activités de détection et de diagnostic de dérive, il semble pertinent de s'appuyer sur les travaux issus du PHM.

Par rapport à ces activités de la norme ISO 13374 et du PHM, le Pôle Processus et Innovation de SEW USOCOME a déjà déployé des moyens pour résoudre les problématiques rattachées à l'étape d'acquisition des données. Ceci s'est concrétisé par le développement d'un outil de collecte de données terrain précis et efficace. Les données collectables par l'outil incluent des informations sur les états des processus, l'état de tous les capteurs et actionneurs des lignes de production, et d'autres variables critiques. Ces informations sont essentielles pour établir un état de référence et suivre l'évolution des processus en temps réel.

Au-delà de cette première étape, aucun travail « validé » n'a été déployé en regard des autres étapes. Par conséquent, cette thèse, en cohérence avec les objectifs du Pôle, se positionne sur les étapes de surveillance et de diagnostic. En effet, la détection de dérive et son diagnostic constituent les piliers à dresser dans une démarche de maintenance prévisionnelle, car ils permettent de repérer les écarts de fonctionnement et d'en identifier les causes. Ces étapes sont fondamentales pour établir une compréhension précise de l'état du système, sur laquelle reposent les étapes suivantes, comme la prévision de l'évolution des dérives et la prise de décision sur les actions de maintenance.

1.2.7 Contraintes et verrous industriels pour une détection de dérive et un diagnostic pertinent

Vis-à-vis de ces deux piliers, plusieurs contraintes industrielles sont obligatoirement à considérer pour faire en sorte que les propositions de leurs fondements soient conformes aux besoins posés. Ces contraintes découlent des réalités technologiques, matérielles et organisationnelles de l'entreprise. Elles imposent, par conséquent, une orientation « des solutions » vers des choix de méthodes appropriées, de la gestion des ressources matérielles et des réseaux, ainsi que des exigences d'interopérabilité et de technologies ouvertes. En ce sens, les contraintes listées ci-dessous vont directement conditionner certains aspects du travail de recherche « en restreignant » potentiellement le domaine scientifique d'investigation.

1. **Utilisation imposée des méthodes d'apprentissage automatique :** SEW USOCOME impose l'utilisation de méthodes d'apprentissage automatique pour la détection des dérives dans les processus de production. Les alternatives plus classiques, telles que l'élaboration de modèles physiques ou de règles prédéfinies, sont exclues. En effet, ces méthodes nécessitent des ajustements manuels constants pour tenir compte des évolutions des processus surveillés. Ces ajustements permanents ont pour impact d'augmenter la charge de maintenance et la complexité opérationnelle des solutions de détection de dérive. Les méthodes d'apprentissage automatique, elles, s'adaptent dynamiquement aux changements sans nécessiter de redéfinir le modèle manuellement. La capacité à apprendre et à améliorer les modèles en continu à partir des données collectées réduit considérablement le besoin

d'intervention manuelle. Concernant les données collectées, elles incluent des variables automatiques telles que définies en Section 2.4.

2. **Contraintes matérielles orientées « hardware/software »** : Les solutions algorithmiques développées doivent impérativement pouvoir s'exécuter sur l'architecture matérielle standard de SEW USOCOME composée d'automates industriels et de PC à faible puissance de calcul. Ces infrastructures limitées imposent donc l'utilisation de logiciels et d'algorithmes légers, capables de traiter les données de manière efficace sans dépendre de ressources matérielles supplémentaires. Cette contrainte exclut l'utilisation de technologies gourmandes en ressources (CPU, GPU, mémoire...).
3. **Contraintes liées aux réseaux** : Pour des raisons de cybersécurité strictes, l'utilisation de technologies sans fil est formellement interdite. Toutes les communications système doivent être effectuées par des réseaux filaires pour garantir un contrôle total sur les flux de données. De plus, aucune solution basée sur le cloud ne peut être envisagée, les données doivent impérativement rester sur les serveurs internes de l'entreprise. Cela impose que toutes les opérations de traitement et d'analyse des données soient effectuées localement, sans recours à des infrastructures externes afin d'assurer une gestion sécurisée des informations de production considérées comme sensibles.
4. **Technologies ouvertes** : Toutes les solutions développées doivent être basées sur des technologies ouvertes, c'est-à-dire des outils et des logiciels qui ne sont pas propriétaires et qui reposent sur des standards accessibles à tous. Cette exigence vise à éviter toute dépendance vis-à-vis d'un fournisseur particulier ou d'une solution fermée. L'objectif est de permettre à l'entreprise de garder la maîtrise totale des développements, des évolutions, et de la maintenance en interne. L'utilisation de technologies ouvertes garantit aussi une meilleure flexibilité et assure la pérennité du projet, même en cas de changements technologiques ou économiques.
5. **Interopérabilité des systèmes** : Les solutions doivent être entièrement interopérables, c'est-à-dire capables de communiquer et d'échanger des données facilement avec les autres systèmes en place via des interfaces de programmation (API) dédiées. Aucune partie du système ne doit fonctionner comme une "boîte noire" où les données seraient inaccessibles ou non exploitables par d'autres applications. Cette interopérabilité est cruciale pour garantir une intégration fluide avec les infrastructures existantes et permettre une gestion centralisée et transparente des processus de maintenance prévisionnelle sans créer de silos technologiques.

Ces contraintes représentent par conséquent une typologie d'exigences à respecter pour la mise en œuvre pertinente des étapes de surveillance et de diagnostic de l'ISO 13374, qui vont nécessiter aussi pour le Pôle Processus et Innovation d'adresser plusieurs problématiques industrielles (ou verrous industriels) qu'il ne sait pas résoudre. Les verrous industriels majeurs attaqués par la thèse se réfèrent à :

1. **La collecte de données de dérives** : Il y a un manque de données historiques sur les dérives, car les processus de production ne sont pas surveillés de manière continue. Même lorsque les données sont collectées, il est difficile de les contextualiser et d'isoler les informations spécifiques aux dérives, compliquant ainsi leur exploitation pour les modèles de détection de dérive et de diagnostic.
2. **Le diagnostic des causes racines** : Diagnostiquer les causes racines des défaillances dans les lignes de production est difficile en raison de la complexité des interactions entre les processus et les composants.
3. **Le développement des modèles de détection et de diagnostic de dérive** : Le développement, la validation/vérification des modèles de détection et de diagnostic sont rendus complexes par la variabilité des conditions de production et le manque de savoir-faire spécifique en interne. Ce manque de compétences spécialisées constitue un obstacle majeur à la réussite du déploiement de la maintenance prévisionnelle.

À ces verrous majeurs s'ajoutent d'autres verrous industriels importants pour SEW USOCOME, mais qui ne font pas l'objet de ce travail de thèse borné sur 3 ans (ce sont des axes de perspective) :

4. **La gestion des coûts et retours sur investissement** : La mise en place de la maintenance prévisionnelle chez SEW USOCOME entraîne des coûts élevés liés à la recherche et développement, l'installation des systèmes, et leur maintenance continue. Il est essentiel d'évaluer ces dépenses et de réaliser des analyses de coûts-bénéfice pour s'assurer que les économies réalisées (réduction des pannes, optimisation des processus, qualité produit) justifient l'investissement. Sans cette évaluation, la viabilité économique de la maintenance prévisionnelle pourrait être compromise.
5. **L'intégration des données et interopérabilité** : Les systèmes de production de SEW USOCOME utilisent des technologies diverses, ce qui complique l'intégration et l'interopérabilité des données. Un standard de données commun (en lien avec la contrainte d'interopérabilité) doit être défini pour permettre la

collecte, la centralisation et l'exploitation cohérente des informations provenant de plusieurs sources. Sans une intégration aboutie, les données ne pourront pas être pleinement exploitées pour les analyses prédictives et la gestion efficace des processus de maintenance.

6. **La sécurité et confidentialité des données** : Il est crucial de repenser l'infrastructure et de s'assurer que la gestion des données respecte les normes de sécurité et de conformité (en lien avec la contrainte de sécurité réseau).

Les contraintes et verrous industriels identifiés chez SEW USOCOME sont essentiels et valides dans leur contexte spécifique, mais ils reflètent aussi des défis communs à de nombreuses industries évoluant vers l'industrie 4.0. Ces industries sont de plus en plus confrontées à des environnements de production complexes et évolutifs. Par exemple, la nécessité d'utiliser des technologies ouvertes, d'éviter le cloud pour des raisons de cybersécurité, ou encore de garantir l'interopérabilité des systèmes sont des problématiques majeures dans des secteurs comme l'aéronautique, l'automobile ou l'énergie. (Silvestri *et al.*, 2020; Kans and Campos, 2024) ont montré que l'implantation de la maintenance prévisionnelle est souvent freinée par la complexité des processus et l'interaction entre les composants dans l'industrie manufacturière au sens large.

De plus (C. Li *et al.*, 2024), montrent que le manque de données sur les modes de dégradation et les difficultés d'adaptation des modèles de maintenance prévisionnelle aux spécificités des équipements sont des obstacles importants dans l'énergie. Dans l'industrie automobile (Einabadi *et al.*, 2023) soulignent que les systèmes de production hétérogènes posent des défis similaires à ceux rencontrés chez SEW USOCOME, notamment en ce qui concerne la gestion des flux de données et leur intégration en temps réel pour des analyses efficaces.

Ainsi, les contraintes et verrous rencontrés chez SEW USOCOME ne sont pas vrais uniquement dans le contexte de l'entreprise. Ils sont représentatifs des défis que rencontrent de nombreuses industries (se devant d'évoluer vers des industries du futur) dans la gestion de la complexité des processus, la sécurité des données, les coûts, et l'intégration des systèmes. Cette situation souligne l'importance d'adapter les solutions de maintenance prévisionnelle aux besoins et aux réalités spécifiques de chaque secteur.

1.2.8 Résumé du contexte industriel

En synthèse, le contexte industriel de SEW USOCOME se caractérise par la production de motoréducteurs électriques dans un cadre de fabrication complexe et diversifié. Pour atteindre ses objectifs stratégiques (qui incluent la personnalisation des produits, la production à grande échelle, des délais de livraison courts et une qualité irréprochable) l'entreprise doit répondre à des contraintes opérationnelles liées à la complexité des processus de production, à la gestion de flux de production, et à la maintenance des équipements.

Afin d'assurer l'atteinte de ses objectifs stratégiques, SEW a élaboré un plan d'innovation. Ce plan vise à améliorer les performances de l'entreprise, notamment par l'introduction de nouvelles technologies. L'un des axes clés de ce plan est la mise en œuvre de la maintenance prévisionnelle à l'échelle des lignes de production via l'analyse de la dégradation des processus de production hétérogène. Toutefois, la variabilité des processus et la complexité des lignes de production rendent difficile la mise en place d'une telle maintenance prévisionnelle.

Ces difficultés ont été formalisées par SEW USOCOME sous forme de verrous industriels rattachés plus spécifiquement aux activités de détection de dérive et de diagnostic telles que proposées dans la norme ISO-13374 et le PHM.

En ce sens, un état de l'art sur les travaux actuels en PHM doit permettre d'identifier, si elles existent, les approches les plus adaptées pour détecter et diagnostiquer les dérives des processus de production hétérogènes et de mettre en évidence, si besoin, des travaux à mener pour « combler » scientifiquement les manques (voir contributions développées au Chapitre 2)

1.3 Contexte scientifique : PHM et Maintenance

1.3.1 Contexte scientifique général

Le PHM est un domaine de recherche interdisciplinaire qui cherche à fournir les outils méthodologiques et technologiques pour surveiller, anticiper et gérer l'état de santé des systèmes et des composants physiques (Qi, Gryllias and Mauricio, 2021). L'objectif global du PHM est d'optimiser la performance, la fiabilité et la durée de vie des systèmes en détectant leurs dégradations, en prévoyant les défaillances futures et en planifiant des

actions de maintenance appropriées. Les secteurs d'application du PHM sont multiples (aérospatial, santé, l'énergie, l'industrie au sens large ou encore dans la défense). (Vanek *et al.*, 2012; Olsson *et al.*, 2021; Es-sakali *et al.*, 2022)

Donc, globalement l'objectif du PHM est d'évaluer l'état actuel d'un système afin de diagnostiquer d'éventuels problèmes et faire une prévision de son état futur (Thomas *et al.*, 2021). Cette approche s'appuie sur une analyse des données représentatives passées, présentes et anticipées du système ou du composant surveillé, tout en prenant en compte des données contextuelles, comme les conditions environnementales, opérationnelles et d'utilisation (Weiss *et al.*, 2016; Bender and Sextro, 2021).

Au sein de la communauté scientifique, diverses études ont été menées pour appliquer le paradigme du PHM à différents niveaux d'échelle. Certaines se concentrent sur l'évaluation et la prédiction de la santé de composants individuels, telles que les études menées par (Xu *et al.*, 2022) sur les roulements dans les moteurs électriques. D'autres études explorent des systèmes multicomposants, analysant l'impact du remplacement d'un élément sur d'autres éléments et les implications logistiques de la maintenance associée (Dinh, Do and lung, 2020). Enfin, des recherches adoptent une approche multisystème, en exploitant des données de flottes similaires pour des pronostics basés sur les données (Medina-Oliva *et al.*, 2014).

Plusieurs cadres théoriques et méthodologiques ont été développés pour déployer le PHM de manière efficace. Parmi ces approches, l'architecture ouverte pour la maintenance conditionnelle (OSA-CBM) s'est imposée comme l'une des plus couramment utilisées (Teixeira, Lopes and Braga, 2020b; C. Huang *et al.*, 2024), en raison de son adaptabilité et de son interopérabilité avec divers systèmes. L'OSA-CBM est une architecture standardisée construite autour des étapes de la norme ISO-13374 qui offre une structure modulaire pour le déploiement de la maintenance conditionnelle (CBM) dans divers environnements industriels. Là où la norme ISO-13374 définit les blocs de fonctionnalité d'un système de maintenance prévisionnelle, ainsi que les entrées et sorties générales de ces blocs, OSA-CBM en propose une implantation opérationnelle. Ainsi, OSA-CBM ajoute des structures de données et définit des méthodes d'interface pour les blocs de fonctionnalité définis par la norme ISO-13374.

Cependant, au regard des travaux PHM orientés sur la formalisation des processus de production hétérogènes, beaucoup de problématiques restent encore relativement peu explorées. Par exemple, bien que de nombreuses études se soient concentrées sur l'évaluation et la prévision de la santé des composants et de systèmes individuels, comme dans les travaux de (Aldeoes Yasser N. and Gokhale, 2023) sur les machines rotatives, peu d'entre elles se sont penchées sur le passage à l'échelle des solutions algorithmiques existantes pour déployer la maintenance prévisionnelle à grande échelle sur des écosystèmes hétérogènes. De même, il existe à ce jour peu de travaux traitant de façon pertinente de la maintenance prévisionnelle pilotée par l'étude de la dégradation des processus de production industriels plutôt que par l'étude de la dégradation des composants eux-mêmes (au regard des contraintes de SEW USOCOME).

En outre, des analyses telles que celle proposée par (Vogl, Weiss and Donmez, 2015; Montero Jimenez *et al.*, 2020) cherchent à mettre en évidence les méthodes et meilleures pratiques pour le développement des systèmes PHM. Leurs travaux mettent en évidence l'importance des activités de diagnostic et de pronostic dans la conception de systèmes PHM efficaces et fiables. En consolidant les connaissances existantes et en proposant des lignes directrices pour les développements futurs, de telles initiatives contribuent à renforcer la base méthodologique de la PHM et à promouvoir son adoption dans un large éventail d'applications industrielles.

À partir de ce positionnement PHM global, la suite de ce chapitre a pour objet de se focaliser sur les travaux existants en lien avec les deux activités du PHM adressées dans cette thèse, à savoir la détection et le diagnostic de dérive.

1.3.2 Détection de dérive des processus de production hétérogènes

La détection de dérive est une étape clé de la maintenance prévisionnelle et du PHM qui consiste à détecter les comportements anormaux des processus ou système surveillé. Cette étape de détection de dérive s'intègre dans l'étape de surveillance de l'ISO 13374. Dans le contexte de cette thèse, la détection de dérive est faite sur les processus de production hétérogène tels que définis en Section 2.4. Dans les environnements de production industriels, les processus peuvent subir des dégradations dans leurs fonctionnements dus à l'usure mécanique des composants physiques qui les composent, les fluctuations des matières premières ou les changements dans les paramètres de production, par exemple. Détecter les dérives de processus doit permettre d'anticiper les défaillances des composants physiques, de maintenir la qualité des produits et d'optimiser l'efficacité des lignes de production.

1.3.2.1 Caractérisation des données de processus hétérogènes

Avant même de s'intéresser aux méthodes d'apprentissage automatique adaptées à la détection de dérive de processus de production hétérogènes (contrainte imposée par le contexte SEW USOCOME ; Section 1.2.7), il est nécessaire de caractériser les données disponibles, entre autres à l'aide d'indicateurs/métriques afin d'alimenter ces méthodes d'apprentissage.

1.3.2.1.1 Type de données étudié

Une analyse de l'état de l'art des travaux récents du PHM met en évidence que les séries temporelles sont les formats de données les plus aptes à représenter les processus de production hétérogènes. En effet, les séries temporelles permettent de suivre l'évolution dynamique de divers paramètres au fil du temps, ce qui est essentiel pour capturer la dégradation progressive des équipements et les changements dans le comportement des systèmes industriels. Ceci est corroboré par des travaux (Mateus *et al.*, 2022; Kashpruk, Piskor-Ignatowicz and Baranowski, 2023) montrant que pour un ensemble très varié de méthodes (statistiques, agrégées, fréquentielles ou par imagerie), les séries temporelles constituent souvent le point de départ incontournable pour modéliser des processus dynamiques.

Par exemple (Qi, Chu and He, 2018), démontrent que l'utilisation de séries temporelles brutes dans les systèmes industriels permet de capturer les tendances sous-jacentes, les fluctuations et les anomalies, facilitant ainsi la mise en place de stratégies de maintenance efficaces. De plus (Gao *et al.*, 2023) montrent que l'analyse statistique, bien qu'elle repose souvent sur des données agrégées ou des résumés statistiques dérivés de séries temporelles, permet de simplifier l'évaluation des systèmes complexes tout en gardant une trace des tendances générales et des comportements aberrants. Cette flexibilité est renforcée par l'utilisation d'analyses fréquentielles, comme le souligne (Ye *et al.*, 2023) qui utilisent la transformée de Fourier pour détecter des anomalies cycliques dans les données industrielles, tout en partant de séries temporelles brutes.

Par ailleurs, d'autres approches reposant sur des transformations ou des compressions de séries temporelles sont couramment employées. Par exemple (Islam, Pears and Bacic, 2018) utilisent des méthodes de wavelet pour capturer les phénomènes transitoires dans les données temporelles des machines industrielles, afin d'anticiper les pannes. Ces méthodes permettent d'analyser la fréquence des événements ou des signaux tout en conservant la dimension temporelle.

Un autre axe de recherche concerne l'intégration des séries temporelles dans des systèmes d'apprentissage automatique. (Mao *et al.*, 2020) démontrent que les modèles basés sur des réseaux de neurones récurrents (RNN) ou des réseaux neuronaux à mémoire longue et courte (LSTM) sont particulièrement efficaces pour traiter les séries temporelles dans le cadre du PHM, car ces modèles exploitent la séquence et l'historique des données pour affiner les prédictions de maintenance. De plus (Mejri *et al.*, 2024), montrent que des modèles d'apprentissage automatique peuvent être utilisés pour classer des états de santé de systèmes industriels après avoir dérivé des caractéristiques pertinentes à partir de séries temporelles.

Malgré cette diversité méthodologique, tout converge vers l'idée que, dans la plupart des cas, les séries temporelles constituent la base des analyses. Même lorsqu'on s'éloigne des séries temporelles brutes pour utiliser des modèles de réseaux bayésiens ou des analyses de corrélation entre sous-systèmes, comme illustré par (Castellano-Quero *et al.*, 2023), les données d'origine restent temporelles, transformées ensuite pour s'adapter aux besoins spécifiques du modèle.

En conclusion, les séries temporelles semblent constituer la représentation la plus adaptée pour capturer les dynamiques des processus de production hétérogènes. Ces séries temporelles permettent non seulement de suivre l'évolution des paramètres critique dans le temps, mais aussi de détecter des anomalies, d'anticiper des dérives, et de construire des modèles de maintenance prévisionnelle. De plus, que les données soient analysées brutes ou transformées (statistiques, fréquentielles, etc.), elles restent représentatives dans une certaine mesure de la série temporelle initiale, validant ainsi leur pertinence en tant que représentation principale pour les systèmes de production industriels.

Il est également important de distinguer deux types de séries temporelles étudiées dans les analyses du PHM : les séries temporelles univariées et les séries temporelles multivariées. Une série temporelle univariée capture une seule variable en fonction du temps. Cette approche peut être suffisante pour des systèmes simples et des processus monovariés. Toutefois, dans le cas des processus de production hétérogènes, les systèmes industriels complexes sont caractérisés par des interactions entre plusieurs variables simultanément mesurées.

Dans ce contexte, les séries temporelles multivariées, qui enregistrent plusieurs variables en parallèle sur la même période, sont plus adaptées pour représenter les données des systèmes industriels. En effet, des travaux récents confirment que les processus de production modernes impliquent souvent un ensemble de variables interdépendantes, et que ces interdépendances doivent être modélisées pour une détection efficace

des dérivées. Par exemple (Xie, Li and Zhu, 2024), démontrent l'efficacité des séries temporelles multivariées pour modéliser les processus industriels où plusieurs capteurs surveillent des paramètres différents, mais corrélés, comme la température, la pression, et les vibrations des machines. De même (Sürmeli and Tümer, 2020), montrent que les séries temporelles multivariées offrent une vision plus complète des systèmes complexes en intégrant des relations entre variables, ce qui améliore les performances des modèles de maintenance prévisionnelle par rapport aux modèles basés sur des séries univariées.

Ainsi, au vu de la nature hétérogène des processus de production industriels, il devient évident que les séries temporelles multivariées sont l'option à privilégier pour représenter les données dans cette thèse. Les séries temporelles multivariées permettent de capturer non seulement l'évolution temporelle de chaque variable, mais aussi les relations complexes et interdépendantes entre ces variables, donnant une image plus riche et plus représentative du système dans son ensemble.

1.3.2.1.2 Caractérisation des séries temporelles

Les travaux cités précédemment pour illustrer la pertinence des séries temporelles pour représenter les processus de production montrent aussi que pour maximiser l'efficacité des algorithmes, il est nécessaire d'adapter le choix des algorithmes en fonction des caractéristiques isolées des données (Lei *et al.*, 2020). Ainsi, plusieurs indicateurs de caractérisation des données sont à prendre en compte. Parmi ces indicateurs on retrouve notamment le **coefficient de corrélation de Pearson**, le **coefficient de corrélation de Spearman**, l'**Autocorrelation Function (ACF)** et le **volume des données**. Ces indicateurs aident à identifier les relations entre les variables, à évaluer les dépendances temporelles et à gérer la complexité des séries temporelles.

- Coefficient de corrélation de Pearson (Wilhelm Kirch, 2008)
Le coefficient de corrélation de Pearson est un outil statistique classique qui mesure la relation linéaire entre deux variables. Il est particulièrement utile dans le cadre des séries temporelles pour identifier des corrélations directes entre les variables. Une valeur élevée indique que les deux variables évoluent de manière similaire, ce qui peut être exploité pour la réduction de la dimensionnalité.
- Coefficient de corrélation de Spearman (Weisburd *et al.*, 2020)
Le coefficient de corrélation de Spearman, quant à lui, mesure les relations monotones, ce qui le rend adapté lorsque les variables ont des relations non linéaires. Il est utile pour les séries temporelles où des tendances ordonnées, mais non linéaires sont observées, permettant d'identifier des corrélations même si les relations entre les variables ne sont pas strictement linéaires.
- Fonction d'autocorrélation (ACF) (Aragon, 2011)
L'ACF mesure la dépendance d'une série temporelle avec ses propres valeurs à différents décalages dans le temps. Elle permet de capturer la persistance des motifs temporels et est essentielle pour comprendre si des structures répétitives existent dans les données.
- Volume des données
Le volume de données est un facteur crucial qui peut influencer la performance des algorithmes. Lorsqu'une série temporelle contient de nombreuses observations ou lorsque le nombre total de points dépasse une certaine limite, il devient nécessaire d'appliquer des techniques de réduction pour alléger la charge computationnelle.

1.3.2.1.3 Traitement des données en fonction des indicateurs

En fonction des valeurs de ces indicateurs, il est possible de déterminer s'il est nécessaire d'appliquer des transformations aux données pour maximiser les résultats des algorithmes d'apprentissage automatique en rendant les données plus adaptées aux spécificités des algorithmes. Ces transformations incluent des actions telles que la **réduction de dimensionnalité**, la **gestion des dépendances temporelles**, ou encore l'**optimisation du volume des données**.

A) Réduction de dimensionnalité (Fejjari, Saheb Etabaa and Korbaa, 2021) :

La réduction de dimensionnalité est souvent la première action à envisager lorsque les indicateurs révèlent une forte redondance entre les variables. En effet, lorsqu'une corrélation élevée est détectée (indicateur > 0.7) via le coefficient de corrélation de Pearson ou de Spearman, cela signifie que certaines variables transportent des informations similaires. Pour éviter que cette redondance n'alourdisse inutilement le modèle et les bases de données, il est conseillé de combiner ces variables.

L'Analyse en composantes principales (PCA) est l'une des techniques les plus communes pour la réduction de dimensionnalité dans ce type de situation. Comme l'ont montré Wu *et al.* (2020), le PCA permet de réduire la complexité des données tout en conservant la majorité de la variance présente dans le jeu de données.

En combinant les variables corrélées en composantes principales, le PCA diminue la charge computationnelle des algorithmes et améliore leur efficacité, tout en réduisant le risque de surapprentissage.

Cependant, lorsque les relations entre les variables sont non linéaires, détectées par exemple par un coefficient de Spearman élevé, il devient nécessaire d'utiliser des méthodes plus avancées. Des techniques comme le t-SNE ou les autoencodeurs peuvent être mis en œuvre pour capturer les structures complexes des données non linéaires. Gupta et al. (2022) montrent que ces méthodes permettent de maintenir l'intégrité des données tout en réduisant leur dimensionnalité, rendant ainsi les séries temporelles plus faciles à manipuler pour les algorithmes tout en préservant leurs caractéristiques essentielles.

B) *Gestion des dépendances temporelles* (Aragon, 2011)

Lorsque les indicateurs temporels, comme l'ACF, ont une forte persistance des motifs dans le temps (indicateur > 0.5), il est préférable de conserver les séries temporelles dans leur forme brute. Cela permet aux algorithmes de capturer les dynamiques temporelles et les régularités intrinsèques du processus étudié.

En revanche, lorsque l'ACF montre une faible dépendance temporelle (indicateurs < 0.5), il est préférable d'appliquer une transformation des données pour extraire des caractéristiques statiques plus exploitables. Dans ces situations, l'utilisation d'algorithmes comme les autoencodeurs devient plus pertinente, car ils se concentrent sur la reconstruction de caractéristiques globales des données sans se préoccuper des relations temporelles locales. Cela permet aux modèles de se focaliser sur les aspects structurels ou non linéaires des données, plutôt que sur la continuité temporelle.

C) *Réduction du volume des données* (Hoteit and Köhl, 2006)

Enfin, lorsque les séries temporelles deviennent volumineuses, des ajustements sont nécessaires pour éviter une surcharge computationnelle et une dilution des dérives dans un nombre d'échantillons très important difficilement capturable par les modèles, ainsi que pour garantir l'efficacité des algorithmes, étant donné que les techniques recommandées pour effectuer les actions de réduction du volume des données incluent le sous-échantillonnage ou l'agrégation des séries temporelles.

L'échantillonnage consiste à sélectionner une partie représentative des données pour en réduire la taille sans perdre les informations essentielles. Il peut s'agir de la sélection de points de données aléatoires ou à intervalles réguliers, afin de conserver la structure temporelle tout en réduisant la quantité de données à traiter.

L'agrégation, quant à elle, regroupe plusieurs points en une seule valeur statistique, comme la moyenne ou la médiane, pour résumer les périodes clés des séries temporelles. Cela permet de réduire le volume tout en conservant les tendances importantes, comme les variations journalières ou les anomalies locales.

Dans les cas où la réduction de la taille des séries pourrait altérer la qualité des données, une approche hybride peut être envisagée. L'utilisation conjointe de techniques de réduction de dimensionnalité, comme le PCA ou les autoencodeurs, avec l'échantillonnage permet de conserver une grande partie de l'information critique tout en facilitant l'analyse. Cela assure une scalabilité optimale sans compromettre les performances des algorithmes.

1.3.2.2 Méthodes d'apprentissage automatique pour la détection de dérive de processus de production hétérogène

Les séries temporelles multivariées ont été définies comme les données à surveiller, et des indicateurs pour les caractériser ont été présentés. Il est maintenant possible d'aborder l'état de l'art sur les techniques d'apprentissage automatique de détection de dérives adaptées à ce type de données. Il est à noter que, comme déjà mentionnée en section 1.2.7, l'utilisation de méthodes d'apprentissage automatique est imposée par l'industriel proposant cette thèse ... mais ce choix peut se justifier par rapport aux données à analyser dans le but de détecter et diagnostiquer des dérives.

1.3.2.2.1 Définition

Comme le montrent les résultats d'études, comme celles de (Fahle, Prinz and Kuhlenkötter, 2020a; Lei et al., 2020; Nacchia et al., 2021) les méthodes d'apprentissage automatique sont très répandues pour l'application des activités de détection de dérives et de diagnostic en PHM.

En effet, l'apprentissage automatique, aussi appelé Machine Learning (ML), apparaît comme une technique clé dans la maintenance prévisionnelle et le PHM. L'apprentissage automatique, sous-domaine de l'intelligence artificielle, permet dans le contexte du PHM d'analyser les données historiques pour détecter des dérives dans les performances des équipements et des processus afin d'anticiper des défaillances. L'apprentissage automatique utilise divers algorithmes pour apprendre à partir des données et faire des prévisions sur l'état des systèmes, détecter des dérives ou classifier des informations. L'apprentissage automatique est très efficace sur les cas simples et permet aux experts de comprendre et d'expliquer les

décisions prises par le modèle. Cependant, cela peut être limité en termes de performance dans des contextes plus complexes.

L'apprentissage automatique comporte un sous-domaine appelé l'apprentissage profond (deep learning) (Heaton, 2018). Contrairement à l'apprentissage automatique qui s'appuie sur des algorithmes statistiques ou mathématiques, l'apprentissage profond utilise des réseaux de neurones artificiels. Ces réseaux, composés de plusieurs couches, sont capables d'apprendre des représentations complexes et abstraites à partir d'un grand volume de données. L'apprentissage profond est souvent préféré dans des cas complexes où une grande quantité de données est disponible. Selon le contexte, l'apprentissage profond peut offrir des performances supérieures à celles de l'apprentissage automatique plus traditionnel. Toutefois, cette performance accrue s'accompagne d'une complexité plus élevée et d'une nature plus opaque dans les mécanismes de prise de décision. L'effet « boîte noire » de l'apprentissage profond rend les modèles moins interprétables et plus difficiles à paramétrer, maintenir et expliquer. (Ibs *et al.*, 2024)

Bien que l'usage de l'apprentissage automatique pour l'activité de détection de dérive soit demandé par l'industriel dans le cadre de cette thèse, il est nécessaire de s'interroger sur la légitimité et la pertinence de ce choix. En effet, dans le domaine de la détection de dérives, plusieurs approches alternatives existent :

Les méthodes statistiques : Ces méthodes, telles que le test CUSUM, le test de Page-Hinkley et le test de Kolmogorov-Smirnov, permettent de détecter des changements dans la moyenne, la variance ou la distribution des données. Elles sont sensibles aux dérives graduelles, mais leur efficacité dépend de l'hypothèse de stationnarité des séries temporelles, ce qui limite leur application dans les environnements industriels où les données sont souvent non stationnaires.

Les méthodes basées sur les seuils : Ces méthodes surveillent un indicateur (ex. : une erreur ou une distance statistique) et déclenchent une alerte lorsque cet indicateur dépasse un seuil prédéfini. L'approche Shewhart est un exemple classique. Bien qu'elles soient simples et couramment utilisées en maintenance, elles souffrent de la difficulté à définir des seuils adéquats et de leur sensibilité aux bruits dans les données. Des méthodes adaptatives ont été développées pour ajuster dynamiquement les seuils selon les variations observées, améliorant leur adaptabilité.

Les méthodes basées sur les modèles physiques : Ces méthodes reposent sur des connaissances expertes et des équations décrivant le comportement attendu du système surveillé. Elles comparent les prévisions basées sur le modèle aux données réelles pour détecter des écarts. Bien qu'elles soient précises, ces méthodes nécessitent une connaissance approfondie du système et peuvent être coûteuses en termes de calcul et de modélisation, en plus d'être affectées par l'usure naturelle des équipements.

Par conséquent, par rapport aux limitations des méthodes statistiques, des seuils et des modèles physiques, les techniques d'apprentissage automatique sont une solution pertinente pour la détection de dérives des processus de production hétérogènes. En effet, les caractéristiques propres aux systèmes de production industriels imposent des défis que les méthodes alternatives à l'apprentissage automatique ne peuvent pleinement relever. Les processus de production évoluent constamment : ajout de nouvelles références à traiter, modifications structurelles, changements dans les composants physiques. Cette variabilité remet en question les hypothèses de stationnarité, de stabilité des seuils ou de prévisibilité des modèles physiques, rendant ces méthodes moins adaptées.

L'apprentissage automatique apporte une réponse à cette caractéristique par sa capacité à apprendre des données pour générer le modèle de détection de dérive, ce qui ne peut être fait par les autres méthodes qu'à travers des modifications manuelles du modèle. Les méthodes statistiques, par exemple, exigent de recalibrer les paramètres ou de revoir les hypothèses de stationnarité chaque fois que les données évoluent. Cette dépendance limite leur adaptabilité et les rend peu pratiques dans des environnements en constante évolution. Les méthodes basées sur les seuils, quant à elles, demandent des ajustements manuels des seuils à chaque variation des conditions de production. Ce besoin d'intervention continue augmente le risque de fausses alertes et rend le suivi coûteux en temps, surtout dans des contextes où les changements sont fréquents. Enfin, les méthodes reposant sur des modèles physiques nécessitent des révisions profondes à chaque modification du processus ou des composants physiques. À chaque ajustement, il faut redévelopper ou adapter les équations, ce qui mobilise une expertise poussée et des ressources importantes pour maintenir la précision du système.

En termes d'avantages supplémentaires, les méthodes d'apprentissage automatique permettent de capturer automatiquement à partir des données des relations complexes et non linéaires entre les variables de production, répondant ainsi à la diversité des comportements observés. Les méthodes alternatives, elles, ne permettent pas d'automatiser cette modélisation et exigeraient de définir manuellement chaque relation complexe, ce qui serait à la fois chronophage et difficile à réaliser avec précision.

De plus, en l'absence de données historiques ou de connaissances précises sur les dérives, l'apprentissage automatique permet de créer des modèles de détection de dérive sur les données de bon

fonctionnement uniquement, même en faible quantité. Contrairement aux méthodes basées sur des seuils, qui nécessitent de définir manuellement des seuils spécifiques pour chaque indicateur et implique donc d'avoir des données ou de la connaissance experte sur les modes de dérives.

Enfin, l'apprentissage automatique permet de créer facilement des modèles de détection de dérive en utilisant simplement les données disponibles. Ainsi, il est possible de déployer des modèles de détection pour des milliers de processus sans intervention manuelle, car chaque modèle s'entraîne automatiquement avec les données spécifiques du processus qu'il surveille. En revanche, les méthodes basées sur des seuils ou des modèles physiques nécessitent de construire manuellement chaque modèle, ce qui serait complexe et chronophage pour détecter les dérives sur des milliers de processus. Avec l'apprentissage automatique, la détection de dérive à grande échelle devient simple et efficace.

Par conséquent, cette comparaison de l'apprentissage automatique avec d'autres alternatives méthodologiques vient conforter l'acceptation de la contrainte inhérente aux choix de l'apprentissage automatique pour la détection de dérives des processus de production hétérogène à grande échelle.

1.3.2.2.2 Les typologies d'apprentissage automatique

Les études de (Alshorman *et al.*, 2020; Fahle, Prinz and Kuhlenkötter, 2020b; Lei *et al.*, 2020; Montero Jimenez *et al.*, 2020; Mian *et al.*, 2024) cartographient des approches et des algorithmes existant pour l'implantation de modèles d'apprentissage automatique dans le contexte de la détection de dérives et du diagnostic au sens de la norme ISO 13374. De ces études il ressort qu'il existe différentes approches d'apprentissage : l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage semi-supervisé et l'apprentissage par renforcement.

- **Apprentissage supervisé** : Il repose sur l'utilisation de données étiquetées pour former des modèles prévisionnelles. Cela signifie que les états normaux et anormaux d'une machine sont connus et que le modèle est formé à partir de ces exemples pour apprendre à détecter les dérives dans le futur. En d'autres termes, un algorithme supervise l'apprentissage en fournissant des réponses correctes lors de la formation. Ce type d'apprentissage est très efficace lorsque l'on dispose de données historiques suffisamment riches et bien étiquetées (Khanzadeh *et al.*, 2018), permettant au modèle de reconnaître les schémas associés à une dérive ou à une défaillance.
- **Apprentissage non supervisé** : Dans ce dernier, les données ne sont pas étiquetées. L'algorithme doit donc identifier par lui-même des schémas ou des comportements à partir des données disponibles. Ce type d'apprentissage est utilisé lorsqu'il n'existe pas d'exemples précis de ce qui constitue un état normal ou anormal, et où l'algorithme doit découvrir des schémas cachés dans les données. Les méthodes non supervisées, comme le regroupement (clustering), permettent de regrouper les données en fonction de leur similarité et d'identifier des comportements qui ne correspondent pas à ceux attendus, ce qui peut signaler une dérive (Mühlbauer *et al.*, 2022).
- **Apprentissage semi-supervisé** : C'est une approche hybride qui combine une petite quantité de données étiquetées avec un grand volume de données non étiquetées. Cela est particulièrement utile dans les environnements industriels où il est coûteux ou difficile de collecter et d'étiqueter l'intégralité des données. L'algorithme utilise les données étiquetées pour apprendre des modèles initiaux et les données non étiquetées pour affiner et généraliser ces modèles, permettant ainsi de détecter des dérives même lorsque seules quelques informations sur les défaillances passées sont disponibles. (Wang and Wang, 2023)
- **Apprentissage par renforcement** : Quant à lui est un type d'apprentissage où un agent (algorithme) interagit avec un environnement et apprend par essais et erreurs. Cet agent reçoit des récompenses ou des punitions en fonction des actions qu'il entreprend, et l'objectif est de maximiser les récompenses à long terme. Dans le contexte de la maintenance industrielle, l'apprentissage par renforcement est utilisé pour optimiser les interventions de maintenance ou ajuster les paramètres des machines afin de prolonger leur durée de vie tout en minimisant les coûts d'entretien et les temps d'arrêt. Le modèle apprend progressivement quelles actions sont les plus bénéfiques pour maintenir les performances optimales des machines rotatives. Cette stratégie d'apprentissage ne s'applique généralement pas aux étapes de détection de dérive et de diagnostic au sens du PHM, car elle se concentre sur l'optimisation d'actions plutôt que sur l'identification de données de dérives et de causes racines. (Siraskar *et al.*, 2023)

En résumé, l'**apprentissage supervisé** repose sur l'utilisation de données étiquetées, permettant de prédire efficacement les dérives grâce à des exemples clairs d'états normaux et anormaux. Cependant, il nécessite des données représentatives et bien étiquetées. L'**apprentissage non supervisé**, quant à lui, est utilisé lorsque les données ne sont pas étiquetées, ce qui oblige l'algorithme à découvrir des schémas ou des anomalies

de manière autonome. Il est particulièrement utile lorsque les dérives possibles ne sont pas connues à l'avance. L'**apprentissage semi-supervisé** combine les approches supervisées et non supervisées, exploitant à la fois des données étiquetées limitées et un grand volume de données non étiquetées. Enfin, l'**apprentissage par renforcement** permet à un algorithme d'optimiser les interventions de maintenance et les paramètres des machines en apprenant par essais et erreurs, maximisant ainsi la durée de vie des équipements et minimisant les coûts et les temps d'arrêt. Ensemble, ces techniques permettent une gestion proactive de la maintenance, améliorant la fiabilité des machines rotatives tout en optimisant leur performance.

1.3.2.2.3 Approche pour la détection de dérive

De ces mêmes études concernant l'apprentissage automatique, il ressort qu'il existe plusieurs approches de détection de dérive de processus hétérogènes finis. Parmi elles on retrouve : la classification, le regroupement et la détection de nouveautés.

La **classification** (Galar *et al.*, 2011; Jha, Dave and Madan, 2019) est une méthode supervisée qui repose sur l'étiquetage préalable des données. Un modèle de classification est entraîné à partir de données étiquetées, où des exemples de séries temporelles normales et anormales sont fournis. Le modèle apprend à distinguer les différents profils de données étiquetés, tels que "normal", "en dérive" ou "défaillant" par exemple. Une fois entraîné, le modèle peut être utilisé pour classer de nouvelles séries temporelles non étiquetées en fonction des catégories apprises. Cette approche est particulièrement puissante lorsqu'il existe un historique riche de données étiquetées sur les comportements de la machine/processus. Cela permet d'identifier avec précision des dérives connues. Cependant, la principale faiblesse de la classification est qu'elle ne peut détecter que des dérives similaires à celles rencontrées lors de l'apprentissage. Si un nouveau type de dérive apparaît, non observé lors de l'entraînement, le modèle ne pourra pas le reconnaître, limitant ainsi sa capacité à détecter des comportements inédits.

Le **regroupement (clustering)** (Saxena *et al.*, 2017; Sürmeli and Tümer, 2020), quant à lui, est une technique non supervisée qui regroupe les données en fonction de leurs similarités sans que des étiquettes soient préalablement attribuées. L'algorithme apprend à regrouper les segments de données ayant des caractéristiques similaires, et toute série temporelle qui s'éloigne de ces regroupements standards est alors considérée comme potentiellement anormale, signalant ainsi une dérive. L'un des avantages majeurs du regroupement est qu'il ne nécessite pas de données étiquetées, ce qui est utile dans des environnements industriels où l'acquisition de données étiquetées est coûteuse ou difficile. Il permet également d'identifier des schémas cachés dans les données et de repérer des dérives sans avoir une connaissance préalable des types de défaillances possibles. Cependant, le regroupement n'offre pas d'étiquettes explicites aux données : il se contente d'indiquer que certaines séries temporelles ne correspondent plus aux schémas habituels, ce qui nécessite une analyse supplémentaire pour interpréter ces anomalies et confirmer la présence d'une dérive réelle.

Enfin, la **détection de nouveauté** (Ma and Perkins, 2003; Pimentel *et al.*, 2014) est une technique semi-supervisée/non supervisée axée sur l'identification de comportements ou de schémas anormaux qui n'ont jamais été observés auparavant. Cette approche est particulièrement efficace pour la détection de dérives inédites, où l'algorithme apprend à partir de données historiques étiquetées de ce qui constitue un comportement normal. Lorsqu'une nouvelle série temporelle s'écarte significativement de ce modèle de normalité, elle est signalée comme une nouveauté, indiquant une dérive potentielle. Cette technique est idéale pour repérer des comportements inattendus, et permet d'anticiper des dégradations que les autres méthodes supervisées ne pourraient pas détecter faute de données étiquetées adéquates. Cependant, l'un des défis de la détection de nouveauté est la gestion des faux positifs : dans certains cas, l'algorithme peut signaler des comportements légèrement différents qui ne sont pas nécessairement des dérives dangereuses, ce qui peut entraîner des alertes excessives.

En résumé, la **classification** est particulièrement adaptée lorsque des données étiquetées sont disponibles et que les types de dérives sont bien définis, mais elle est limitée face aux nouvelles dérives. Le **regroupement** permet de détecter des changements subtils sans avoir besoin d'étiquettes, mais il nécessite une analyse approfondie pour interpréter les résultats. La **détection de nouveauté**, quant à elle, est idéale pour identifier des comportements totalement nouveaux, bien qu'elle puisse parfois entraîner des faux positifs. Ces trois méthodes offrent des approches complémentaires pour la surveillance des séries temporelles dans des environnements industriels complexes et la détection précoce des dérives avant que celles-ci ne se transforment en pannes.

1.3.2.2.4 Algorithmes de détection de dérive

La thèse portant spécifiquement sur les activités de détection de dérive au sens du PHM, le Tableau 1 regroupant les algorithmes de détection de dérive identifiés durant l'état de l'art du PHM est proposé afin de compléter l'état de l'art des techniques d'apprentissage. Le tableau présente les algorithmes en précisant leur domaine d'application (machine learning ou deep learning), ainsi que le type d'apprentissage (supervisé, non supervisé, ou les deux). Il indique également si l'algorithme est capable de réaliser des tâches de classification, de détection de nouveauté (anomalies ou dérives), et de regroupement, c'est-à-dire de regroupement de données similaires. Ces différentes informations permettent de mieux comprendre les capacités spécifiques de chaque algorithme identifié.

Cet état de l'art des algorithmes synthétisé par le Tableau 1 met en évidence qu'il existe plusieurs travaux du PHM qui traitent déjà de la détection de dérive des processus de production hétérogène. En revanche, des travaux comme ceux de (Vanek *et al.*, 2012; Alshorman *et al.*, 2020; Bourgana *et al.*, 2021; Hiruta *et al.*, 2021; Rehab *et al.*, 2021; Es-sakali *et al.*, 2022; Xu *et al.*, 2022) sont tous optimisés pour un seul composant ou processus. En effet l'état de l'art manque de généralisation. Les méthodes de détection de dérive sont très souvent calibrées et utilisées dans un contexte bien particulier, sur des données bien maîtrisées et contextualisées. Très peu de travaux proposent des solutions capables de s'adapter à la variabilité inhérente aux processus de production hétérogène et manque de données concernant les profils de dérives (voir Section 1.2.4 et 1.2.5). **Une première question scientifique concerne donc le développement ou l'adaptation de méthodologies spécifiquement adaptées à la surveillance de processus hétérogènes dans un contexte pauvre en données représentatives des modes de dérives des processus de production industriels.**

Première problématique scientifique

- ***Comment développer une méthode de détection des dérives de processus de production hétérogènes ne nécessitant pas de données étiquetées et suffisamment évolutive pour s'adapter à toute dérive apparaissant tout au long de la vie du processus ?***

Tableau 1. Algorithmes d'apprentissage automatique appliqués à la détection de dérive de série temporelle utilisée dans le domaine du PHM

Nom d'algorithme	Type d'apprentissage	Classification	Détection de nouveautés	Regroupement	Références
Réseaux de neurones convolutifs (CNN)	Supervisé	✓			(Canizo <i>et al.</i> , 2019; Ruan <i>et al.</i> , 2023; Zhang, Liang and Liu, 2024)
Réseaux de neurones récurrents (RNN)	Supervisé	✓			(Canizo <i>et al.</i> , 2019; Mohammadi <i>et al.</i> , 2023; Al-Selwi <i>et al.</i> , 2024)
Réseaux de neurones à base de capsules (CapsNet)	Supervisé	✓			(D. Li <i>et al.</i> , 2024; Lyu <i>et al.</i> , 2024; S. Li <i>et al.</i> , 2024)
Autoencodeurs (AE)	Non supervisé, Supervisé		✓		(Ren <i>et al.</i> , 2018; Chen <i>et al.</i> , 2019; Xu <i>et al.</i> , 2022)
Deep Support Vector Data Description (Deep SVDD)	Non supervisé, Supervisé	✓	✓		(Chan, Han and Pan, 2024; M. Li <i>et al.</i> , 2024)
Long Short-Term Memory (LSTM)	Supervisé	✓			(Dong, Li and Sun, 2017; Luo, Huang and Zhou, 2018; Aydemir <i>et al.</i> , 2021)
Variational Autoencoders (VAE)	Non supervisé		✓		(Yang, Yu and Wang, 2021; Yu and Li, 2021; Jing <i>et al.</i> , 2022)
Arbres de décision	Supervisé	✓			(Tso and Yau, 2007; Dhakar, Singh and Gupta, 2023; El Mawas, Cappelle and El Badaoui El Najjar, 2023; Yıldız and Kalayci, 2024)

Forêts aléatoires (RF)	Supervisé	✓			(Breiman, 2001; Wescoat, Krugh and Mears, 2021; K. Huang <i>et al.</i> , 2024; Yan <i>et al.</i> , 2024)
K-plus proches voisins (KNN)	Supervisé	✓			(Sarmadi and Karamodin, 2020; Hamidi, Robertson and Ilow, 2023)
Machines à vecteurs de support (SVM)	Supervisé	✓			(Xiao <i>et al.</i> , 2016; Abedi, Aghamohammadi and Ameli, 2022; B. Wang <i>et al.</i> , 2024)
Gaussian Mixture Models (GMM)	Non supervisé		✓	✓	(Liu <i>et al.</i> , 2019)
Isolation Forest (iForest)	Non supervisé		✓		(Ding and Fei, 2013; Safari, Sabahi and Oshnoei, 2024)
One-Class SVM (OC-SVM)	Non supervisé		✓		(Xiao <i>et al.</i> , 2016)
Principal Component Analysis (PCA)	Non supervisé		✓	✓	(Cao <i>et al.</i> , 2021)
Ensemble Learning	Supervisé	✓			(Ngo, Beard and Chandra, 2022; Khan, Chaudhari and Chandra, 2024; Mian <i>et al.</i> , 2024)
K-means	Non supervisé			✓	(Chen <i>et al.</i> , 2024; H. Wang <i>et al.</i> , 2024; Wu, 2024)
DBSCAN	Non supervisé			✓	(Hajihosseini, Maghsoudi and Ghezelbash, 2024; Raja <i>et al.</i> , 2024)

1.3.3 Diagnostic des dérives des processus de production hétérogènes

En complément à l'étape de détection des dérives, l'étape de diagnostic consiste à identifier les causes racines des dérives détectées au cours de l'étape de détection de dérive. D'après les travaux les plus récents du PHM concernant le diagnostic, il existe trois types de modèle de diagnostics : les modèles basés sur la physique (PBM) (Figure 4a), les modèles basés sur les connaissances (KBM) (Figure 4b) et les modèles basés sur les données (DDM) (Figure 4c). Ces modèles peuvent être utilisés de manière isolée ou en collaboration. Lorsqu'un seul modèle est utilisé, on parle d'une approche par modèle unique (SM). Les approches SM sont simples à implanter et à interpréter, mais peuvent manquer de flexibilité face à la complexité croissante des systèmes de production industriels. Lorsque plusieurs modèles sont utilisés en collaboration, on parle d'approche multimodèle (MM). Les MM cherchent à combiner les avantages des différents types de modèles (PBM, KBM, DDM) pour surmonter les limites individuelles de chaque méthode. L'ensemble des interactions MM possibles est représenté par la Figure 4.

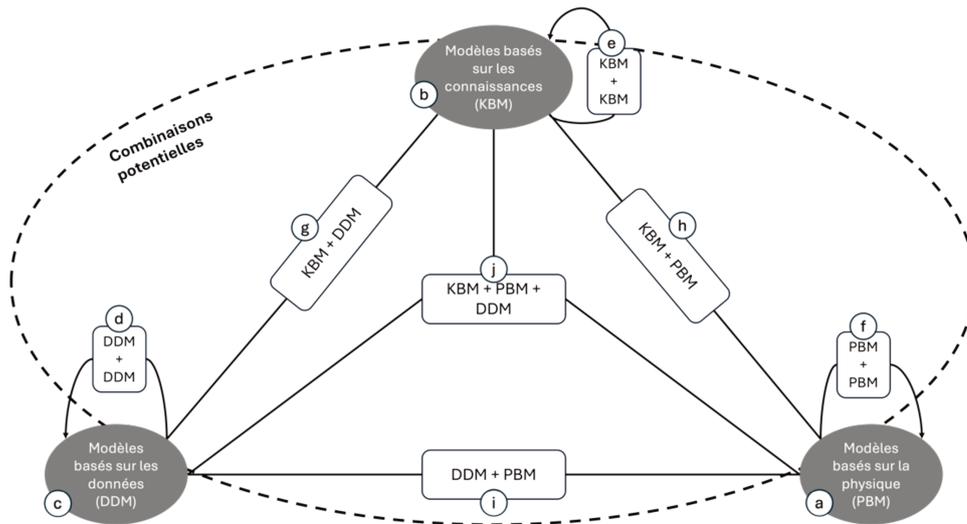


Figure 4. Combinaisons potentielles des méthodes de diagnostics adaptés de (Montero Jimenez *et al.*, 2020)

1.3.3.1 Modèles basés sur la physique (PBM)

Les PBM utilisent les lois de la physique pour évaluer la dégradation des composants. Les PBM demandent des compétences poussées en mathématiques et en physique pour leur développement. Avec une modélisation de la physique précise d'un système, il est possible d'effectuer des simulations qui permettent d'étudier le comportement de dégradation de composants ou de processus. Dans la littérature du PHM, on trouve des applications des PBM pour différentes applications de diagnostic. Notamment les travaux de (Vachtsevanos *et al.*, 2007; Nasution, Sævik and Gjøsteen, 2012) qui modélisent des PBM pour diagnostiquer la fatigue et la propagation de fissures sur des composants. (Climente-Alarcon *et al.*, 2017; Li *et al.*, 2018) appliquent les PBM pour diagnostiquer la défaillance de machine rotative et la défaillance d'électrovannes. (Cholette *et al.*, 2019) utilise les PBM pour diagnostiquer l'encrassement de tube échangeur dans des chaudières. Enfin (Qiao and Brian A. Weiss, 2018) diagnostic l'état de dégradation de robots industriels avec les PBM.

Les PBM offrent la possibilité de créer des modèles de diagnostic très précis. Cependant, cette précision a un coût. Les PBM nécessitent une connaissance approfondie du système, ce qui peut poser un problème en présence de dérives non documentées ou inconnues. En effet, les modèles physiques ne peuvent diagnostiquer correctement que les phénomènes déjà intégrés dans le modèle. De plus, le développement et la calibration de ces modèles sont souvent coûteux et requièrent un investissement en temps et en expertise conséquent. Les travaux de (Downey *et al.*, 2019; Jin and Ni, 2019; Aivaliotis *et al.*, 2021) soulignent ses limitations liées à la complexité croissante des systèmes complexes où les phénomènes physiques deviennent de plus en plus difficiles à modéliser.

1.3.3.2 Modèles basés sur la connaissance (KBM)

Les modèles basés sur les connaissances reposent sur une connaissance experte acquise par l'expérience. En effet, la connaissance experte des systèmes accumulés au fil des années d'exploitation et de

maintenance du système technique peut être représentée par des règles. Ces règles sont ensuite utilisées pour identifier les défauts, décrire la dégradation et anticiper une défaillance potentielle des composants, systèmes ou processus surveillés. Des publications telles que (Majstorović and Milačić, 1990; Freyermuth, 1991; Vingerhoeds *et al.*, 1995) décrivent l'utilisation des KBM pour réaliser des diagnostics dans des systèmes industriels.

Pour le KBM on retrouve trois grands types de modèles : les modèles à base de règles, les modèles à base de cas et les modèles à base de connaissances floue.

- **Modèles à base de règles** : Dans ces modèles, la connaissance est représentée par des règles sous la forme « SI-ALORS », permettant d'effectuer une inférence censée simuler un mécanisme de raisonnement humain simplifié (Aimiyekagbon *et al.*, 2021). Les systèmes à base de règles se composent d'une base de connaissances contenant toutes les règles, d'un moteur d'inférence et de données d'entrées. L'inférence est un processus itératif. Le moteur d'inférence compare les données d'entrées avec l'ensemble de règles de la base de connaissances et produit des conclusions comme sorties. Ce processus se répète jusqu'à ce que le raisonnement soit terminé.
- **Modèles à base de cas** : Ces modèles représentent la connaissance sous forme de cas, obtenus à partir de situations de problèmes concrets rencontrés précédemment. Lorsqu'un nouveau problème survient, le cas le plus similaire est extrait de la base de cas et sa "solution" est réutilisée pour résoudre le nouveau problème.
- **Modèles flous basés sur les connaissances** : Les systèmes à base de connaissances floue utilisent essentiellement le même format de règles « SI-ALORS » que les systèmes à base de règles, mais avec une logique floue qui permet d'introduire un degré d'appartenance dans les inférences, ce qui est utile pour les environnements complexes où la précision est difficile à obtenir (Sikorska, Hodkiewicz and Ma, 2011).

Dans les faits, les KBM se distinguent par leur capacité à intégrer facilement les connaissances humaines, ce qui les rend particulièrement utiles dans des contextes où les données sont limitées et où il existe une forte expertise humaine. Cependant, leur performance est intrinsèquement liée à la qualité et à l'étendue de la base de connaissances disponibles. Ainsi, dans des situations non anticipées ou dans des systèmes évolutifs, où la connaissance des causes racines est incomplète et évolue constamment, ces modèles peuvent manquer de flexibilité. En effet, des études, comme celles de (Medina-Oliva *et al.*, 2014), soulignent la nécessité d'une mise à jour continue des bases de connaissances pour maintenir la pertinence et l'efficacité des diagnostics fournis.

1.3.3.3 Modèles basés sur les données (DDM)

Les DDM ont gagné beaucoup d'importance ces dernières années, notamment grâce à la disponibilité accrue de la puissance de calcul et à la production de grandes quantités de données provenant des systèmes industriels modernes comme préconisés par l'industrie 4.0. En effet, ces systèmes intègrent un grand nombre de paramètres opérationnels mesurés et enregistrés parfois en continu. Le volume important de données qui en résulte peut être utilisé, de manière explicite ou implicite, à diverses fins, y compris la maintenance prévisionnelle. Les informations obtenues à partir de ces données peuvent être utilisées pour étudier la dégradation des composants et diagnostiquer l'état de santé actuel du système. D'après l'état de l'art, pour les DDM trois méthodes de diagnostic sont mises en avant : les modèles statistiques, les modèles stochastiques et les modèles d'apprentissage automatique.

- **Modèles statistiques** (Fassois and Kopsaftopoulos, 2013a): Ces modèles considèrent les données issues de capteurs comme des variables aléatoires dont on peut calculer les statistiques ou effectuer des tests sur la base d'historiques. Pour la maintenance prévisionnelle, ils permettent d'évaluer la dégradation et la durée de vie restante du système en comparant leur comportement actuel avec des comportements connus représentés par des séries de données. Des méthodes comme l'analyse de régression, les modèles autorégressifs et les modèles bayésiens sont couramment utilisés dans ce cadre.
- **Modèles stochastiques** (Bagheri, Siegel, *et al.*, 2015): Les processus stochastiques, tels que les processus de Wiener, les chaînes de Markov ou les processus de Lévy, sont utilisés pour modéliser l'évolution des variables aléatoires dans le temps. Ces modèles sont particulièrement adaptés pour la modélisation de la dégradation progressive des systèmes techniques et l'estimation de la durée de vie restante, car ils

permettent de modéliser des phénomènes aléatoires complexes qui affectent la performance des systèmes.

- **Modèles d'apprentissage automatique (machine learning)** (Fahle, Prinz and Kuhlenkötter, 2020a): Ces modèles sont capables de gérer et de capturer des relations complexes entre les données, souvent difficiles à obtenir à l'aide de modèles physiques ou statistiques. Les réseaux de neurones artificiels (ANN), y compris les perceptrons multicouches (MLP), les réseaux récurrents (RNN) et les réseaux convolutionnels (CNN), sont des exemples courants utilisés dans les applications de maintenance prévisionnelle. Ils permettent de diagnostiquer des pannes et de prédire la durée de vie restante en analysant des volumes massifs de données générées par les systèmes industriels.

L'un des avantages majeurs des DDM est leur capacité à identifier les données, sans nécessiter une compréhension approfondie des relations physiques sous-jacentes. Cela les rend particulièrement adaptés à des environnements de données massives où les relations entre les variables sont complexes et non linéaires.

Toutefois, les DDM sont dépendants de la qualité et de la quantité de données disponibles. Une mauvaise qualité des données peut entraîner des résultats de diagnostic erronés, et une quantité insuffisante de données peut limiter la capacité du modèle à généraliser correctement. De plus, la gestion de l'incertitude provenant des données (Jardine, Lin and Banjevic, 2006; Sankararaman, Daigle and Goebel, 2014; Lei *et al.*, 2018) est une des difficultés de mise en œuvre des DDM. D'autres études utilisent les modèles de Dempster-Shafer (Verbert, De Schutter and Babuška, 2015; Verbert, Babuška and De Schutter, 2017) (théorie des croyances), la logique floue (Vafaei, Ribeiro and Camarinha-Matos, 2019) ou encore la théorie des possibilités pour gérer l'incertitude des données. Par ailleurs, les modèles d'apprentissage automatique peuvent parfois être perçus comme des "boîtes noires", ce qui complique leur interprétation et leur validation dans des environnements critiques.

1.3.3.4 Approche multimodèle (MM)

Les approches monomodèles ne permettent généralement pas de traiter toutes les tâches de diagnostic des systèmes et processus industriels. Les études consultées qui utilisent des approches monomodèles (KBM, DDM et PBM) proposent souvent des modèles complémentaires pour surmonter les points faibles de certains modèles. Il est de plus en plus courant de trouver des études de recherche qui combinent différents modèles dans des approches multimodèles pour surmonter la complexité des tâches de maintenance prévisionnelle. Cette complexité croissante inclut le nombre de pannes potentielles et de modes de défaillance du système ou du processus surveillé, le type et le nombre de sources d'information et/ou de données obtenues, ainsi que le nombre de tâches de diagnostic, en plus de la complexité de conception du modèle sélectionné. La plupart des études consultées présentent des cas d'usages limités à seulement quelques modes de défaillance, parfois un seul, ce qui rend difficile l'extrapolation des approches monomodèles aux applications de systèmes complexes de l'industrie 4.0 (Alaswad and Xiang, 2017).

En réponse, les MM combinent plusieurs techniques pour surmonter les limites des approches monomodèles.

Les approches multimodèles peuvent se structurer de trois façons principales :

- **Modèles en série** : La sortie d'un modèle sert d'entrée à un autre. Par exemple, une carte auto-organisatrice (SOM) combinée à un modèle statistique permet de détecter les pannes en série (Schwartz *et al.*, 2020).
- **Modèles en parallèle** : Les modèles traitent des données simultanément et leurs sorties sont combinées. Un modèle basé sur les données et un modèle basé sur la physique peuvent fonctionner ensemble pour estimer la durée de vie restante des composants (Al-Dulaimi *et al.*, 2019).
- **Modèles imbriqués** : Un modèle est inclus dans un autre. Par exemple, un modèle neuro-flou avec un modèle de Markov caché (Soualhi *et al.*, 2014).

Et en ce qui concerne les combinaisons possibles pour les approches MM, elles sont illustrées en Figure 4. Un MM est caractérisé par l'association de deux modèles, qu'il soit de même type ou différent, et il n'y a pas de limites pour le nombre de modèles composant une approche MM.

En conclusion, les approches MM, bien qu'elles apportent des avantages en matière de précision et de gestion de la complexité, posent des défis supplémentaires en termes de conception de complexité et

d'intégration des différents modèles. Cependant, elles offrent des perspectives prometteuses pour améliorer la maintenance prévisionnelle, en intégrant données, connaissances expertes et modèles physiques (Javed, Gouriveau and Zerhouni, 2017).

1.3.3.5 Bilan des méthodes de diagnostic

Sur la base des analyses précédentes, le Tableau 2 résume les avantages et inconvénients de chaque méthode de diagnostic vis-à-vis de l'adaptabilité à l'hétérogénéité rencontré dans les processus de production hétérogènes.

Tableau 2. Résumé des méthodes de diagnostics existantes

Nom de la méthode de diagnostic	Description	Composants clés	Sous-méthodes	Avantages vis-à-vis de l'adaptabilité à l'hétérogénéité	Désavantages vis-à-vis de l'adaptabilité à l'hétérogénéité
Modèles basés sur la physique (PBM)	Utilisation des lois fondamentales de la physique pour les simulations de diagnostic	Fondement mathématique et outils de calcul	Méthodes des éléments finis, modèles de propagation de fissures, et modèles de fatigue	Fourniture d'une haute précision et d'une grande exactitude dans le domaine physique spécifique pour lequel ils sont formulés ; capacité à assurer des performances fiables dans des conditions bien définies	Faible généralité ; limitation causée par les hypothèses et les simplifications inhérentes aux lois physiques, réduisant ainsi l'adaptabilité aux environnements hétérogènes
Modèles basés sur la connaissance (KBM)	Utilisation de règles, de cas et de logique pour le raisonnement diagnostique	Règles, cas et logique	Modèles basés sur les règles, modèles basés sur les cas, et modèles basés sur la logique floue	Adaptabilité pilotée par l'expertise ; capacité à intégrer systématiquement les connaissances d'experts, permettant des ajustements et des raffinements pour améliorer la réponse à des conditions hétérogènes spécifiques	Dépendance à l'humain ; dépendance aux règles et aux cas prédéfinis, ce qui limite l'adaptabilité à des scénarios imprévus et à des conditions nouvelles
Modèles basés sur les données (DDM)	Utilisation de vastes données pour les analyses de diagnostic	Techniques statistiques et algorithmes d'apprentissage automatique/profond	Modèles statistiques, modèles stochastiques, et modèles d'apprentissage automatique	Flexibilité significative dans la modélisation de jeux de données divers et complexes ; capacité à capturer des motifs complexes et à généraliser à des conditions variées	Forte dépendance à la qualité des données ; susceptibilité à une performance réduite lorsqu'il est confronté à des données bruitées, éparses ou biaisées, ce qui impacte l'adaptabilité à l'hétérogénéité

En résumé, dans le contexte du diagnostic de dérive des processus de production hétérogènes :

- Les PBM reposent sur une modélisation détaillée des phénomènes physiques selon les lois spécifiques des systèmes étudiés. Cette approche offre une grande précision, mais reste limitée en termes d'adaptabilité à la variabilité et à l'hétérogénéité des systèmes. En effet, ces modèles nécessitent une connaissance approfondie des principes physiques et des compétences avancées en mathématiques pour assurer une représentation précise. Leur application peut être coûteuse en termes de développement et les résultats instables, particulièrement dans des environnements où les phénomènes ne sont pas entièrement compris ou documentés.
- Les KBM s'appuient sur l'intégration des connaissances expertes à travers des règles prédéfinies. Ils offrent une adaptabilité modérée face à la variabilité et à l'hétérogénéité des systèmes. Ils sont principalement limités par l'étendue et la qualité des connaissances disponibles. La formulation des règles et la sélection des cas dépendent fortement de l'expertise disponible, ce qui peut restreindre leur efficacité dans des situations non anticipées ou dans des systèmes évolutifs.
- Les DDM exploitent de vastes ensembles de données à l'aide de techniques d'intelligence artificielle et d'algorithmes d'apprentissage automatique et d'apprentissage profond. Ces modèles se distinguent par leur grande capacité d'adaptation à la variabilité et à l'hétérogénéité. Cette particularité les rend particulièrement efficaces pour traiter des systèmes complexes où les relations physiques sont difficiles à modéliser. Bien que ces modèles ne dépendent pas autant de l'expertise spécifique au système, ils requièrent néanmoins des compétences avancées en science des données pour le prétraitement, la sélection des algorithmes appropriés, et l'interprétation des résultats. La qualité des diagnostics produits est étroitement liée à la qualité et à la quantité des données disponibles.

Ainsi, sur la base des potentialités soulignées, le développement d'un diagnostic de dérive efficace nécessite souvent une connaissance experte de l'actif ou du processus étudié. C'est pourquoi les MM qui combinent les avantages de plusieurs méthodes devraient offrir une solution plus robuste.

En effet, les DDM permettent d'identifier les dérives des processus de production hétérogènes, mais ne peuvent pas identifier les causes racines physiques entraînant la dérive. Ces derniers pouvant être liés à un nombre variable d'éléments physiques, il est nécessaire d'intégrer de la connaissance via des KBM afin de pouvoir lier les dérives identifiées à des causes racines physiques.

En ce sens, pour développer un diagnostic évolutif et pertinent des dérives dans des processus de production hétérogènes, capable de s'adapter aux évolutions futures des processus observés, il est crucial de répondre à **une deuxième question scientifique concernant le développement d'une méthode de diagnostic d'un processus hétérogène basée sur une approche MM efficace.**

Deuxième problématique scientifique

- ***Comment concevoir un système de diagnostic des dérives pour les processus de production hétérogènes, combinant des modèles basés sur les données (DDM) et des modèles basés sur la connaissance (KBM) ?***

1.3.4 Intégration des étapes de détection de dérive et de diagnostic

Comme démontré dans les sections précédentes, la détection de dérive et le diagnostic sont deux activités du PHM étroitement liées. En effet, la détection identifie les écarts significatifs par rapport aux performances de référence, tandis que le diagnostic vise à en déterminer les causes racines.

L'un des principaux défis du diagnostic est d'avoir accès à des données représentatives des différentes dérives potentielles ou à une expertise approfondie permettant de construire des modèles de diagnostic fiables. Sur un équipement neuf, les données de dérive peuvent être limitées. En effet, les premiers signes de dégradation n'apparaissent souvent qu'après plusieurs mois ou années en raison de facteurs comme l'usure mécanique. Sans ces données, établir des diagnostics précis devient difficile. Ce constat va dans le sens des travaux de (Nunes, Santos and Rocha, 2023) qui soulignent que l'absence d'intégration entre ces deux étapes (les deux étapes sont « penser » en complémentarité en fonction des données disponibles) limite fortement la capacité des systèmes à poser des diagnostics de dérive fiables et précis.

Pour surmonter ce défi d'intégration, il est nécessaire de développer une approche méthodologique qui propose une considération conjointe de la détection et du diagnostic, et ceci dans le cadre d'une boucle d'apprentissage continu permettant de s'adapter à la dynamique, à la vie du système et donc des situations rencontrées et non nécessairement connues à l'avance.

Définition : Une **approche méthodologique** est un ensemble structuré de principes, de méthodes et de procédures destinés à guider la réalisation d'un processus ou d'un projet. Elle fournit une démarche systématique permettant d'organiser les étapes, d'assurer la cohérence des actions et de faciliter l'analyse et l'amélioration continue des résultats.

Dans le contexte de la détection de dérive et du diagnostic, l'approche méthodologique doit structurer la mise en place de moyens/activités/modèles permettant de détecter et diagnostiquer les dérives dans les processus de production hétérogènes. Elle doit également intégrer des étapes d'apprentissage continu pour mettre à jour les modèles proposés pour la détection/diagnostic en utilisant les dernières données disponibles représentatives de la dynamique/évolution des processus de production.

Cependant, pour maximiser son impact vis-à-vis du contexte d'application, cette approche méthodologique se doit d'être **générique** et adaptable à des environnements industriels hétérogènes sans nécessiter des adaptations complexes pour chaque contexte spécifique. En ce sens (Lee, Ghaffari and Elmeligy, 2011; Kusiak, 2019; Morgan *et al.*, 2021) insiste sur l'importance d'une approche méthodologique flexible qui intègre l'analyse des données en temps réel et s'ajustent automatiquement aux variations des processus.

De plus (Bagheri, Yang, *et al.*, 2015), montrent que des méthodologies 'holistiques' au sein de l'industrie 4.0 doivent gérer non seulement la surveillance, mais aussi la détection de dérive et le diagnostic de manière intégrée, tout en étant capables de s'améliorer continuellement grâce à l'apprentissage basé sur les données collectées.

Par conséquent, la troisième question scientifique pour cette thèse qui en découle porte sur le développement d'une approche méthodologique pour la détection des dérives et le diagnostic des processus de production hétérogènes de manière intégrée et évolutive.

Troisième problématique scientifique

- ***Comment élaborer une approche méthodologique unifié pour la détection et le diagnostic des dérives dans les processus de production hétérogènes ?***

1.4 Conclusion

Ce chapitre a eu pour objectif de présenter de façon complémentaire, d'une part, le contexte industriel genèse de cette thèse et référée à l'entreprise SEW USOCOME, et d'autre part, le contexte scientifique référé au PHM.

En ce sens, dans la première partie a été mis en évidence la complexité des produits fabriqués (des motoréducteurs électriques ... très personnalisés) et les exigences qui y sont associées, et par voie de conséquence la complexité des processus de production en support de cette fabrication. Ceci a permis de caractériser ces processus comme hétérogènes, et de souligner toute l'importance de la maîtrise, en opération de ces processus, surtout en regard de leurs dérives pouvant conduire à des situations de production non nominales. Cette maîtrise fait l'objet actuellement d'actions du Pôle Processus et Innovation, commanditaire de cette thèse, et plus précisément de la mise en œuvre d'une maintenance prévisionnelle en référence à l'ISO 13374 avec une focalisation pour la thèse, sur les activités de détection et de diagnostic de dérives. Ces activités sont difficiles à déployer en regard des caractéristiques des processus à assurer et ont conduit à la formalisation d'un ensemble de verrous industriels et de contraintes matérielles/logicielles à respecter pour l'implantation des solutions associées.

Sur cette base orientée industrielle, la deuxième partie du chapitre a cherché à repositionner ces problématiques de détection/diagnostic en lien avec la maintenance prévisionnelle dans un contexte plus scientifique, celui du PHM. Ceci a permis de développer des états de l'art permettant de mettre en évidence les travaux les plus pertinents du PHM relatif à la détection et diagnostic de dérive de processus de production hétérogènes. Plus précisément ont été abordés les problèmes de caractérisation des données de ce type de processus (p. ex. séries temporelles multivariées), l'exploitation des techniques d'apprentissage automatique

pour la détection de dérives à partir de ce type de données, les différents types de modèles pour le diagnostic de dérive en recensant les PBM, les KBM, les DDM et leurs combinaisons et, enfin, les problèmes d'intégration de ces deux activités en un tout cohérent.

À partir de ces travaux identifiés ont été soulevés plusieurs manques dans la littérature pour répondre totalement aux verrous industriels isolés en première partie. Ceci a conduit dans le chapitre à formaliser 3 problématiques scientifiques :

1. **Comment développer une méthode de détection des dérives de processus de production hétérogènes ne nécessitant pas de données étiquetées et suffisamment évolutive pour s'adapter à toute dérive apparaissant tout au long de la vie du processus ?**
2. **Comment concevoir un système de diagnostic des dérives pour les processus de production hétérogènes, combinant des modèles basés sur les données (DDM) et des modèles basés sur la connaissance (KBM) ?**
3. **Comment élaborer une approche méthodologique unifiée pour la détection et le diagnostic des dérives dans les processus de production hétérogènes ?**

Ces problématiques sont reprises dans le chapitre 2 pour proposer 3 contributions scientifiques majeures dans le cadre de cette thèse.

Chapitre 2

Vers une approche méthodologique générique pour la détection de dérive et le diagnostic des processus de production hétérogènes

2.1 Introduction

Le chapitre précédent a présenté le contexte industriel et scientifique de cette thèse. Ceci a permis d'une part de mettre en évidence les verrous industriels liés à la surveillance et au diagnostic des processus de production chez SEW USOCOME, puis d'autre part, de replacer ces verrous dans un contexte scientifique plus large, celui de la maintenance prévisionnelle et du PHM. Cette recontextualisation a mené à l'identification de trois problématiques scientifiques.

Ce deuxième chapitre présente les contributions apportées à chacune des trois problématiques scientifiques en détaillant chaque contribution dans une section spécifique.

En ce sens, la Section 0 traite de la problématique scientifique concernant **l'adaptation des méthodes de détection des dérives aux spécificités des processus de production hétérogènes**. Dans cette section une solution de détection de dérive capable de s'adapter à la variabilité et à la complexité des données industrielles est proposée. La solution de détection de dérive se base sur le principe de la détection de nouveautés pour détecter toute dérive par rapport à une référence nominale sans connaissance préalable des modes de dérives des processus.

La Section 2.3 se concentre sur la problématique scientifique concernant **le diagnostic des dérives de processus de production hétérogènes**. Dans cette section une solution de diagnostic permettant de caractériser les dérives des processus et de les associer à des causes racines physiques est proposée. La solution de diagnostic se base sur une architecture couplant l'expertise humaine à des techniques d'IA capable d'identifier avec précision des profils de données connus.

La Section 2.4 aborde la problématique scientifique **d'intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes**. Dans cette section, une approche méthodologique intégrant de manière complémentaire les activités de détection et de diagnostic des dérives est proposée. L'approche méthodologique reprend les spécificités des solutions présentées en Sections 2.2 et 2.3 pour les combiner. De plus, l'approche méthodologique intègre une boucle d'apprentissage continue s'enrichissant de nouvelles données et connaissances pour améliorer la couverture et les performances de la détection et du diagnostic de dérive.

Le chapitre se conclut par la Section 2.5 qui propose un résumé des contributions majeures de la thèse en soulevant l'application de cette approche méthodologique à un cas applicatif industriel concret de la société SEW USOCOME (objectif du chapitre 3).

2.2 Détection de dérive des processus de production hétérogènes.

Comme introduit dans le chapitre 1, les systèmes de production numériques prônés par l'industrie 4.0 sont souvent caractérisés par leur hétérogénéité, leur complexité et leur nature non stationnaire. Ces caractéristiques posent des défis scientifiques majeurs pour la détection de dérives. À titre d'exemple, contrairement à des systèmes bien modélisés, où il est possible d'établir des lois physiques stables, les processus

de production hétérogènes comme ceux de SEW USOCOME se caractérisent par des données multivariées non linéaires, évoluant au fil du temps sous l'influence de multiples paramètres (Kamm, Jazdi and Weyrich, 2021; CHAPELIN *et al.*, 2023). La difficulté majeure dans ces systèmes réside dans l'impossibilité de prévoir à l'avance toutes les dérives possibles, mais aussi dans un nombre d'occurrences faible des dérives observées, ce qui réduit la pertinence d'une approche de modélisation statique basée sur des données labellisées (Fassois and Kopsaftopoulos, 2013b). Par ailleurs, la non-stationnarité des environnements industriels signifie que les méthodes qui reposent sur des hypothèses fixes concernant les distributions des données deviennent rapidement obsolètes. Par ailleurs, la non-stationnarité des environnements industriels concernés implique que les méthodes qui reposent sur des hypothèses fixes relativement aux distributions des données deviennent rapidement obsolètes. Il faut donc s'orienter vers des méthodes plus flexibles, donc plus facilement maintenables et adaptables à un large spectre d'application par un passage à l'échelle maîtrisé.

Ainsi, la première contribution de cette thèse est une contribution méthodologique et non fondamentale qui repose sur la sélection et l'adaptation de méthodes d'apprentissage automatique existantes, plus particulièrement la **détection de nouveauté**, pour répondre à la problématique de la **détection de dérives dans des processus de production hétérogènes**. Dans l'objectif de répondre aux défis liés à l'hétérogénéité, à la non-stationnarité des environnements industriels et à la problématique de passage à l'échelle, la contribution prend la forme d'un logigramme de décision. Ce logigramme guide l'utilisateur à travers plusieurs étapes clés, en fonction d'indicateurs représentatifs des caractéristiques des données brutes (types séries temporelles représentatives des processus de production hétérogènes). En fonction des indicateurs, le logigramme détermine si les données doivent être transformées ou conservées brutes, puis propose l'algorithme de détection de nouveautés le plus adapté. Grâce à cette approche, la méthodologie offre une solution générique à tout type de séries temporelles représentatives de processus de production complexes. En automatisant la sélection des algorithmes et en guidant sur la nécessité de transformer ou non les données, la méthodologie réduit les ajustements manuels. Cela facilite ainsi le passage à l'échelle, quelle que soit la taille des ensembles de données ou la diversité des environnements industriels.

2.2.1 Choix de la technique d'apprentissage

La technique d'apprentissage automatique, bien que contrainte imposée par l'industriel, est également le choix le plus pertinent pour notre recherche, comme le confirme l'analyse critique des méthodes de détection de dérive en Section 1.3.2. Comme introduit dans cette même Section, parmi les techniques d'apprentissage automatique adapté à la détection de dérive, plusieurs solutions peuvent être envisagées : la classification supervisée, le regroupement et la détection de nouveauté.

Dans le chapitre 1, ces solutions ont été surtout caractérisées pour en déterminer les éléments discriminants. Sur cette base de caractérisation, cette sous-section propose maintenant une analyse critique de ces différentes solutions pour déterminer la plus appropriée aux processus de production hétérogènes, dans un contexte industriel caractérisé par la rareté des données de dérives.

2.2.1.1 Classification supervisée

La classification supervisée est souvent utilisée dans l'industrie pour la détection d'anomalies, afin d'identifier des observations qui diffèrent du comportement attendu. (Chalapathy and Chawla, 2019) expliquent que ces méthodes reposent sur des modèles qui nécessitent des données labellisées pour distinguer les comportements normaux des anomalies. Ces modèles sont efficaces dans des environnements où les anomalies sont fréquentes et bien documentées. Cependant, dans des environnements industriels complexes et non stationnaires, les dérives et anomalies sont souvent rares et imprévisibles. Cette dépendance aux données labellisées limite l'efficacité des modèles supervisés (Campos *et al.*, 2016; Munir *et al.*, 2019) soulignent que ces modèles sont inefficaces lorsque les environnements évoluent rapidement. Leur rigidité pose un problème majeur dans des systèmes où les profils de dérive sont difficiles à anticiper. Par exemple, dans le secteur de l'énergie éolienne, des modèles supervisés surveillent en temps réel la performance des turbines. Les capteurs collectent les données des machines, et les modèles détectent les anomalies lorsqu'elles surviennent. Cette approche est performante lorsque les dérives sont connues à l'avance, mais dans des environnements où elles sont rares et imprévisibles, l'efficacité de ces méthodes diminue (Chen *et al.*, 2019; Panthi, 2020; Zhang and Yang, 2023).

2.2.1.2 Regroupement

Le regroupement est une méthode non supervisée qui regroupe les données en fonction de leur similarité. (Jain, Murty and Flynn, 1999) montrent que des techniques comme K-means ou GMM sont

couramment utilisées pour cette tâche. Ces méthodes sont particulièrement efficaces lorsque les comportements normaux et anormaux ne sont pas définis à l'avance.

Le principal inconvénient du regroupement est qu'il nécessite des hypothèses initiales sur la structure des données. (Fahad *et al.*, 2014) expliquent que déterminer le nombre de clusters ou leur distribution peut être difficile dans des environnements dynamiques. Dans des systèmes industriels où les variables changent constamment, le regroupement perd souvent en efficacité. (Bohte, Poutré and Kok, 2002) notent que ces modèles dépendent fortement des paramètres initiaux et de la qualité des données. Cela les rend vulnérables aux erreurs dans des environnements complexes.

Dans la surveillance des machines industrielles, le regroupement définit les profils de performance selon les différents régimes de fonctionnement. Cependant, les ajustements réguliers des modèles sont nécessaires pour s'adapter aux évolutions des processus, ce qui limite leur efficacité à grande échelle (Aghabozorgi, Seyed Shirkorshidi and Ying Wah, 2015).

Une variante des méthodes de regroupement statique appelée regroupement de flux de données (data stream clustering) permet d'adapter le regroupement aux flux de données continus et évolutifs. Cette méthode regroupe les données en clusters qui changent au fil du temps et permet d'identifier des changements structurels dans les groupes de données. Ces changements peuvent être interprétés comme des dérives dans les processus industriels. (Amini, Wah and Saboohi, 2014) démontrent l'efficacité de cette méthode pour traiter des données en temps réel. Cependant cette même étude met en lumière la dépendance directe des résultats à la qualité paramétrage de l'algorithme. De plus, cette solution peut être coûteuse en termes de calculs et de ressources d'exploitation.

2.2.1.3 Détection de nouveauté

Tout comme la classification supervisée, **la détection de nouveauté** repose sur l'entraînement d'un classificateur, mais avec une différence clé : au lieu d'utiliser plusieurs classes de données labellisées, elle se concentre uniquement sur une classe, généralement le comportement normal. Cela permet au modèle de détecter tout écart par rapport à cette classe de référence. Contrairement à la détection d'anomalies, qui dépend de la distinction entre des comportements normaux et anormaux connus a priori, la détection de nouveauté se concentre sur la modélisation du comportement normal, sans nécessiter d'exemples de comportements anormaux. Cela en fait une approche particulièrement adaptée pour les environnements industriels complexes, où les dérives sont rares et souvent mal documentées.

(Pimentel *et al.*, 2014) ont montré que la détection de nouveauté est efficace dans des environnements non stationnaires, où la collecte de données étiquetées est difficile ou coûteuse. Cette approche est particulièrement pertinente dans les processus industriels hétérogènes, car elle permet d'identifier des comportements déviants sans avoir besoin de données sur les dérives antérieures.

Des algorithmes comme le **One-Class SVM** et l'**Isolation Forest**, qui se concentrent sur la séparation des comportements normaux de tout écart, sont fréquemment utilisés pour la détection de nouveauté dans les environnements industriels (Ruff *et al.*, 2018). Ces méthodes sont robustes face aux changements dynamiques dans les processus, permettant une surveillance continue sans besoin de réajuster fréquemment le modèle.

D'autres travaux, comme ceux de (Campos *et al.*, 2016) ont montré que les méthodes de détection de nouveauté, appliquées à des flux de données, sont capables de capturer des schémas nouveaux dans des environnements où les données évoluent constamment.

Dans la pratique, la détection de nouveauté est utilisée pour surveiller des systèmes tels que les réseaux énergétiques, où les dérives sont rares, mais critiques. (Khan and Madden, 2009) ont démontré que ces systèmes permettent une détection précise des écarts, en particulier dans des environnements industriels dynamiques. De plus, des approches plus récentes, dont la **détection de nouveauté basée sur l'apprentissage profond**, a montré leurs efficacités dans des environnements non stationnaires, en particulier lorsqu'il est difficile de collecter des données étiquetées (Hossen *et al.*, 2024).

Toutefois, certains algorithmes de détection de nouveauté, comme les autoencodeurs, nécessitent des volumes importants de données pour l'entraînement, ce qui peut poser des défis de mise en œuvre lorsque les données sont limitées. D'autres approches, telles que les **One-Class SVM**, sont plus efficaces dans des environnements où les ressources sont contraintes, tout en offrant une détection rapide des nouveautés.

2.2.1.4 Synthèse de l'analyse critique des méthodes d'apprentissage automatique de détection de dérives

Dans le cadre de la détection de dérives des processus de production hétérogènes, et compte tenu des différentes caractéristiques des processus, des dérives et des données à considérer (voir sous-section

précédente), la **détection de nouveautés** apparaît comme la solution la plus adaptée pour répondre aux besoins spécifiques du contexte industriel. Ce choix est justifié par les éléments de synthèse majeurs suivants :

- Contrairement aux méthodes de **classification supervisée**, qui reposent souvent sur des données étiquetées comme anormales, la détection de nouveauté permet d'apprendre uniquement à partir de données normales, ce qui la rend particulièrement efficace pour traiter des dérives peu documentées (Pimentel *et al.*, 2014)
- Les méthodes de regroupement, bien qu'utiles pour regrouper des comportements similaires, sont souvent limitées par la nécessité de définir des paramètres initiaux, comme le nombre de clusters, et par leur sensibilité aux données bruitées (Zimek, Campello and Sander, 2014). En revanche, la **détection de nouveauté** est plus robuste face à ces limitations grâce à sa capacité à modéliser uniquement les comportements normaux et à détecter toute nouveauté, sans avoir besoin de spécifier des clusters a priori.
- Dans les environnements non stationnaires, où les processus évoluent rapidement, la détection de nouveauté se distingue par sa capacité à s'adapter aux changements rapides tout en nécessitant peu de réajustements avec des algorithmes comme l'**Isolation Forest** ou les **One-Class SVM**.
- Les algorithmes modernes de détection de nouveauté, comme les **Autoencodeurs Variants (VAE)**, permettent d'intégrer des données complexes en modélisant leurs distributions internes avec une grande précision. (Chen *et al.*, 2019; Zhang and Yang, 2023) ont démontré que ces modèles offrent une grande flexibilité pour s'adapter aux environnements dynamiques, tout en réduisant la nécessité de réentraînement fréquent.

Ainsi, la **détection de nouveauté** est la technique de base retenue en support de notre première contribution scientifique.

2.2.2 Proposition d'algorithmes adaptée à la détection de nouveautés

Les algorithmes de détection de nouveautés reposent sur un large éventail d'algorithmes d'apprentissage automatique et profond, chacun offrant des avantages et des limitations spécifiques. La sélection d'un algorithme approprié dépend de divers facteurs directement définis par l'application. Parmi ces facteurs on retrouve la nature des données, la complexité des relations entre les variables, et les besoins en termes de précision et de performance computationnelle. Il n'existe donc pas d'algorithme universel, il faut composer avec les caractéristiques du processus surveillé et choisir l'algorithme approprié en conséquence. L'état de l'art des algorithmes de classification adapté aux séries temporelles multivariées a permis d'identifier un espace d'algorithmes synthétisé par le Tableau 1.

À partir de cette synthèse, les algorithmes de classification adaptés à la détection de nouveautés ont pu être extraits et résumés dans le Tableau 3, parmi eux on retrouve : **L'Isolation Forest, les AE, les VAE, les One-Class SVM, les LSTM, les GMM et la PCA.**

Tableau 3. Récapitulatif des algorithmes de classification adapté à la détection de nouveautés et de leurs spécificités vis-à-vis des données

ALGORITHME	AVANTAGES SPÉCIFIQUES EN DÉTECTION DE NOUVEAUTÉS	LIMITES POTENTIELLES
Isolation Forest	Efficace pour isoler les anomalies dans de grands ensembles de données, avec une bonne performance computationnelle.	Moins performant avec des distributions de données non homogènes ou complexes.
Autoencodeurs (AE) / Autoencodeurs Variationnels (VAE)	Adaptés aux relations non linéaires complexes et capables de détecter des écarts subtils dans le comportement des processus.	Exigent un grand volume de données pour un entraînement optimal et sont coûteux en calcul.
One-Class SVM	Utile pour les distributions non linéaires, sensibles aux paramètres.	Sensible aux choix de paramètres et inefficace pour les grands ensembles de données.

LSTM	Conçu pour capturer les dépendances temporelles dans les séries temporelles.	Nécessite des réglages complexes et consomme beaucoup de ressources pour l'entraînement.
Modèles de Mélange Gaussien (GMM)	Offre une approche probabiliste pour la détection d'anomalies.	Limité avec des données non linéaires et sensibles aux dimensions élevées.
Analyse en Composantes Principales (PCA)	Fournis une réduction de dimensionnalité efficace.	Inefficace avec des données non linéaires et sensibles aux valeurs extrêmes.

L'objectif de ce nouveau tableau de synthèse est de guider la sélection de l'algorithme le plus approprié en fonction des caractéristiques spécifiques des données du processus de production surveillé. Chaque algorithme présente des spécificités en termes de fonctionnement, de capacité à gérer des données multivariées, et de performance computationnelle, ce qui influe directement sur leur applicabilité dans des environnements industriels. Ainsi, cette sélection d'algorithmes fournit une boîte à outils essentielle pour la mise en œuvre pratique de la détection de nouveauté. **Cette liste d'algorithmes n'est pas exhaustive en raison de l'évolution rapide du domaine de l'apprentissage automatique et de l'apparition constante de nouveaux algorithmes, mais elle fournit un point de départ pertinent en sachant que certains besoins spécifiques peuvent nécessiter des solutions supplémentaires ou alternatives à celles proposées.** À plus grande échelle, cette proposition d'algorithmes vise à soutenir une approche flexible et générique de la détection de dérives, en permettant une mise en œuvre adaptée aux contraintes des environnements industriels et aux objectifs de surveillance et de maintenance prévisionnelle.

2.2.3 Méthodologie guidée pour l'optimisation du traitement des données et la sélection des algorithmes de détection de nouveautés.

Pour faciliter le choix des algorithmes de détection de dérives identifiés dans le tableau précédent, en fonction des données caractérisant ces dérives (séries temporelles multivariées), une méthodologie de sélection a été développée. Celle-ci repose par conséquent sur une analyse rigoureuse des caractéristiques des données telles que résumées dans le Tableau 3 en prenant en compte majoritairement : l'analyse des volumes de données, l'analyse des corrélations entre variables et l'analyse des dépendances temporelles. À partir de ces analyses, des indicateurs sont extraits, notamment le nombre de points dans les séries temporelles, les coefficients de corrélation de Pearson et Spearman, ainsi que les valeurs d'autocorrélation (ACF). Les seuils et valeurs utilisés (ACF = 0,5, corrélations de Pearson et Spearman à 0,7) découlent de ce même état de l'art. Ces indicateurs jouent un rôle central dans la décision de conserver les données brutes ou d'appliquer des réductions de dimension ou des transformations statistiques.

Par conséquent, la méthodologie (de type logigramme de décision) représentée en Figure 5 suit une démarche en plusieurs étapes afin de guider un utilisateur, sur la base du calcul des indicateurs mentionnés, pour le choix de l'algorithme de détection de nouveautés le plus adapté à la série temporelle des données qui caractérise son processus de production. Ces étapes sont :

1. **Évaluation des dépendances temporelles (calcul de l'ACF)** : Le processus commence par le calcul de la **fonction d'autocorrélation (ACF)**, qui mesure la corrélation entre les valeurs d'une série temporelle et leurs versions décalées dans le temps (lags). L'ACF est calculée pour plusieurs lags, mais, dans le contexte de **processus finis et répétitifs**, il est particulièrement pertinent de se concentrer sur les **petits lags**. En effet, ces processus se répètent souvent sur des intervalles réguliers, ce qui signifie que les dépendances temporelles à court terme sont les plus significatives pour capturer ces répétitions cycliques. Si **l'ACF est supérieure à 0,5** pour des petits lags, cela indique une forte dépendance temporelle à court terme, typique des processus répétitifs. Ce seuil de 0,5, identifié à partir de l'état de l'art, permet de distinguer les séries avec une structure temporelle significative. Dans ce cas, il est recommandé de **garder les données brutes** afin de préserver cette information temporelle, car les algorithmes spécialisés dans la gestion des dépendances temporelles, tels que les **Réseaux de Neurones Récurrents (RNN)** ou les **réseaux LSTM**, sont particulièrement adaptés à ces situations. En se concentrant sur les petits lags, la méthodologie permet de capturer efficacement les cycles répétitifs typiques des processus industriels hétérogènes, tandis que les grands lags sont généralement moins pertinents dans ce contexte où les processus se répètent sur des périodes plus courtes.
2. **Vérification de la taille des séries temporelles et analyse des corrélations linéaires (calcul du coefficient de Pearson)** : Si l'ACF est inférieure à 0,5, cela signifie que la dépendance temporelle est

faible. Dans ce cas, l'analyse continue avec le **coefficient de corrélation de Pearson**, qui mesure la relation linéaire entre les variables. Si ce coefficient est supérieur à 0,7, comme suggéré par l'état de l'art, cela montre une forte corrélation linéaire entre les variables. Une **réduction de dimension linéaire** par **PCA** est alors appliquée pour simplifier les données sans perdre l'essence de leur structure. Les algorithmes linéaires, tels que **KNN** ou **SVM**, sont ensuite recommandés, car ils sont mieux adaptés à ces relations linéaires.

3. **Analyse des corrélations non linéaires (calcul du coefficient de Spearman)** : Si le coefficient de Pearson est inférieur à 0,7, cela suggère que les relations linéaires sont faibles. Dans ce cas, on calcule le **coefficient de corrélation de Spearman**, qui détecte les relations monotones (linéaires ou non). Si le coefficient de Spearman dépasse 0,7, un seuil défini dans l'état de l'art pour les relations monotones, cela indique la présence de relations non linéaires fortes entre les variables. Une **réduction de dimension non linéaire** via des techniques telles que **t-SNE** ou des **autoencodeurs** est alors effectuée. Ces transformations permettent de simplifier la structure complexe des données, tout en préservant leurs caractéristiques essentielles. Des algorithmes non linéaires, tels que **Random Forest (RF)** ou les **réseaux de neurones profonds (DNN)**, sont alors suggérés pour capturer ces relations non linéaires.
4. **Transformation des données (extraction statistique)** : Si aucun des coefficients de corrélation (Pearson ou Spearman) ne dépasse les seuils définis (Pearson < 0,7 et Spearman < 0,7), cela signifie que les relations entre les variables sont trop faibles pour que les approches de réduction de dimension classique soient efficaces. Dans ce cas, il est nécessaire de **transformer les données** en extrayant des caractéristiques statistiques (par exemple via **FFT**, analyse fréquentielle ou autres). Cette étape permet de représenter les données sous une forme plus exploitable. Une fois cette transformation effectuée, des algorithmes comme **SVM**, **LSTM**, ou **RNN** sont proposés en fonction des nouvelles caractéristiques des données transformées.

Ainsi, la méthodologie guide l'utilisateur à travers plusieurs étapes pour l'aider à décider s'il est nécessaire de transformer les données. À chaque étape, des solutions de transformation des données sont proposées en fonction des indicateurs extraits des caractéristiques des données. Ensuite, en fonction des décisions prises sur la transformation des données, le logigramme oriente l'utilisateur vers l'algorithme de détection de nouveautés le plus pertinent, adapté au nouveau format des données. Cette méthodologie est générique à tout type de séries temporelles, car elle repose sur des indicateurs pertinents et universels. De plus, le logigramme offre une approche cohérente pour le passage à l'échelle des algorithmes de détection de nouveautés dans des contextes industriels, en permettant de choisir la solution la plus adaptée aux données étudiées.

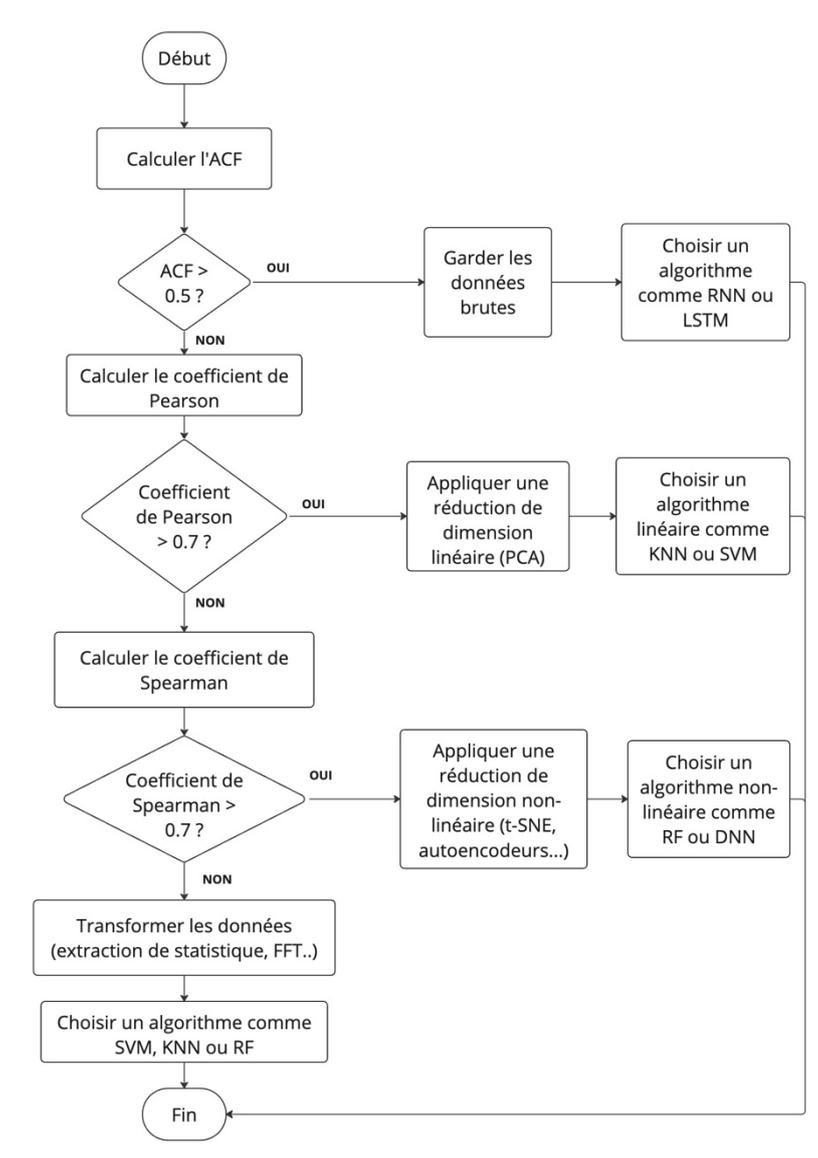


Figure 5. Logigramme de décision pour le traitement des données et le choix d'algorithme de détection de nouveautés associé.

2.2.4 Synthèse de la contribution scientifique à la problématique d'adaptation des méthodes de détection des dérives aux spécificités des processus de production hétérogènes

En synthèse de la contribution scientifique proposée dans la section suivante, il est pertinent de dégager les axes de valeur ajoutée suivants :

1. Une analyse critique des méthodes de détection de dérives, avec une justification de l'apprentissage automatique comme solution optimale.
2. La démonstration de la pertinence de la détection de nouveautés pour les processus de production hétérogènes.
3. L'identification d'algorithmes de classification adaptés à la mise en œuvre de la détection de nouveautés, appliqués à la détection de dérives des processus de production hétérogènes.
4. La proposition d'une méthodologie structurée pour guider le choix des traitements de données et des algorithmes de détection de nouveautés, afin de maximiser l'efficacité des processus de surveillance.

Finalement, en facilitant le choix de l'algorithme le plus pertinent en fonction des indicateurs propres aux données (comme la corrélation, la dimensionnalité ou la dépendance temporelle), la méthodologie permet

une réponse adaptée à des environnements industriels complexes et hétérogènes, où les processus ne suivent pas des schémas prévisibles ou uniformes. De plus, elle simplifie le passage à l'échelle en permettant de sélectionner facilement et rapidement l'algorithme approprié aux données. Ainsi, la contribution résultant de ces 4 axes répond directement à la problématique scientifique posée en début de section, à savoir **l'adaptation des méthodes de détection des dérives aux spécificités des processus de production hétérogènes**.

2.3 Diagnostic de dérive des processus de production hétérogènes.

Dans la continuité des enjeux abordés dans la section précédente autour de la détection de dérives, le diagnostic de ces dérives des processus de production hétérogènes constitue une seconde problématique majeure afin de pouvoir identifier avec précision les dérives elles-mêmes, et également leurs causes racines.

En ce sens, dans cette section est proposée une seconde contribution scientifique (en lien avec la seconde problématique scientifique) qui repose sur l'hybridation de méthodes de diagnostics existantes, plus particulièrement les modèles basés sur les connaissances (KBM) et les modèles basés sur les données (DDM). Ces méthodes ont été introduites dans le Chapitre 1 pour en identifier globalement les forces et les faiblesses. Ceci a permis d'établir que les faiblesses concernant le diagnostic de dérives de processus de production résident principalement dans la difficulté à identifier les causes racines physiques des dérives des processus de production hétérogènes. En effet ces processus peuvent impliquer des dizaines, voire des centaines d'éléments physiques pour leur fonctionnement. Cette complexité est augmentée par les relations dynamiques et non linéaires entre ces éléments. De plus, tout comme pour les techniques de détection de dérives, il existe une difficulté de passage à l'échelle des méthodes de qualification des dérives. Cela est dû à la diversité des processus, à leur nature non stationnaire, ainsi qu'à la nécessité de personnalisation des solutions limitant leur généralisation à grande échelle dans l'industrie.

Face à ces limitations, la thèse propose donc une approche de diagnostic basée sur les ensembles de classificateurs binaires (DDM), une technique d'apprentissage automatique adaptée à l'identification précise des dérives dans les profils de données (voir section 2.3.1.). Ce choix offre une solution évolutive et flexible, capable de s'adapter aux caractéristiques complexes et dynamiques des environnements de production hétérogènes. En complément, cette approche intègre un modèle basé sur les connaissances (KBM) qui associe les dérives identifiées par les DDM à une arborescence d'éléments physiques potentiellement responsables de ces dérives (voir section 2.3.2.). Cette structure hiérarchique permet d'établir un lien entre les dérives observées et les composants physiques impliqués dans les processus de production surveillés. Grâce à cette association, le diagnostic ne se limite pas à l'identification des dérives, mais fournit également une compréhension des causes physiques sous-jacentes, facilitant ainsi des interventions de maintenance ciblées et efficaces.

2.3.1 Choix d'une méthode de diagnostic basée sur les données pour l'identification précise des dérives

Dans le cadre du diagnostic des dérives, l'une des premières étapes essentielles consiste à identifier précisément les dérives observées, afin de pouvoir établir des liens entre les données et les causes racines. Cela nécessite l'utilisation d'une méthode capable d'apprendre à partir des données disponibles et de classer les dérives en fonction de leurs caractéristiques. Dans cette optique, une voie logique est de s'appuyer sur une méthode de diagnostic basée sur les données via l'apprentissage automatique. Les classificateurs sont les approches privilégiées dans ce contexte, car ils permettent de catégoriser les profils de dérives à partir de grandes quantités de données hétérogènes.

En effet (Sisode and Devare, 2023) a mis en évidence, dans le contexte du PHM, un intérêt croissant pour l'utilisation de méthodes de classificateurs pour le diagnostic des dérives dans les environnements industriels. Parmi ces méthodes, on retrouve principalement des classificateurs binaires et multiclassés. Par exemple (Jin *et al.*, 2024), ont utilisé des classificateurs multi-classes pour détecter et identifier différents types de défauts dans des moteurs électriques. Ces méthodes ont montré une bonne capacité à gérer la complexité des systèmes industriels. Toutefois, l'utilisation des classificateurs binaires tend à être privilégiée dans les travaux récents pour leur modularité et leur flexibilité.

Par exemple (Galar *et al.*, 2011), ont démontré que les classificateurs binaires permettent une approche plus granulaire et évolutive du diagnostic en divisant la tâche en plusieurs sous-classifications spécifiques. Cela simplifie l'ajout de nouvelles classes de dérives, ce qui répond au besoin d'adaptabilité dans des environnements

industriels évolutifs. De plus (Jha, Dave and Madan, 2019), montrent que l'utilisation de classificateurs binaires facilite la gestion des données déséquilibrées, une situation fréquente dans les systèmes industriels.

2.3.1.1 Choix des classificateurs binaires pour une évolutivité et flexibilité

Par conséquent, notre contribution en regard du diagnostic s'est positionnée sur l'utilisation des classificateurs binaires, surtout vis-à-vis de leur capacité à offrir une solution flexible et évolutive. En effet, contrairement aux classificateurs multiclassés, qui nécessitent la gestion simultanée de plusieurs classes, les classificateurs binaires permettent de segmenter le problème en plusieurs sous-problèmes indépendants. Cela favorise une meilleure gestion des données déséquilibrées et permet un ajout progressif de nouvelles classes de dérives sans remettre en cause l'intégrité du système existant (Galar *et al.*, 2011).

Cette modularité est cruciale dans des environnements industriels hétérogènes, où les dérives peuvent évoluer et changer au fil du temps. En séparant les problèmes de classification en décisions binaires, il devient alors possible d'adapter le système en fonction des besoins spécifiques de chaque contexte de production.

2.3.1.2 Proposition d'une solution d'apprentissage ensembliste pour le diagnostic

À partir de ce choix de classificateurs binaires, et afin d'améliorer la robustesse et la précision des diagnostics, il est proposé dans notre seconde contribution scientifique d'adopter une approche basée sur des **ensembles** de classificateurs binaires. En effet, l'**apprentissage d'ensemble** consiste à **combinaison plusieurs modèles** pour tirer parti de leurs forces respectives et compenser leurs faiblesses. Comme le montrent (Mian *et al.*, 2024), les méthodes d'ensemble offrent des performances accrues en améliorant à la fois la résilience face à la diversité des dérives et la précision des prédictions.

L'utilisation d'ensemble de classificateurs binaires permet également d'adapter les modèles aux variations spécifiques des données, ce qui est crucial dans des environnements industriels complexes. Cependant, cette approche présente aussi certaines limites. La coordination des réponses entre plusieurs classificateurs peut entraîner des incohérences, en particulier lorsque plusieurs modèles retournent des résultats contradictoires. Si un ou plusieurs classificateurs moins performants que les autres du modèle ensembliste influencent trop la décision finale, cela peut diluer la pertinence globale de la réponse. Ce cas de figure intervient surtout lorsque certains modèles ne sont pas adaptés à certaines situations spécifiques vis-à-vis des données classifiées.

2.3.1.2.1 Stratégie D'apprentissage d'ensemble : bagging, boosting et stacking

Par rapport à cette caractérisation d'apprentissage d'ensemble, trois stratégies d'apprentissage ensembliste sont généralement utilisées : le bagging, le boosting et le stacking (Ngo, Beard and Chandra, 2022; Sun *et al.*, 2023; Satoła and Satoła, 2024).

- Le **bagging** est une approche efficace lorsque les données sont bruitées ou qu'il existe une variance importante dans les processus surveillés. En générant plusieurs sous-ensembles de données à partir de l'ensemble initial, puis en combinant les prévisions des modèles individuels, le bagging permet de réduire la variance des prévisions. Dans notre contexte industriel où les données peuvent être hétérogènes et affectées par du bruit, cette méthode garantit une meilleure stabilité des diagnostics en évitant que les modèles soient trop sensibles à des anomalies locales dans les données. Le bagging est donc adapté pour les environnements où l'interprétabilité et la stabilité sont prioritaires, notamment si les données de dérives peuvent présenter de l'incertitude.
- Le **boosting**, en revanche, est plus adapté lorsque l'on souhaite améliorer la précision des modèles en ciblant spécifiquement les erreurs. Chaque modèle est entraîné successivement en se concentrant sur les échantillons mal classés par les modèles précédents. Cette architecture d'apprentissage permet de mieux traiter les erreurs systématiques. Cela fait du boosting un choix pertinent dans les environnements où certaines dérives sont rares, mais critiques et nécessitent une détection fine et précise. Cependant, cette approche est plus coûteuse en termes de temps de calcul et peut être sensible au surapprentissage, ce qui pourrait être un inconvénient dans notre contexte où certains processus sont très dynamiques et surtout évolutifs dans le temps.
- Enfin, le **stacking** combine plusieurs modèles de base en utilisant un modèle méta pour faire la prédiction finale. Cette méthode est idéale pour tirer parti des forces complémentaires de différents algorithmes, ce qui peut être pertinent dans notre contexte où nous avons des données complexes et

hétérogènes. Le stacking permet d'améliorer la précision des diagnostics en combinant plusieurs perspectives, mais il présente une plus grande complexité en termes de gestion des modèles. Son utilisation peut se justifier si les gains en précision sont essentiels et si le coût computationnel n'est pas un facteur limitant.

En complément de ces stratégies, le **vote** peut être utilisé comme une méthode simple, mais efficace pour combiner les prédictions des différents classificateurs dans un ensemble. Le vote peut être **majoritaire ou pondéré**, selon l'importance accordée à chaque modèle.

Au global, il n'existe pas de stratégie universelle pour l'apprentissage ensembliste. Le choix entre le bagging, le boosting, le stacking ou le vote doit dépendre des objectifs spécifiques attendus du modèle. Chaque méthode présente des avantages et des inconvénients en termes de précision, de complexité et de coût computationnel. Ainsi, la sélection de la stratégie la plus appropriée est avant tout un problème d'ingénierie avancée, devant être alignée avec les contraintes industrielles et les exigences de performance du système de diagnostic.

2.3.1.3 Proposition d'algorithmes de classification binaire adaptés au diagnostic de dérives

Aux vues des positionnements scientifiques décrits dans les sous-sections précédentes, il est donc acté que le diagnostic des dérives proposé dans notre contribution repose sur l'utilisation d'algorithmes de classification binaire. Cependant, comme pour le choix de l'algorithme de détection de nouveautés, le choix de l'algorithme de classification binaire le plus approprié dépend de divers facteurs, notamment la nature des données, la complexité des relations entre les variables, et les exigences en termes de précision et de performance. Il n'existe donc pas d'algorithme universel, ce qui nécessite d'adapter le choix à chaque cas d'application en fonction des données surveillées.

L'état de l'art des algorithmes de détection de dérive adapté aux séries temporelles multivariées, synthétisées dans le Tableau 1, a permis d'identifier un ensemble d'algorithmes utilisé pour des activités de détection de dérive dans la littérature du PHM. Certains de ces algorithmes sont adaptés aux activités de classification et sont par conséquent pertinents pour le diagnostic de dérives. Ces algorithmes sont résumés dans le Tableau 4, qui met en avant leurs spécificités en fonction des caractéristiques des données. Parmi eux, on retrouve : les forêts aléatoires (Random Forest), les machines à vecteurs de support (SVM), les k-plus proches voisins (KNN), les arbres de décision, les Gradient Boosting Machines (GBM), les réseaux de neurones convolutifs (CNN), les réseaux de neurones Long Short-Term Memory (LSTM), les réseaux de neurones récurrents (RNN), les autoencodeurs, et les réseaux de neurones à base de capsules (CapsNet).

Tableau 4. Récapitulatif des algorithmes de classification binaire adaptés au diagnostic de dérives et à leurs spécificités vis-à-vis des données.

Nom de l'algorithme	Spécificité en diagnostic de dérives vis-à-vis des données	Limitations potentielles vis-à-vis des données
Forêts aléatoires (Random Forest)	Robuste au surapprentissage et efficace pour des séries temporelles bruitées.	Moins efficace pour des données très hétérogènes ou fortement corrélées.
Machines à vecteurs de support (SVM)	Performant pour des séries temporelles bien séparées.	Moins performant pour des données bruitées ou avec des chevauchements entre classes.
K-plus proches voisins (KNN)	Flexible pour divers types de données.	Perd en précision avec des séries temporelles multivariées ou de grandes dimensions.
Arbres de décision	Interprétabilité élevée et faible coût de calcul, idéal pour des données facilement séparables.	Sensible au bruit et aux petites variations dans les données.

Gradient Boosting Machines (GBM)	Efficace pour des données complexes et bruitées grâce à une amélioration itérative des prédictions.	Peut perdre en efficacité si les données sont très déséquilibrées ou contiennent des valeurs extrêmes fréquentes.
Réseaux de neurones convolutifs (CNN)	Efficace pour extraire automatiquement des caractéristiques complexes, adaptée aux séries multivariées avec des patterns récurrents.	Moins performant sur des séries non structurées ou sans motifs spatiaux définis.
Long Short-Term Memory (LSTM)	Capture les longues dépendances temporelles.	Sensible aux données manquantes et aux séries temporelles très irrégulières.
Réseaux de neurones récurrents (RNN)	Capacité à modéliser les dépendances temporelles.	Moins efficace pour les séries longues avec dépendances complexes.
Autoencodeurs	Efficace pour la détection d'anomalies dans des données non linéaires.	Moins précis si les anomalies sont faibles ou si les données contiennent des variations subtiles.
Réseaux de neurones à base de capsules (CapsNet)	Efficace pour capturer des relations complexes entre des caractéristiques temporelles et spatiales.	Moins adapté pour des données sans structure spatiale ou sans relations complexes.

Globalement, chaque algorithme présenté dans le tableau précédent a des spécificités en termes de gestion des données et de performance computationnelle, ce qui influe directement sur leur applicabilité dans des environnements industriels complexes.

Ainsi, cette liste d'algorithmes, bien qu'elle ne soit pas exhaustive en raison de l'évolution rapide du domaine de l'apprentissage automatique, fournit une base essentielle pour isoler l'algorithme le plus adapté en regard du contexte de diagnostic de dérive visé. À plus grande échelle, cette proposition d'algorithmes soutient une approche flexible et générique, facilitant la mise en œuvre du diagnostic des dérives dans des contextes de production hétérogènes.

2.3.2 Complémentarité des algorithmes de classification binaire pour l'élaboration d'un classificateur ensembliste binaire de diagnostic

L'une des questions essentielles dans la mise en place d'une architecture d'apprentissage ensembliste telle que positionnée en section 2.3.1.2 réside dans la **complémentarité des algorithmes** utilisés. En effet, chaque algorithme possède des forces et des faiblesses spécifiques, et l'efficacité d'un modèle d'apprentissage ensembliste repose sur la capacité à combiner ces algorithmes de manière à ce que les faiblesses de certains soient compensées par les forces d'autres. Cette approche permet de maximiser la robustesse et les performances globales du modèle, en particulier dans des contextes complexes, tels que la classification binaire de séries temporelles multivariées.

Dans ce contexte, il est pertinent de réfléchir aux **indicateurs de complémentarité** entre les algorithmes afin de guider le choix des combinaisons à envisager. Ces indicateurs peuvent être extraits directement des problématiques de traitement des données explorées en section 1. En effet les méthodes d'apprentissage dépendent quasi intégralement des données et ce sont elles qui guident les choix.

Ainsi, le Tableau 5 présente les complémentarités potentielles entre les algorithmes en fonction des indicateurs définis. Les indices de complémentarité attribués permettent de quantifier la capacité d'un algorithme à compenser les faiblesses d'un autre de manière visuelle.

Tableau 5. Tableau de complémentarité des algorithmes de classification binaire au regard des indicateurs définis.

	1. Gestion des dépendances temporelles	2. Grands volumes de données	3. Capture des relations non linéaires	4. Gestion des données bruitées
Random Forest (1)	7, 8	5, 6	2, 3, 4, 5, 6	3, 4, 5, 9
SVM (2)	7, 8	5, 6	1, 3, 4, 5, 6	1, 3, 4, 5, 9
KNN (3)	7, 8	1, 5,6	1, 5, 9	1, 5, 9
Arbre de Décision (4)	7, 8	1, 5	1, 5, 9	5, 9
GBM (5)	7, 8	1, 6	1, 3, 4, 6, 9	1, 3, 4, 9
CNN (6)	—	1,5	1, 3, 4, 5	1, 3, 5
LSTM (7)	—	6	5, 6	1, 5
RNN (8)	—	6	5, 6	1, 5
Autoencodeurs (9)	7, 8	6	5, 6	1, 5
CapsNet (10)	—	—	9	—

Pour lire ce tableau, il faut suivre la démarche suivante :

- **Colonne 1** : Cette colonne liste les algorithmes d'apprentissage automatique pour la classification binaire (par exemple, Random Forest, SVM, KNN).
- **Colonnes suivantes** : Chaque colonne représente un critère de complémentarité, comme la gestion des dépendances temporelles, les grands volumes de données, la capture des relations non linéaires, et la gestion des données bruitées.
- **Cellules** : Les chiffres dans chaque cellule indiquent les indices des algorithmes (de la colonne 1) qui complètent bien l'algorithme de la ligne pour le critère en question. Par exemple, pour Random Forest sous "Gestion des dépendances temporelles," les chiffres "7, 8" indiquent que les algorithmes **LSTM (7)** et **RNN (8)** sont complémentaires pour ce critère.

En résumé, le Tableau 5 offre une représentation synthétique des complémentarités potentielles entre les algorithmes, facilitant ainsi la prise de décision pour la construction d'un modèle d'apprentissage ensembliste optimisé.

2.3.3 Méthodologie guidée pour le choix des algorithmes de classification binaire pour l'élaboration d'un classificateur ensembliste binaire de diagnostic

La méthodologie de choix des algorithmes à retenir pour le diagnostic des dérives repose sur une approche similaire à celle présentée en section 2.2.3 pour la sélection des algorithmes de détection de nouveauté. En effet, les données et les indicateurs extraits pour caractériser les séries temporelles sont identiques, qu'il s'agisse de la détection ou du diagnostic des dérives. Cette méthodologie (logigramme de décision) permet ainsi de maximiser les performances des algorithmes de classification binaire en fonction des caractéristiques des données disponibles, telles que mesurées par l'**Auto-Correlation Function (ACF)** et les **coefficients de corrélation de Pearson et Spearman**.

Cette méthodologie recoupe les résultats du Tableau 4, qui propose une sélection d'algorithmes de classification binaire adaptés au diagnostic des dérives, ainsi que ceux du Tableau 5 qui présente la complémentarité des algorithmes et s'appuie sur l'état de l'art des approches de diagnostic décrites dans le **Chapitre 1**.

La Figure 6 illustre cette méthodologie sous forme de logigramme, structurant les décisions de l'utilisateur à travers plusieurs étapes clés. Ces étapes incluent l'évaluation des dépendances temporelles, l'analyse des corrélations linéaires et non linéaires, ainsi que la nécessité ou non de transformer les données en fonction des indicateurs extraits.

- **Évaluation des dépendances temporelles (calcul de l'ACF)** : Si l'ACF est supérieure à 0,5, cela indique une dépendance temporelle significative dans les données. Il est alors recommandé de conserver les

données brutes et d'utiliser des algorithmes comme les RNN ou les LSTM, qui sont particulièrement adaptés à la gestion des relations temporelles.

- **Analyse des corrélations linéaires (coefficient de Pearson)** : Si l'ACF est inférieure à 0,5, on évalue le coefficient de Pearson. Si celui-ci est supérieur à 0,7, cela indique des relations linéaires fortes. Dans ce cas, une réduction de dimension linéaire (PCA) est appliquée, et des algorithmes, tels que KNN, SVM, ou Random Forest (RF) sont recommandés pour modéliser ces relations.
- **Analyse des corrélations non linéaires (coefficient de Spearman)** : Si le coefficient de Pearson est inférieur à 0,7, on poursuit avec le calcul du coefficient de Spearman. Si ce dernier dépasse 0,7, cela suggère des relations non linéaires importantes entre les variables. Une réduction de dimension non linéaire est alors appliquée (t-SNE, autoencodeurs), et des algorithmes non linéaires, comme Random Forest, GBM, ou autoencodeurs sont proposés.
- **Transformation des données (extraction statistique)** : Si les coefficients de Pearson et de Spearman sont tous deux inférieurs à 0,7, cela indique des relations faibles entre les variables. Dans ce cas, il est nécessaire de transformer les données en utilisant des techniques comme la FFT. Après transformation, des algorithmes tels que SVM, RF, KNN, ainsi que des réseaux de neurones plus complexes comme CNN et CapsNet peuvent être utilisés pour capturer les motifs présents dans les données transformées.
- **Sélection des algorithmes complémentaires pour un ensemble de classificateurs binaire** : Une fois les étapes précédentes finalisées, il est recommandé de combiner plusieurs algorithmes pour créer un classificateur binaire basé sur l'apprentissage ensembliste. Cette décision permet de renforcer la robustesse du modèle en associant plusieurs algorithmes complémentaires, par exemple en combinant Random Forest, GBM, ou une combinaison d'algorithmes linéaires et non linéaires. Le choix des algorithmes complémentaires se fait à partir du Tableau 5, qui récapitule les algorithmes les plus pertinents pour chaque situation, en fonction des relations identifiées et des caractéristiques des données.

Ainsi, la méthodologie proposée guide l'utilisateur à travers plusieurs étapes afin de décider si une transformation des données est nécessaire et, le cas échéant, quels algorithmes sont les plus appropriés en fonction des indicateurs extraits. Tout comme pour la méthodologie proposée en Section 2.4, la méthodologie présentée dans cette Section est générique à tout type de séries temporelles, car elle repose sur des indicateurs pertinents et universels. De plus, le logigramme de décision offre une approche cohérente pour le passage à l'échelle des algorithmes d'identification des dérives dans des contextes industriels, en permettant de choisir la solution la plus adaptée aux données étudiées et en guidant le choix d'algorithmes complémentaires nécessaires pour créer un ensemble de classificateurs binaires.

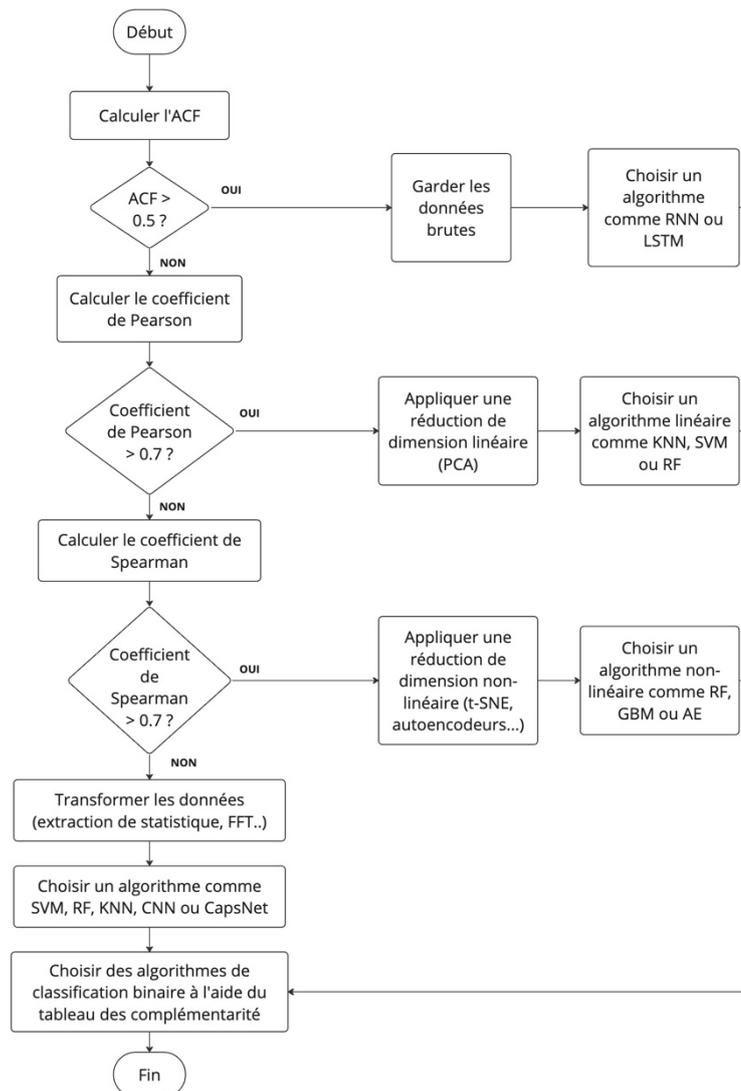


Figure 6. Méthodologie (Logigramme de décision) pour le traitement des données et le choix d'algorithme de diagnostic de dérive associé.

2.3.4 Choix d'une méthode de diagnostic basé sur les connaissances pour le diagnostic des dérives identifiées

En complément de la contribution proposée en section 2.3.3 sur une méthode de diagnostic basée donnée (DDM) pour une identification claire des dérives, il est nécessaire de compléter cette contribution en regard des attendus industriels de SEW USOCOME, par une méthode d'identification des causes de ces dérives (étape essentielle aussi du diagnostic). Les modèles basés sur les connaissances (KBM) sont a priori pertinents pour cette tâche, car ils offrent une interprétabilité et une capacité à établir des liens entre les dérives observées et les éléments physiques du système. Contrairement aux DDM, qui se concentrent principalement sur la détection des dérives à partir de données historiques, les KBM fournissent une structure explicite pour modéliser les relations entre les différents composants du système et les phénomènes observés (Chiachío *et al.*, 2019). Les KBM s'appuient sur des règles expertes, des arbres de décision, des ontologies ou des réseaux bayésiens pour représenter les relations causales entre les dérives et leurs causes potentielles (Zhou, Yu and Zhang, 2015; Verbert *et al.*, 2017). Cette approche est particulièrement efficace pour interpréter les dérives dans des systèmes complexes, où les composants sont fortement interdépendants. En associant les dérives identifiées par les DDM à des éléments physiques spécifiques, les KBM permettent de formuler des diagnostics exploitables et d'orienter les actions de maintenance à mener. En ce sens, dans les environnements industriels, la capacité à comprendre les causes des dérives est cruciale pour la maintenance prévisionnelle et la prise de décision.

Vis-à-vis des capacités mentionnées précédemment, l'approche KBM a donc été choisie, au niveau contribution scientifique, pour compléter la partie initiale d'identification des dérives par l'exploitation de DDM en raison de sa capacité à fournir des diagnostics détaillés et interprétables. Les dérives identifiées par les classificateurs binaires peuvent ainsi être reliées à une arborescence d'éléments physiques potentiellement impliqués. Cette approche est essentielle dans des environnements où l'interprétabilité et la compréhension des causes physiques sont primordiales pour garantir une intervention (Studer, Benjamins and Fensel, 1998).

L'approche KBM est également alignée avec la nécessité de passage à l'échelle, un enjeu crucial dans les environnements industriels complexes. En permettant de relier les labels de dérive à des composants spécifiques, les KBM facilitent la création de diagnostics modulaires. Chaque diagnostic peut être ajusté ou enrichi en fonction des évolutions des processus de production sans avoir à remanier l'ensemble du système. Cela offre une flexibilité et une adaptabilité importantes, permettant d'intégrer de nouvelles connaissances sans perturber l'intégrité du système de diagnostic global.

2.3.5 Analyse critique des méthodes de KBM

Les KBM proposent différentes méthodes pour diagnostiquer les dérives dans les systèmes industriels complexes. Parmi ces méthodes, et en référence aux éléments donnés pour le KBM dans le Chapitre 1 figurent les règles expertes, les arbres de décision, les ontologies et les réseaux bayésiens. Chacune présente des avantages et des inconvénients spécifiques, mais toutes rencontrent des défis lorsqu'il s'agit de les adapter à des environnements industriels hétérogènes et évolutifs.

2.3.5.1 Règles expertes

Les systèmes basés sur des règles expertes sont appréciés pour leur capacité à capturer et appliquer les connaissances des experts sous forme de règles logiques (Adamczyk et al., 2019). Ils permettent d'établir des liens directs entre les causes et les effets dans un processus donné, offrant une interprétabilité élevée. Cependant, leur principal inconvénient est leur rigidité et la nécessité de mise à jour constante par des experts, ce qui peut être difficile dans des environnements industriels complexes où les dérives évoluent (Aimiyekagbon et al., 2021). De plus, l'efficacité des systèmes basés sur des règles diminue lorsque le nombre de règles devient trop important, créant ainsi une difficulté à passer à l'échelle (Zhao et al., 2024). Cette limitation rend également difficile l'adaptation à de nouveaux scénarios ou l'intégration de connaissances évolutives.

2.3.5.2 Arbres de décision

Les arbres de décision sont largement utilisés pour leur simplicité et leur capacité à modéliser des relations complexes entre les variables d'entrée et de sortie (Dhakar, Singh and Gupta, 2023) ils offrent une structure hiérarchique claire qui facilite l'interprétation des résultats (Tso and Yau, 2007). Cependant, ils peuvent devenir trop spécifiques aux données d'entraînement, perdant en généralisation et étant sensibles aux variations mineures dans les données. Dans les systèmes industriels complexes, cette sensibilité peut entraîner des erreurs de diagnostic ou une perte de performance (El Mawas, Cappelle and El Badaoui El Najjar, 2023). Les arbres de décision nécessitent également des ajustements manuels fréquents pour s'adapter aux changements dans le système, ce qui limite leur efficacité dans des environnements dynamiques. De plus, les arbres de décision seuls peinent à gérer les relations non linéaires entre les variables, ce qui est souvent le cas dans des processus de production hétérogènes.

2.3.5.3 Ontologies

Les ontologies permettent de modéliser les connaissances de manière formelle, en représentant les relations hiérarchiques et sémantiques entre les concepts d'un domaine (He et al., 2020). Elles sont particulièrement adaptées pour intégrer et structurer des connaissances provenant de différentes sources, ce qui est crucial dans des environnements industriels où les processus sont interconnectés (Medina-Oliva et al., 2014). Cependant, leur développement et leur maintenance nécessitent une expertise spécialisée et sont généralement chronophages (Zhou, Yu and Zhang, 2015). Dans des applications en temps réel, la complexité des ontologies peut poser des défis en termes de performance et de rapidité de réponse (Nuñez and Borsato, 2017). De plus, la mise à jour des ontologies pour refléter les changements dans les processus physiques peut être un processus long et difficile, ce qui limite leur adaptabilité.

2.3.5.4 Réseaux bayésiens

Les réseaux bayésiens offrent une approche probabiliste pour modéliser les relations entre les variables, ce qui est crucial dans des environnements où l'incertitude est élevée (Mosallam, Medjaher and Zerhouni, 2016).

Ils permettent de représenter et d'inférer les relations causales entre les différents composants d'un système, offrant une compréhension approfondie des interactions complexes (Serras, Vinga and Carvalho, 2021). Cependant, leur mise en œuvre dans des environnements industriels hétérogènes est complexe, en particulier lorsque le nombre de variables est élevé (Castellano-Quero *et al.*, 2023). La construction d'un réseau bayésien nécessite une connaissance détaillée des relations conditionnelles entre les variables, ce qui peut être difficile à obtenir. De plus, les réseaux bayésiens peuvent être gourmands en ressources computationnelles, et l'interprétation des résultats peut s'avérer compliquée, en particulier pour les opérateurs non experts (Luo *et al.*, 2023). Ils sont également sensibles aux données incomplètes ou bruitées, ce qui peut entraîner des erreurs d'inférence.

2.3.5.5 Synthèse des méthodes de diagnostic basé sur les connaissances

Une synthèse de l'analyse critique des méthodes de diagnostic basées sur les connaissances et présentées dans la sous-section précédente permet de mettre en évidence que les arbres de décision sont le choix le plus approprié dans le contexte de cette recherche, en particulier en combinaison avec des DDM pour l'identification des dérives. Les arbres de décision sont conçus pour modéliser les relations causales entre les différentes composantes d'un système et les événements indésirables. Cette caractérisation est totalement alignée avec les attentes du diagnostic des processus de production hétérogènes, où les dérives peuvent impliquer des interactions entre plusieurs éléments physiques.

Contrairement aux règles expertes, qui s'avèrent limitées en raison de leur rigidité et de la difficulté à les maintenir à jour lorsque les systèmes évoluent, les arbres de décision offrent une plus grande flexibilité. De plus, les règles expertes nécessitent une mise à jour manuelle et dépendent fortement de connaissances expertes préalables, ce qui peut s'avérer difficile dans des environnements complexes. Au contraire, les arbres de décision permettent de modéliser des relations causales même avec des connaissances incomplètes et peuvent être enrichis progressivement à mesure que de nouvelles informations deviennent disponibles.

Dans un environnement où les dérives peuvent varier fréquemment en raison de changements dans les processus de production, les arbres de décision offrent une intégration simple des connaissances physiques liées aux processus, tout en étant faciles à mettre à jour. Une fois la structure causale établie, il est aisé de modifier ou d'ajouter des composants physiques sans nécessiter une révision complète de la structure.

Les ontologies, bien que puissantes pour modéliser les connaissances de manière sémantique, demandent non seulement un investissement considérable en termes de développement et de maintenance, mais elles nécessitent également des connaissances expertes étendues pour être correctement configurées. Leur construction repose fortement sur des experts capables de définir les relations sémantiques entre les différents concepts, et leur mise à jour est complexe lorsqu'il faut intégrer de nouvelles informations ou adapter les modèles aux évolutions des processus. De plus, leur complexité peut ralentir les processus de diagnostic en temps réel, rendant leur utilisation moins pratique dans des environnements où les processus évoluent rapidement.

Les réseaux bayésiens, quant à eux, nécessitent une connaissance approfondie des relations conditionnelles entre les variables pour leur construction. Dans des environnements industriels hétérogènes impliquant de nombreuses variables interdépendantes, il est souvent difficile de disposer de l'ensemble des informations nécessaires pour modéliser ces relations avec précision. Comme pour les ontologies, les réseaux bayésiens dépendent fortement de connaissances expertes préalables, et leur construction est donc plus complexe et chronophage. De plus, les exigences en termes de ressources computationnelles pour leur mise en œuvre et leur utilisation en temps réel sont élevées, limitant leur applicabilité dans des contextes où des diagnostics rapides sont essentiels (contraintes industrielles mises en évidence au Chapitre 1).

En intégrant les arbres de décision dans une approche hybride avec les DDM, l'efficacité des DDM pour l'identification des dérives à partir de données et la capacité des arbres de décision à fournir une interprétation claire des causes physiques des dérives sont combinées. Cette combinaison permet non seulement d'identifier les dérives, mais également de comprendre leur origine, ce qui est crucial pour la prise de décision et les interventions de maintenance dans l'industrie. Les arbres de décision sont particulièrement pertinents dans ce contexte, car ils permettent un diagnostic évolutif et interprétable des dérives dans des processus de production hétérogènes, tout en étant adaptables aux changements et à l'ajout de nouvelles connaissances au fil du temps, même en cas de connaissances expertes limitées.

2.3.6 Proposition d'une solution multimodèle de diagnostic basé sur des ensembles de classificateurs binaires et des arbres de décisions.

À partir des éléments défendus dans les sous-sections précédentes de la section 2.3 axée diagnostic de dérive, pour répondre à la problématique du **diagnostic des dérives** dans les processus de production hétérogènes, cette thèse propose une **approche multimodèle** combinant des **ensembles de classificateurs binaires** et des **arbres de décision**. Cette solution hybride exploite la puissance des modèles basés sur les données (DDM) pour identifier les dérives, tout en s'appuyant sur des arbres de décision pour relier ces dérives aux causes physiques sous-jacentes dans le système.

Cette approche, en intégrant les DDM et les arbres de décision, répond directement aux défis liés à la complexité des environnements industriels hétérogènes. En effet, l'association des DDM avec les KBM permet de réduire l'incertitude des DDM grâce aux informations expertes des KBM, ouvrant ainsi de nouvelles opportunités pour l'innovation dans les systèmes de maintenance prédictive, particulièrement pour l'analyse de données complexes et hétérogènes issues de capteurs ou de bases de données. Les DDM apportent leur capacité à gérer des volumes de données massifs et à identifier des dérives, même dans des scénarios où les données sont incomplètes ou bruitées. Toutefois, ils manquent souvent d'interprétabilité, ce qui rend difficile l'identification des causes racines. À cet égard, les KBM, et notamment les arbres de décision, permettent d'enrichir les DDM en fournissant des informations interprétables sur les causes physiques sous-jacentes. Des études récentes, telles que celles de (El Mawas, Cappelle and El Badaoui El Najjar, 2023; Peng *et al.*, 2023) montrent que cette approche hybride améliore la précision et l'interprétabilité du diagnostic dans des environnements industriels complexes, tout en s'adaptant à des processus dynamiques et évolutifs.

Les arbres de décision sont proposés comme une solution clé pour modéliser les relations causales entre les dérives identifiées par les DDM et les composants physiques d'un processus de production. Ils permettent d'établir des liens explicites entre les événements observés dans les processus de production hétérogènes et les éléments physiques responsables, offrant ainsi un diagnostic robuste et interprétable. Le développement de cet arbre de décision repose sur plusieurs étapes :

1. **Création initiale de l'arbre de décision** : Lorsque des connaissances expertes sont disponibles, un expert est chargé de créer l'arbre initial en reliant les labels de dérive générés par les DDM aux composants physiques impliqués. En l'absence de telles connaissances, il est essentiel de créer une liste des composants physiques du processus surveillé pour servir de base à la construction ultérieure de l'arbre.
2. **Mise à jour par un système de rétroaction** : Un des avantages majeurs des arbres de décision réside dans leur capacité à être mis à jour en continu grâce à un système de rétroaction. Après chaque dérive détectée et une intervention de maintenance, les techniciens renseignent le composant physique en cause ainsi que l'action corrective effectuée. Ce retour d'information permet d'enrichir automatiquement l'arbre de décision, en liant le label de dérive aux composants concernés. Des travaux tels que ceux de (Yildiz and Kalayci, 2024) soulignent l'efficacité de ces systèmes de rétroaction dans l'amélioration continue des diagnostics. Au fil du temps, cette boucle de rétroaction améliore la précision du diagnostic en affinant l'arbre, en pondérant davantage les composants fréquemment impliqués dans des dérives spécifiques.
3. **Association avec les DDM** : Les arbres de décision sont intégrés de manière complémentaire avec les DDM, qui identifient les dérives et génèrent des codes d'erreur représentant les types de défauts. Ces codes sont utilisés pour relier les dérives aux composants physiques pertinents via l'arbre de décision. Ce couplage permet non seulement de diagnostiquer la dérive, mais aussi de fournir des informations interprétables sur les causes racines, facilitant ainsi des actions correctives plus rapides et plus efficaces.
4. **Enrichissement itératif** : Au fil du temps, cette boucle de rétroaction améliore la précision du diagnostic en affinant l'arbre de décision. En pondérant davantage les composants physiques les plus souvent impliqués dans les dérives, les diagnostics deviennent plus précis. Cette approche est similaire aux travaux de (Ran *et al.*, 2019; Peng *et al.*, 2023), qui mettent en avant l'utilisation de la rétroaction continue pour affiner les modèles de diagnostic en maintenance prévisionnelle.

En conclusion, les arbres de décision offrent une solution robuste et flexible pour le diagnostic des dérives dans des environnements industriels complexes. Leur capacité à relier les dérives identifiées à des composants physiques spécifiques, tout en permettant une mise à jour continue via un système de rétroaction, fait d'eux un outil puissant pour améliorer la précision des diagnostics et optimiser les interventions de

maintenance. Cette approche, en intégrant les DDM et les arbres de décision, répond directement aux défis liés à la complexité des environnements industriels hétérogènes.

2.3.7 Synthèse de la contribution scientifique à la problématique de diagnostic des dérives de processus de production hétérogènes.

La contribution scientifique défendue dans cette section 3.1 relativement au diagnostic de dérive est double. Premièrement, elle propose une approche MM combinant les DDM et les KBM. Dans cette approche, le DDM est utilisé pour identifier les dérives dans les processus de production hétérogènes à l'aide d'ensembles de classificateurs binaires, tandis que le KBM permet d'identifier les causes racines des dérives grâce à l'intégration de connaissances expertes, structurées sous la forme d'une arborescence matérielle liée aux processus déviants. Cette hybridation permet non seulement de diagnostiquer les dérives de manière précise, mais aussi de comprendre et d'interpréter les causes physiques sous-jacentes à partir de la connaissance experte.

Deuxièmement, la contribution propose une méthodologie (logigramme de décision) pour la sélection des algorithmes de classification binaire (présentés Tableau 4) optimisant la mise en œuvre des modèles ensemblistes (à partir des complémentarités définies dans le Tableau 5) dans le cadre du DDM en fonction des caractéristiques des données.

Ainsi, cette contribution scientifique répond directement à la problématique du diagnostic des dérives dans les processus de production hétérogènes, identifiée au Chapitre 1. Cette contribution est générique est structurée sur deux aspects clés du diagnostic des dérives : l'identification des dérives et le diagnostic des causes racines. L'identification des dérives est guidée par le logigramme qui oriente l'utilisateur à travers des indicateurs analytiques standardisés, garantissant une flexibilité dans l'application à différents types de données et environnements industriels. Le diagnostic des causes racines repose sur l'intégration des modèles basés sur les connaissances (KBM), qui permettent à chaque utilisateur d'ajouter des connaissances expertes spécifiques à leur contexte industriel. Cette approche offre une grande adaptabilité, car elle peut être ajustée et personnalisée en fonction des besoins spécifiques de chaque secteur ou utilisateur. Ainsi, la méthodologie assure une solution robuste, flexible et facilement adaptable pour le diagnostic des dérives dans des environnements hétérogènes.

2.4 Intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes.

Dans la continuité des deux contributions précédentes axées Détection et Diagnostic de dérives, la troisième problématique scientifique attaquée par cette thèse concerne **l'intégration continue** de ces deux activités. En effet à ce stade, il est pertinent de proposer « un cadre » permettant d'unifier/réunir d'une certaine manière les deux contributions proposées de façon isolée. En ce sens, l'objectif de la contribution proposée dans cette section est de fournir une approche cohérente et évolutive capable de traiter l'ensemble du cycle de gestion des dérives, depuis leur détection jusqu'à leur diagnostic. Cette approche répond donc à la problématique scientifique **d'intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes.**

Par conséquent, cette approche méthodologique s'articule autour de trois phases de vie : **l'apprentissage de nouveauté, l'expertise des diagnostics et l'apprentissage continu.**

Elle cherche ainsi à intégrer les approches de **détection de nouveauté** pour la détection des dérives des processus et les **ensembles de classificateurs binaires** couplés à **la connaissance experte** pour les diagnostics de dérive. En ce sens elle surmonte les limitations observées dans les approches traditionnelles qui traitent ces problématiques de manière isolée. La flexibilité de cette approche permet également d'adapter les solutions aux environnements industriels spécifiques tout en garantissant une scalabilité et une généralisation à d'autres contextes.

Par ailleurs, l'approche méthodologique introduit une dynamique d'**apprentissage continu** qui permet de garantir une évolution constante des modèles utilisés pour la détection de dérives et leurs diagnostics. Cette boucle d'amélioration continue exploite les données collectées au fil du temps pour affiner les capacités des modèles de détection de dérive et de diagnostic, garantissant ainsi leurs performances face aux changements des processus industriels. Ce mécanisme permet également d'enrichir les bases de connaissances, notamment grâce à l'utilisation de techniques de **regroupement** et **des ensembles de classificateurs binaires**, en regroupant les dérives similaires et en formant de nouveaux modèles adaptés aux conditions réelles des processus de production.

Ainsi, l'approche méthodologique proposée, en intégrant de manière continue et complémentaire les contributions 1 et 2, propose une solution évolutive et générique. L'approche méthodologique répond aux besoins d'adaptabilité et de personnalisation des systèmes industriels modernes, tout en offrant un outil puissant pour gérer la complexité des processus de production des lignes de production modernes.

2.4.1 Justification de l'élaboration d'une approche méthodologique

Le choix d'une approche méthodologique pour répondre à la troisième problématique scientifique de cette thèse est justifié à la fois par des considérations scientifiques et industrielles. Une approche méthodologique permet de structurer de manière intégrée et évolutive les processus de détection et de diagnostic de dérives des processus de production hétérogènes. En effet, les environnements de production modernes sont caractérisés par leur hétérogénéité, leur non-stationnarité et leurs interactions complexes entre les processus et les composants (Kamm, Jazdi and Weyrich, 2021). Ces caractéristiques rendent nécessaire une approche méthodologique capable d'organiser et d'orchestrer l'intégration des différentes phases du cycle de vie des dérives, depuis leur détection jusqu'à leur diagnostic.

Les approches méthodologiques offrent plusieurs avantages spécifiques. Premièrement, elles facilitent l'intégration et la réutilisation des composants et des techniques de détection et de diagnostic. Cela est essentiel pour garantir une solution flexible et générique, capable de s'adapter aux spécificités des processus de production variés tout en répondant aux exigences de modularité et de scalabilité (Morgan *et al.*, 2021). Dans les environnements industriels dynamiques, où les processus évoluent constamment et où les conditions de production peuvent varier, une approche peu adaptable basée uniquement sur des techniques spécifiques risque de devenir rapidement obsolète, car ne tenant pas compte du retour d'expérience.

Deuxièmement, une approche méthodologique assure une cohérence entre les différentes étapes de détection et de diagnostic. Dans la littérature scientifique, plusieurs travaux montrent que la fragmentation des approches, c'est-à-dire le traitement de la détection et du diagnostic comme des étapes indépendantes, limite la capacité des systèmes à s'adapter aux environnements complexes et à intégrer les apprentissages passés (Chen *et al.*, 2023). En centralisant l'information et en standardisant les pratiques au sein d'une approche méthodologique, il devient possible d'améliorer la traçabilité des décisions, de renforcer la collaboration entre les équipes, et d'optimiser l'efficacité des diagnostics en temps réel.

Enfin, une approche méthodologique répond aux besoins industriels en matière de maintenabilité et de déploiement à grande échelle. En cohérence avec les contraintes mises en évidence au Chapitre 1 pour SEW USOCOME, il est essentiel que les solutions mises en place soient facilement déployables et maintenables, afin de garantir leur adoption à long terme. En apportant une structure organisée et adaptable, une approche méthodologique favorise la pérennité des solutions et limite les interventions humaines nécessaires pour ajuster ou mettre à jour les modèles. Plus précisément, cette capacité à intégrer des solutions dans une architecture flexible est particulièrement pertinente face aux verrous industriels, tels que le manque de savoir-faire spécifique et la gestion des coûts et du retour sur investissement (verrou industriel 4), car elle permet de réduire le besoin d'expertise et de minimiser les investissements en ressources humaines et techniques nécessaires au déploiement et à la maintenance des systèmes de diagnostic.

Globalement, le choix d'une approche méthodologique pour l'intégration des activités de détection et de diagnostic repose sur la nécessité d'avoir une approche modulaire, évolutive et générique, capable de répondre aux besoins des environnements de production hétérogènes tout en assurant une cohérence méthodologique. Cela permet de surmonter les verrous industriels liés à la complexité des processus (verrou 2), à l'adaptabilité des modèles (verrou 3), et à l'interopérabilité des systèmes (verrou 5), en proposant une solution globale et scalable.

2.4.2 Proposition d'une approche méthodologique en trois phases de vie

La conception de cette approche méthodologique en trois phases de vie résulte d'une prise en compte des spécificités inhérentes aux processus de production hétérogènes. Cette prise en compte repose sur deux grands enjeux. Le premier est l'absence de données exhaustives sur les dérives et les modes de dégradation des processus hétérogènes. Il est donc indispensable de concevoir une approche méthodologique capable de fonctionner même en l'absence de données préalables. Le deuxième enjeu est lié à la dynamique des environnements de production, où les systèmes et les processus sont en constante évolution, ce qui impose aux modèles de diagnostic et de détection d'être continuellement mis à jour pour rester pertinents. Ainsi, la structuration de la méthodologie s'est construite sur 3 phases :

La première phase, l'apprentissage de nouveauté, répond directement au besoin de détecter des dérives sans disposer de données historiques exhaustives sur les dérives spécifiques (verrou industriel 1). Dans

les environnements industriels, il est souvent impossible de connaître à l'avance tous les scénarios de dérives. Il devient donc crucial de détecter de nouveaux comportements anormaux dès qu'ils se manifestent, sans attendre qu'ils soient répertoriés. La détection de nouveauté, démontrée comme une solution optimale dans la première contribution scientifique, permet d'identifier des dérives sans nécessiter de labellisation préalable ou d'hypothèses fixes sur les données. Cette phase permet d'amorcer une surveillance des processus en se basant uniquement sur les données nominales, assurant ainsi une couverture efficace des dérives rares ou imprévisibles. Cette phase est cruciale pour la détection initiale des dérives dans des systèmes où la variabilité et l'évolution des conditions sont permanentes.

La deuxième phase, l'expertise des diagnostics, intervient pour traiter la complexité des interactions entre les composants et les processus de production en contextualisant les dérives identifiées. Cette phase s'appuie sur une étape de regroupement des données de nouveautés par similarité, simplifiant ainsi l'analyse et la classification des dérives. Les experts ajoutent ensuite leur connaissance pour caractériser ces profils et entraîner des modèles de diagnostic qui permettront d'associer chaque dérive à une cause potentielle. L'utilisation de classificateurs binaires en ensemble, démontrée dans la deuxième contribution scientifique, améliore la précision du diagnostic en modulant la classification en plusieurs sous-modèles spécialisés, augmentant ainsi la robustesse et la flexibilité du diagnostic. Cette approche offre également la possibilité d'adapter les modèles de diagnostic en fonction des spécificités des processus surveillés, garantissant ainsi une capacité de généralisation à une large gamme de scénarios industriels. Cette phase d'expertise est essentielle pour l'interprétation des dérives détectées et pour relier ces dérives aux causes physiques sous-jacentes, facilitant ainsi une intervention de maintenance ciblée et efficace.

La troisième phase, l'apprentissage continu, garantit que les modèles de l'approche méthodologique évoluent au fil du temps, en intégrant de nouvelles données et en améliorant les modèles existants. Dans un contexte industriel, les processus et les systèmes sont en constante évolution, et il est crucial que les modèles de détection et de diagnostic ne deviennent pas obsolètes. Cette phase permet non seulement d'enrichir la base de données des nouveautés avec des comportements non reconnus, mais aussi d'améliorer les modèles de diagnostic grâce aux retours d'expérience accumulés. Cette boucle d'apprentissage continu, qui intègre des éléments de la première et de la deuxième contribution scientifique, garantit une amélioration constante de la capacité du système à détecter et diagnostiquer les dérives, tout en assurant une flexibilité et une évolutivité optimales. Le système peut ainsi répondre aux changements dynamiques des processus industriels tout en maintenant une haute précision des diagnostics et une pertinence des interventions.

En somme, la structuration en trois phases de vie répond à la fois aux exigences scientifiques et aux besoins industriels identifiés. Chaque phase s'appuie sur des contributions précédentes, tout en adressant des verrous industriels spécifiques. Cette approche permet de gérer les dérives de manière cohérente, depuis leur détection jusqu'à leur diagnostic et leur résolution, tout en garantissant une évolution continue du système. Ces trois phases, interconnectées et complémentaires, permettent de concevoir une approche méthodologique matérialisée par un enchaînement d'activités adaptées.

2.4.3 Présentation de l'approche méthodologique

Les trois phases de vie de l'approche méthodologique abordent de manière ciblée les trois questions de recherche ayant conduit aux trois contributions scientifiques.

Ainsi, l'approche méthodologique est divisée en trois phases de vie pour mener à bien ses objectifs de détection de dérive et de diagnostic des processus de production hétérogènes, tout en assurant une généralité des méthodes et une adaptation continue des modèles à l'évolution des processus (voir Figure 7). Chacune des phases est construite sur un ensemble d'activités qui s'enchaînent en sachant que les objectifs précis de ces trois phases sont :

1. **Apprentissage de nouveauté** (Figure 7 P1) : Grâce à la détection des « nouveautés », les dérives de fonctionnement sont initialement identifiées à partir d'un ensemble de données réduit et représentatif des cycles nominaux. Cette phase implique l'utilisation d'un modèle de détection des nouveautés pour déterminer de nouvelles données inédites (Figure 7 P1.1), qui sont ensuite utilisées pour enrichir une base de données de « nouveautés » (Figure 7 P1.2). Une fois la base de données de « nouveautés » assez dense, l'approche méthodologique passe à la phase d'expertise des diagnostics.
2. **Expertise des diagnostics** (Figure 7 P2) : La base de données de « nouveautés » est soumise à un processus de regroupement pour regrouper les « nouveautés » similaires (Figure 7 P2.1). Les experts analysent les clusters pour ajouter des connaissances, en caractérisant les « nouveautés » détectées et en les attribuant à des profils de processus, qui peuvent être soit nominaux, soit indicatifs d'une dérive

(Figure 7 P2.2). Un profil est un ensemble de données qui décrit un état de fonctionnement ou un comportement particulier d'un processus. Les modèles de diagnostic sont ensuite (ré) entraînés sur ces profils de dérive à l'aide de classificateurs binaires d'ensemble (Figure 7 P2.3). Pour chaque profil de dérive, une expertise s'appuyant sur les connaissances humaines et les expériences réelles des systèmes d'information (par exemple, un système informatisé de gestion de la maintenance ou GMAO) est ajoutée aux modèles de diagnostic afin de relier la dérive aux actifs physiques. Le modèle de détection des nouveautés est également amélioré avec des profils nominaux qui n'avaient pas été détectés auparavant (Figure 7 P2.4). Enfin, l'approche méthodologique passe à la phase d'apprentissage continu sur décision manuelle d'un expert.

3. **Apprentissage continu** (Figure 7 P3) : Cette phase intègre la détection des dérives et le diagnostic par l'utilisation combinée de modèles de détection et de diagnostic des nouveautés ainsi que l'enrichissement en permanence de la base de données des « nouveautés » avec des données non reconnues par les modèles de détection et de diagnostic (Figure 7 P3.1). L'apprentissage continu garantit que le système évolue et améliore ses capacités de détection et de diagnostic au fil du temps via une boucle d'amélioration continue entre les phases P3 et P2.

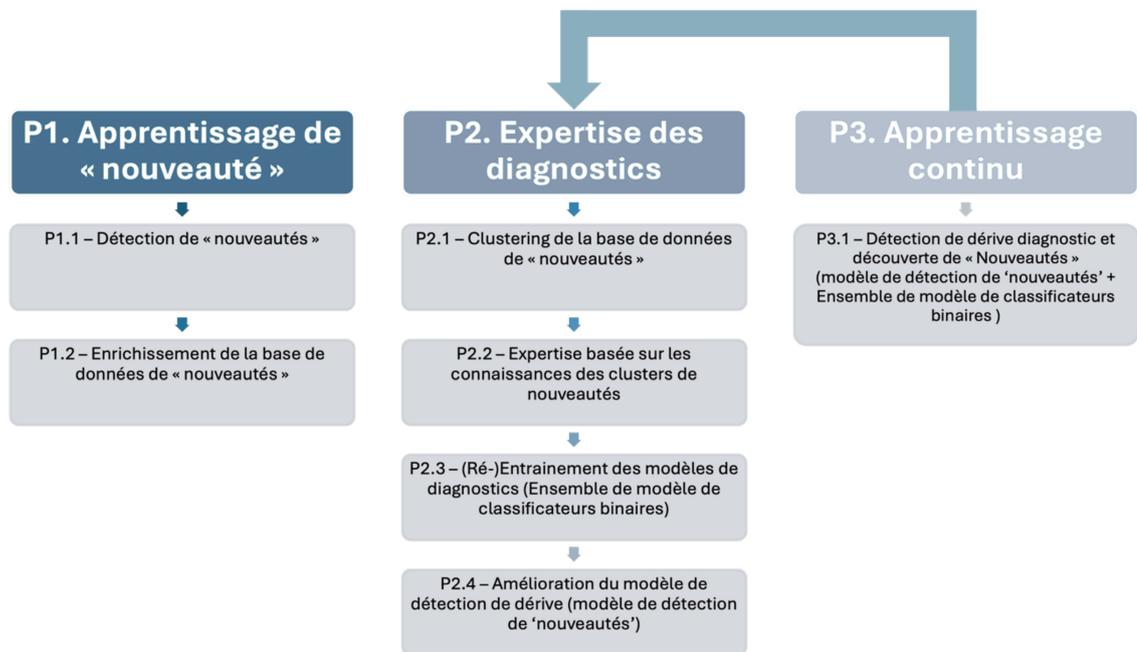


Figure 7. Principales phases de vie de l'approche méthodologique et fonctions associées

Le diagramme de flux de la Figure 8 est une représentation complète de l'approche méthodologique du point de vue opérationnel (enchaînement d'activités). La figure détaille le processus séquentiel et intégré d'apprentissage des nouveautés, d'expertise des diagnostics et d'apprentissage continu.

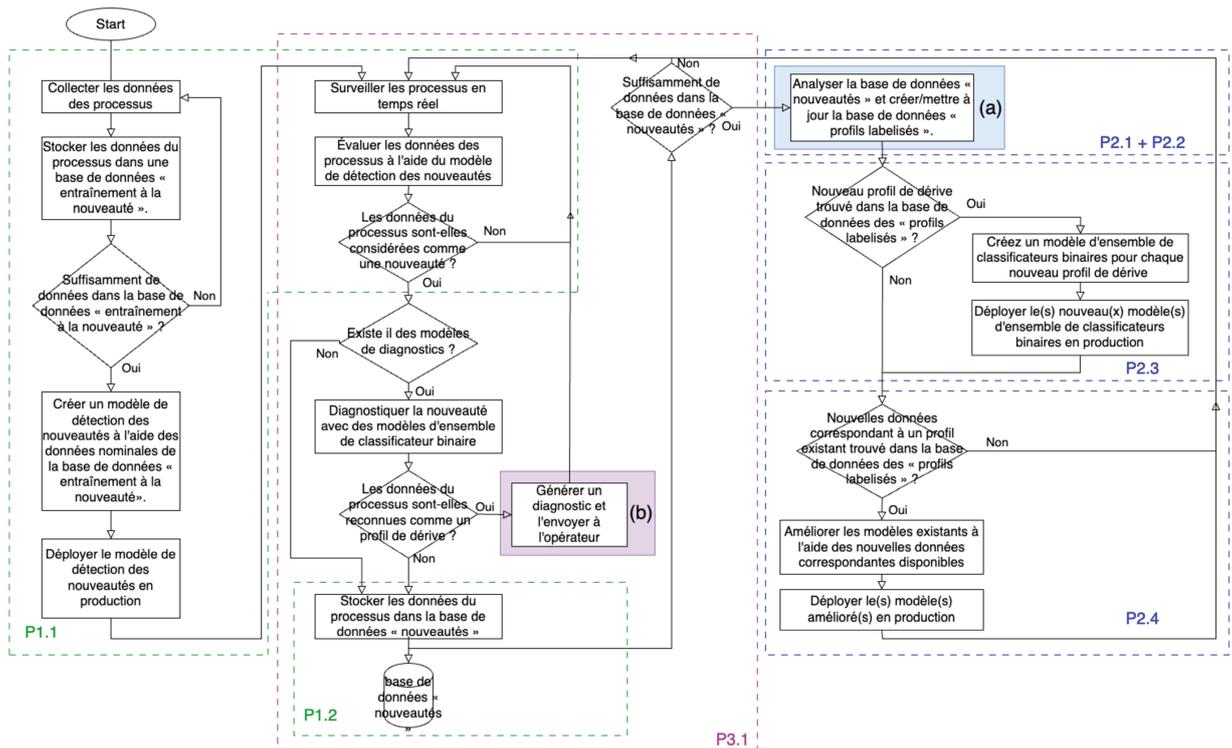


Figure 8. Diagramme de flux du cadre complet déployé (Px.y se réfère aux sous-phases dans la Figure 7). Les actions mises en évidence dans les zones (a) et (b) sont détaillées dans les Figure 12 et Figure 15

L'approche méthodologique a aussi été modélisée en exploitant des principes syntaxiques du langage SADT (mais les figures proposées ne sont que des schémas explicatifs non de véritables diagrammes SADT). Les figures Figure 9 et Figure 10 montrent cette modélisation avec une vue claire des étapes et des interactions de haut niveau du système.

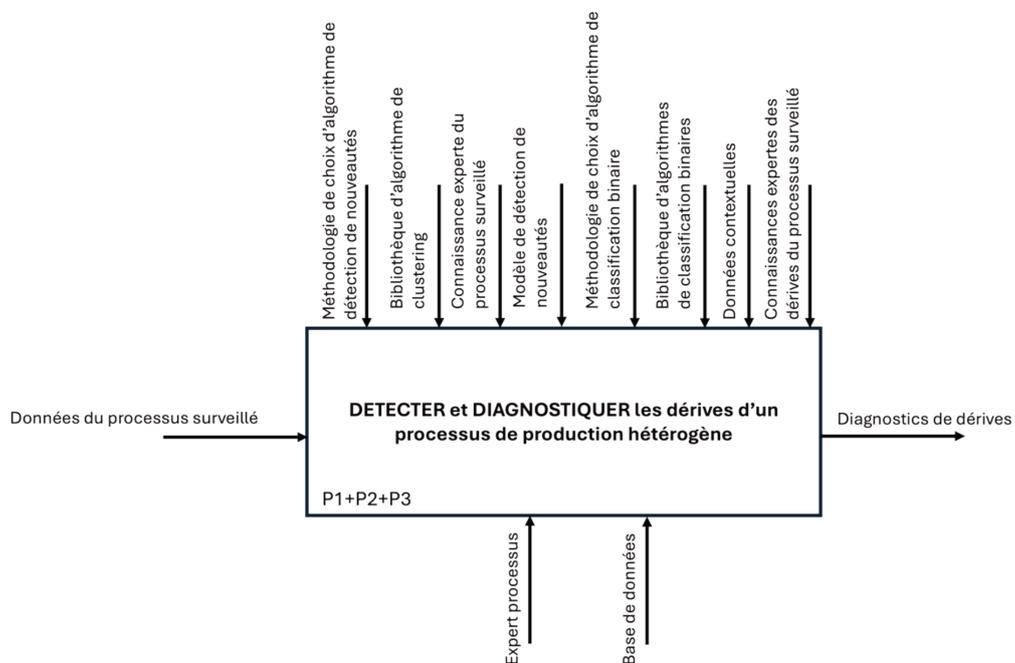


Figure 9. Illustration schématique à base des principes SADT de la méthodologie

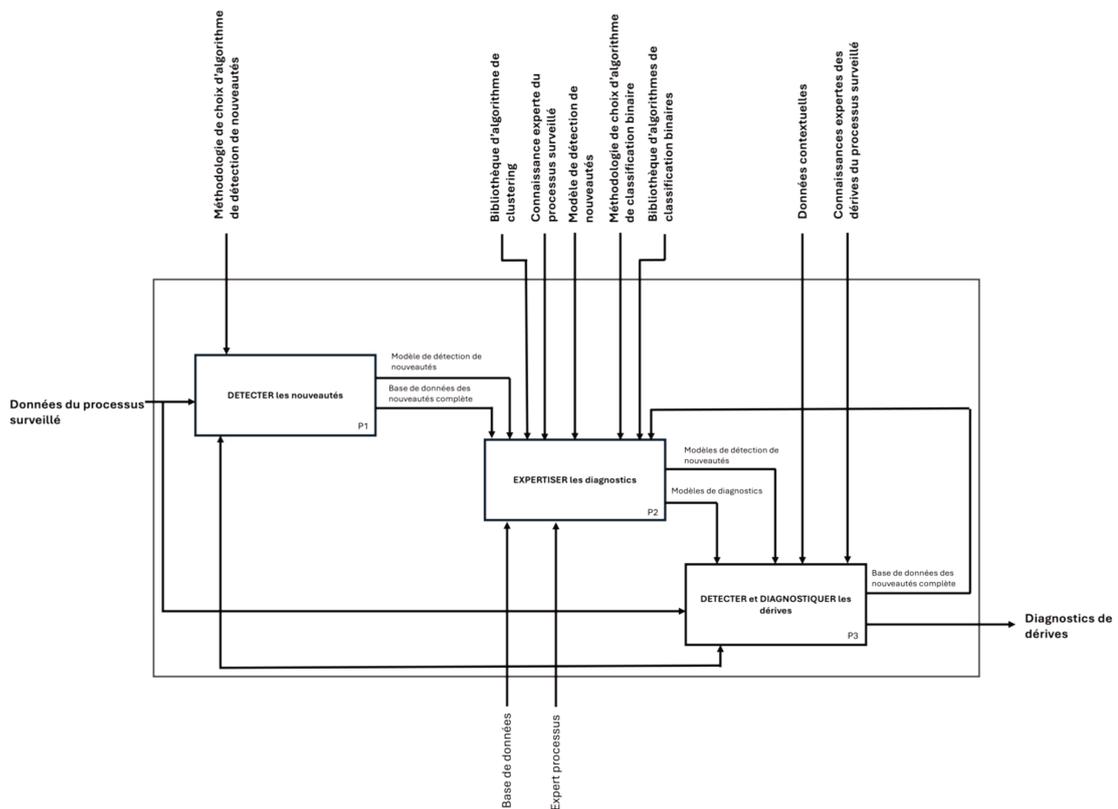


Figure 10. Illustration schématique de la décomposition de la fonction initiale « Détecter/Diagnostiquer » structurant les phases P1, P2 et P3 de l’approche méthodologique.

À partir de ces deux figures schématiques, il est possible ensuite d’expliquer chaque phase de l’approche méthodologique ainsi que ses activités internes respectives en explicitant les choix menés.

2.4.4 Étapes de l’approche méthodologique

2.4.4.1 Phase d’apprentissage de « nouveauté »

La première phase, « apprentissage de « nouveauté » », se concentre sur la collecte de nouveaux schémas de données de processus de production. Cette phase consiste à établir une base de données de référence qui représente les cycles de processus conformes. L’objectif est de développer des mécanismes capables de détecter les déviations dans les données surveillées en temps réel vis-à-vis de la base de données de référence. Ces données déviantes sont désignées ici comme des « nouveautés ». Une fois repérées, ces nouveautés sont enregistrées dans une base de données pour un suivi et une analyse plus approfondie dans les phases suivantes de l’approche méthodologique. Les profils de nouveautés identifiables dans cette phase peuvent inclure de nouveaux comportements conformes, ou bien des dérives de tous types. Ces dérives peuvent être des écarts significatifs par rapport aux normes établies par les profils conformes, ou bien des fluctuations irrégulières ou aléatoires, par exemple.

Comme le montre la Figure 7, la phase d'apprentissage de nouveauté comprend deux étapes : la détection des dérives (P1.1) et l'enrichissement de la base de données des « nouveautés » (P1.2)

En complément, la Figure 11 représente un schéma illustratif inspiré du langage SADT de la phase de vie P1.

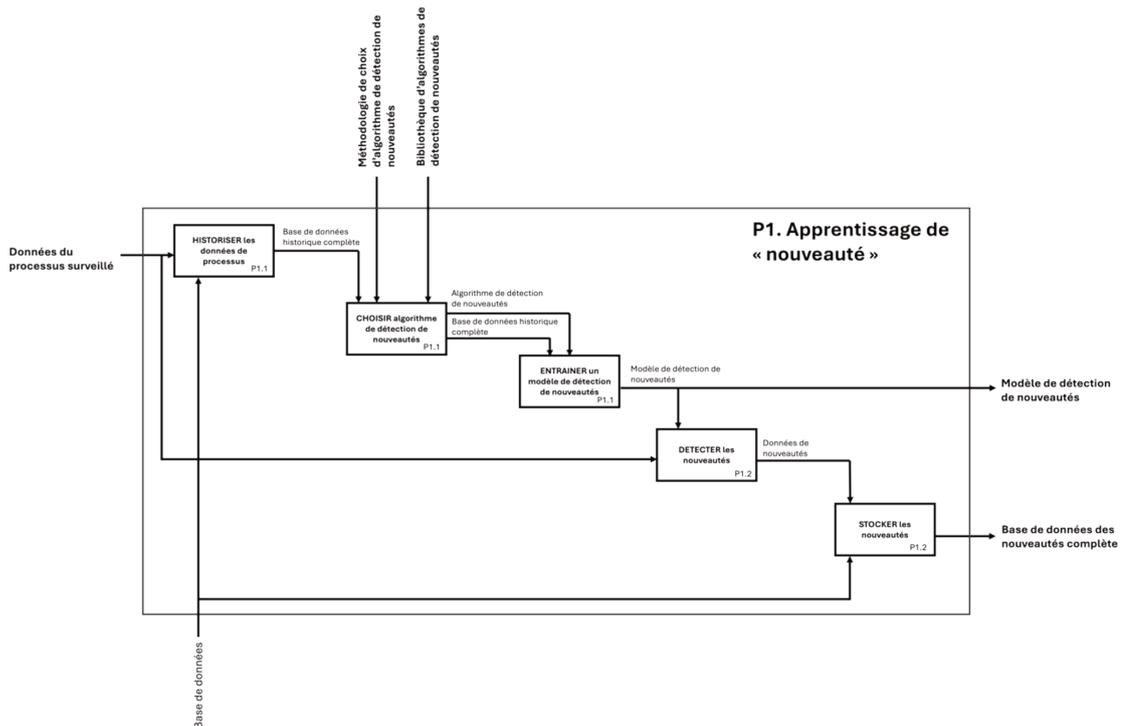


Figure 11. Illustration schématique de l'interaction entre les étapes internes de la phase P1 de l'approche méthodologique.

2.4.4.2 P1.1 - Détection des « nouveautés »

L'étape P1.1, « Détection des « nouveautés » », a pour objectif de mettre en place une solution de surveillance capable de détecter les dérives dans les processus de production en temps réel. Cette étape commence par le choix d'une méthode de détection en exploitant la première contribution scientifique synthétisée par le logigramme de décision représenté en Figure 5. Une fois la solution de détection de dérive sélectionnée, il est nécessaire de collecter des données de processus conformes qui serviront à entraîner le modèle de détection de dérive. Sur cette base, un modèle de *détection de nouveautés* est entraîné. Une fois entraîné, le modèle est déployé en production pour surveiller continuellement les processus et identifier toute déviation par rapport à la nominalité établie par les cycles conformes.

2.4.4.3 P1.2 - Enrichissement de la base de données « Nouveauté »

L'étape P1.2, intitulée « Enrichissement de la base de données des « Nouveautés » », se concentre sur l'historisation des déviations identifiées lors de la détection des « nouveautés » (P1.1). Le but de cette étape est de créer une base de données contenant les « nouveautés » en y ajoutant de manière systématique les données caractérisées comme des « nouveautés » par l'algorithme de détection de nouveautés.

Lorsque la base de données des « nouveauté » contient suffisamment de données pour créer les premiers diagnostics de dérive, l'approche méthodologique peut alors passer à sa phase de vie P2 - Expertise des diagnostics. Le nombre de données nécessaire pour déclencher la transition vers la phase P2 se fait selon un paramètre défini manuellement par un expert ou selon un échancier pouvant évoluer suivant la maturité de la solution (plus fréquent au départ, quand les nouveautés risquent d'être plus fréquemment présentes, puis moins fréquemment).

2.4.5 Phase d'expertise diagnostic

La deuxième phase, « expertise des diagnostics », utilise la base de données de « nouveautés » pour regrouper les données similaires en différents groupes. Ces groupes sont ensuite analysés par des experts qui ajoutent des connaissances expertes et contextuelles afin de caractériser les groupes de nouveautés au mieux. L'objectif est d'attribuer un profil à ces groupes, qu'ils soient conformes ou symptomatiques de dérives. Cette

phase permet de créer/améliorer les modèles de diagnostic de dérives en associant les profils de données avec la connaissance experte issues de l'expertise humaine et de l'expérience terrain. En intégrant l'expertise humaine et l'expérience terrain (par exemple, issues de la GMAO), l'approche méthodologique permet de relier les dérives détectées aux actifs physiques, améliorant ainsi la précision et la pertinence des modèles de diagnostic. Cette phase pouvant être exécutée plusieurs fois garantit la capacité d'adaptation du système via une dynamique d'apprentissage continu, améliorant ainsi sa capacité à détecter et à diagnostiquer les dérives dans les systèmes de production complexes au fil du temps. En effet, les nouveaux profils de données conformes viendront améliorer le modèle de détection de nouveautés, les profils de données non conformes correspondant à des dérives connues viendront améliorer les modèles de diagnostics associés et les nouveaux profils non conformes serviront à créer de nouveaux diagnostics de dérives.

Conformément à la structure de l'approche méthodologique présentée en Figure 7, la phase d'expertise diagnostic comprend 4 étapes : regroupement de la base de données « nouveautés » (P2.1), expertise fondée sur les connaissances sur les clusters de nouveauté (P2.2) (re) entraînement du modèle de diagnostic (P2.3) et amélioration du modèle de détection de dérive (P2.4). Les étapes d'expertise de la base de données « nouveauté » et d'expertise fondée sur les connaissances sont représentées dans la Figure 8 en tant que bloc d'action unique. La Figure 12 se concentre sur ce bloc d'action spécifique et décrit plus en détail le regroupement de la base de données « nouveauté » et des étapes d'expertise basées sur les connaissances.

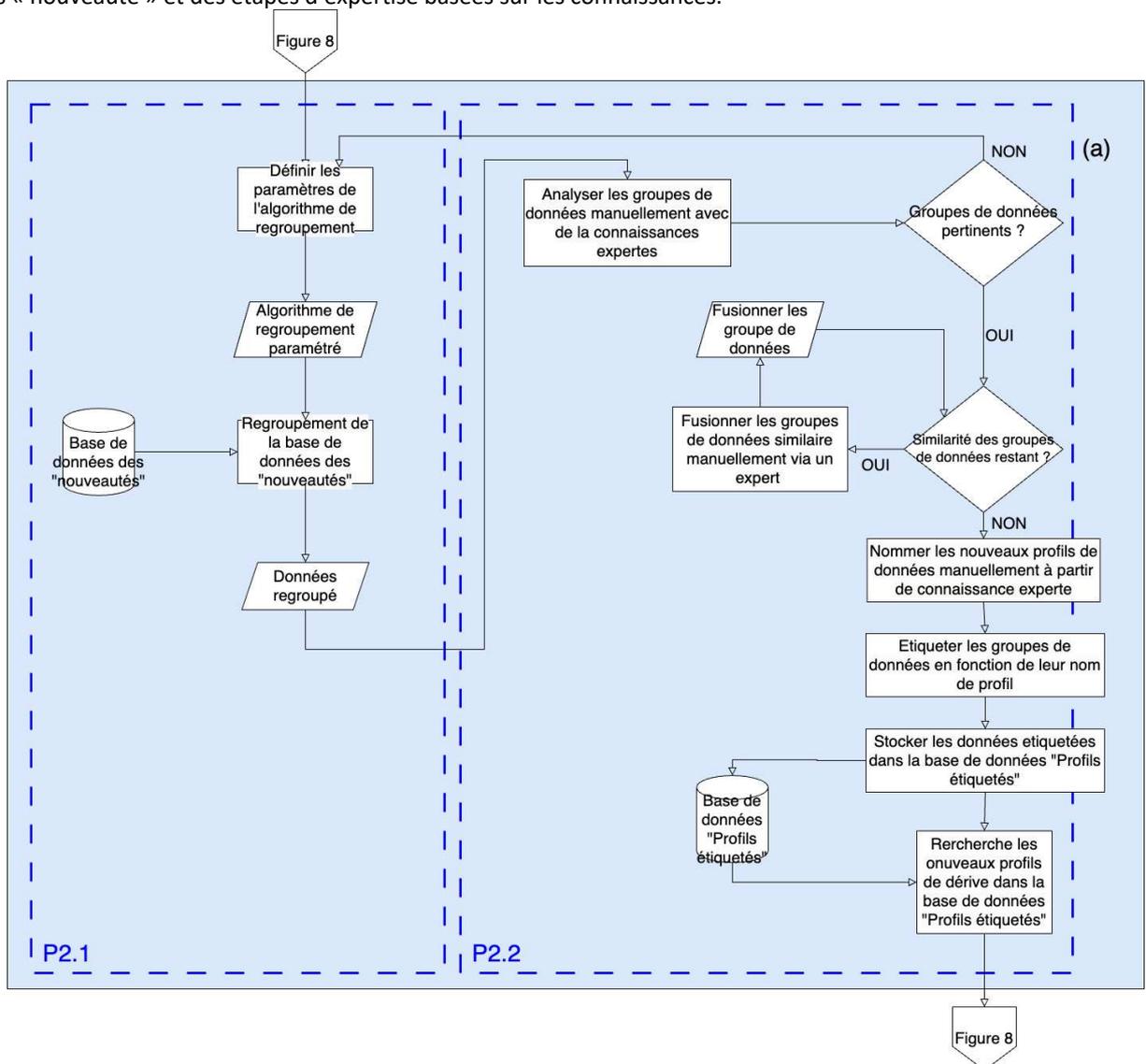


Figure 12. Représentation détaillée sous forme de diagramme de flux du bloc d'action de la Figure 8a, indiquant le regroupement de la base de données "nouveautés" et les étapes d'expertise basée sur la connaissance.

La Figure 13 représente un schéma illustratif basé SADT de la phase de vie P2.

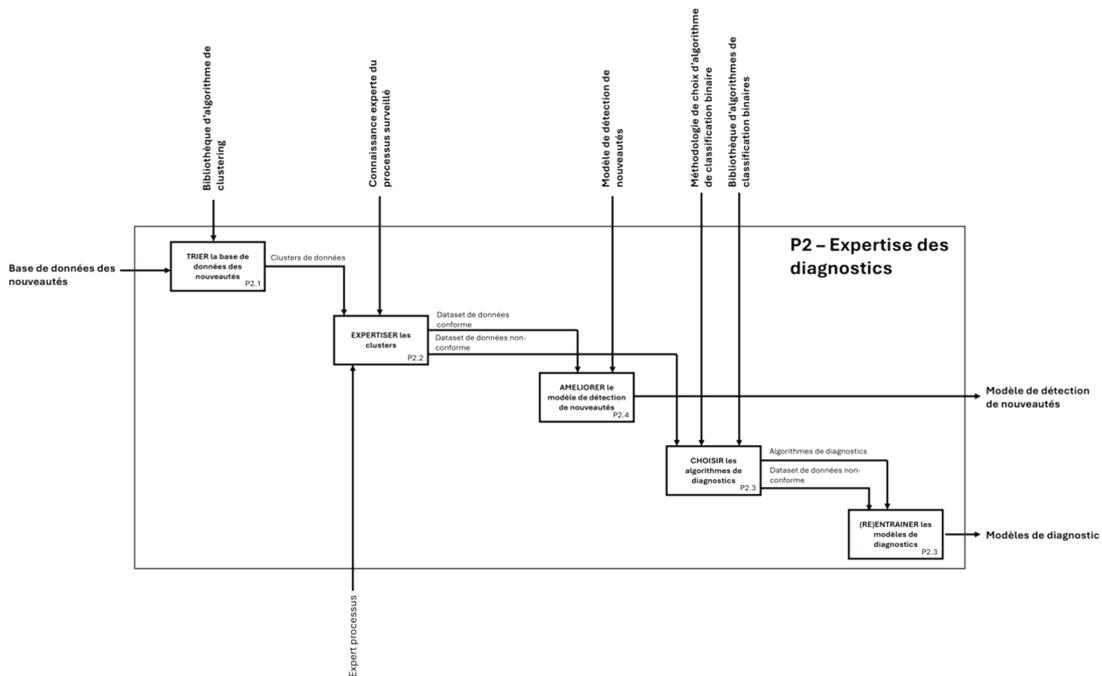


Figure 13. Illustration schématique de l'interaction entre les étapes internes de la phase P2 de l'approche méthodologique.

2.4.5.1 P2.1 - Regroupement de la base de données « nouveauté »

L'étape P2.1, « Regroupement de la base de données de « nouveautés » », a pour objectif de trier les données collectées durant l'étape P1.2 pour faciliter le processus de création des diagnostics des étapes P2.2 et P2.3.

Dans cette étape, la base de données de « nouveauté » peut contenir un énorme volume de données important et plusieurs variétés de nouveaux profils. Cette phase de tri peut être automatisée via des techniques d'intelligence artificielle dite de regroupement, comme le démontrent les travaux de (Fahad *et al.*, 2014). Les techniques de regroupement, telles que le K-means, les iForest et le DBSCAN permettent le traitement automatique des données de la base de données « nouveauté » pour les regrouper en groupes de profils de données similaires appelées « clusters ». Les clusters peuvent ensuite être analysés plus en détail par des experts à l'étape suivante de l'approche méthodologique.

2.4.5.2 P2.2 - Expertise basée sur la connaissance des clusters de nouveauté

L'étape P2.2, « Expertise basée sur la connaissance des clusters de nouveauté », permet aux experts d'analyser et de classer les clusters créés lors de l'étape P2.1. Ces clusters étiquetés serviront à développer des modèles de diagnostic pour chaque dérive à l'étape P2.3 et à améliorer le système de détection des dérives à l'étape P2.4. Durant cette étape, les experts identifient les clusters correspondant à des comportements conformes ou non conformes et, si nécessaire, fusionnent les clusters similaires. Le nom des nouveaux profils est déterminé par un mélange des résultats du regroupement et de l'évaluation des experts.

Concrètement, les experts examinent chaque cluster issu de P2.1, en évaluent les caractéristiques, et décident s'il représente un nouveau profil conforme, non conforme ou une variation d'un profil existant. Ce processus assure que la classification est à la fois guidée par les données et validée par des experts. Leur rôle est essentiel pour affiner et ajuster les clusters, surtout lorsque les méthodes automatiques ne parviennent pas à tout distinguer. Cette démarche itérative permet d'identifier des profils opérationnels plus précis et nuancés.

2.4.5.3 P2.3 - (re) entraînement des modèles de diagnostic

L'étape P2.3, « (re) Entraînement des modèles de diagnostic », propose une méthode pour la création de modèles de diagnostic utilisant des ensembles de classificateurs binaires. Ces modèles sont entraînés à partir des clusters de données expertisés à l'étape P2.2. L'approche se base sur une combinaison de modèles de

diagnostic basé sur les connaissances (KBM) et de modèles basés sur les données (DDM). Le choix de cette approche MM est motivé par l'état de l'art de l'activité de diagnostic du PHM et des considérations pratiques concernant la robustesse, l'adaptabilité et l'exactitude des diagnostics dans un contexte industriel complexe.

2.4.5.3.1 Proposition de combinaisons d'algorithmes pour les ensembles de classificateurs

Pour maximiser la robustesse et la précision des diagnostics de dérives de processus industriels, l'approche méthodologique recommande d'utiliser des combinaisons d'algorithmes aux forces et faiblesses complémentaires. Par exemple, une combinaison de RF et de LSTM peut s'avérer pertinente. Les RF sont reconnues pour leur capacité à gérer des données bruitées et déséquilibrées grâce à leur mécanisme d'échantillonnage et d'agrégation de multiples arbres de décision. Comme le montre (Mian *et al.*, 2024), ce mécanisme permet d'améliorer la robustesse des diagnostics. En parallèle, comme le montrent (Luo, Huang and Zhou, 2018; Aydemir *et al.*, 2021), les LSTM sont spécialisées dans la capture des dépendances temporelles à long terme, ce qui les rend efficaces pour analyser des séquences temporelles complexes. En combinant ces deux approches, l'ensemble de classificateurs devient capable de capturer à la fois des variations locales (gérées par les RF) et des dynamiques temporelles plus globales (gérées par les LSTM), améliorant la capacité globale de diagnostic.

Un autre exemple peut être l'utilisation d'un ensemble de classificateurs qui combinent les SVM, les CNN et les LSTM, offrant des performances encore meilleures. Les SVM sont efficaces pour optimiser la séparation des classes en maximisant les marges, ce qui est essentiel pour une classification précise, en particulier dans des environnements industriels où les limites entre états normaux et anormaux peuvent être subtiles (Xiao *et al.*, 2016). Les CNN, quant à eux, ont une excellente capacité à détecter des motifs complexes et structurés dans les données, ce qui est un atout pour le diagnostic basé sur des données comme des images ou des séries temporelles (Canizo *et al.*, 2019). Enfin, les LSTM apportent des informations contextuelles essentielles en conservant la mémoire des séquences passées, ce qui permet de mieux comprendre les évolutions temporelles des processus (Dong, Li and Sun, 2017).

En combinant ces algorithmes un ensemble de classificateurs entraîné avec une stratégie de stacking par exemple, le modèle de diagnostic sera capable de compenser les faiblesses individuelles de chaque algorithme et d'exploiter leurs forces complémentaires. Dans cet exemple, l'ajout du CNN et du LSTM au modèle de SVM permet d'enrichir les capacités de diagnostic en identifiant les variations subtiles et en s'adaptant aux contextes dynamiques des données du processus de production. Cette approche d'apprentissage d'ensemble permet d'obtenir des modèles de diagnostic plus robuste, plus précis et mieux adapté à la complexité des processus industriels.

2.4.5.4 P2.4 - Amélioration du modèle de détection de dérive

L'étape P2.4, « Amélioration du modèle de détection de dérive », intervient après les trois premières étapes du processus de diagnostic pour intégrer les nouvelles données et connaissances acquises sur les modes de fonctionnement conforme des processus. Les données récoltées à l'étape P1.2 dans la base de données des nouveautés peuvent comporter des nouveaux profils de données correspondant à des comportements conformes. Ces données, analysées et étiquetées par des experts à l'étape P2.2, sont utilisées pour améliorer le modèle de détection de nouveautés. Ce processus d'enrichissement consiste à réentraîner le modèle de détection de nouveautés existant en utilisant un nouveau jeu de données comprenant à la fois les données historiques collectées à l'étape P1.1 et les nouvelles données étiquetées comme conformes. Cette approche permet de s'assurer que les nouveaux modèles de données, désormais reconnus comme conformes, ne seront plus détectés à tort comme des nouveautés par le modèle de détection de nouveautés. L'amélioration continue du modèle de détection à travers cette méthode de réentraînement est essentielle pour maintenir une précision élevée et une adaptabilité du système.

Une fois le modèle de détection de nouveautés mis à jour, l'approche méthodologique peut passer à sa phase P3 d'apprentissage continu. L'objectif de la phase 3 est de permettre à l'approche méthodologique d'améliorer constamment sa précision et sa capacité à détecter les dérives et les diagnostiquer en utilisant les nouvelles données collectées et les retours d'expérience récents.

2.4.6 Phase d'apprentissage continu

La troisième phase, intitulée « apprentissage continu », vise à exploiter les mécanismes de détection et de diagnostic développés dans les phases précédentes pour surveiller les processus en continu. Cette surveillance permet de détecter et diagnostiquer les dérives en temps réel ou à enrichir la base de données de nouveautés lorsque les données ne sont pas identifiées par les mécanismes de diagnostic en place. Cette phase comprend

également un mécanisme de retour vers l'étape d'expertise diagnostic, dans le but de réévaluer et d'améliorer continuellement les approches de détection et de diagnostic. Ce processus assure ainsi une adaptation et une optimisation constantes, renforçant la capacité à réagir de manière précise et efficace face aux évolutions et aux changements des processus de production.

Selon la structure de l'approche méthodologique présentée dans la Figure 7, cette phase comprend une étape : Détection de dérive diagnostic et découverte de « Nouveautés » (P3.1.)

En complément, la Figure 14 représente un schéma illustratif basé SADT de la phase de vie P3.

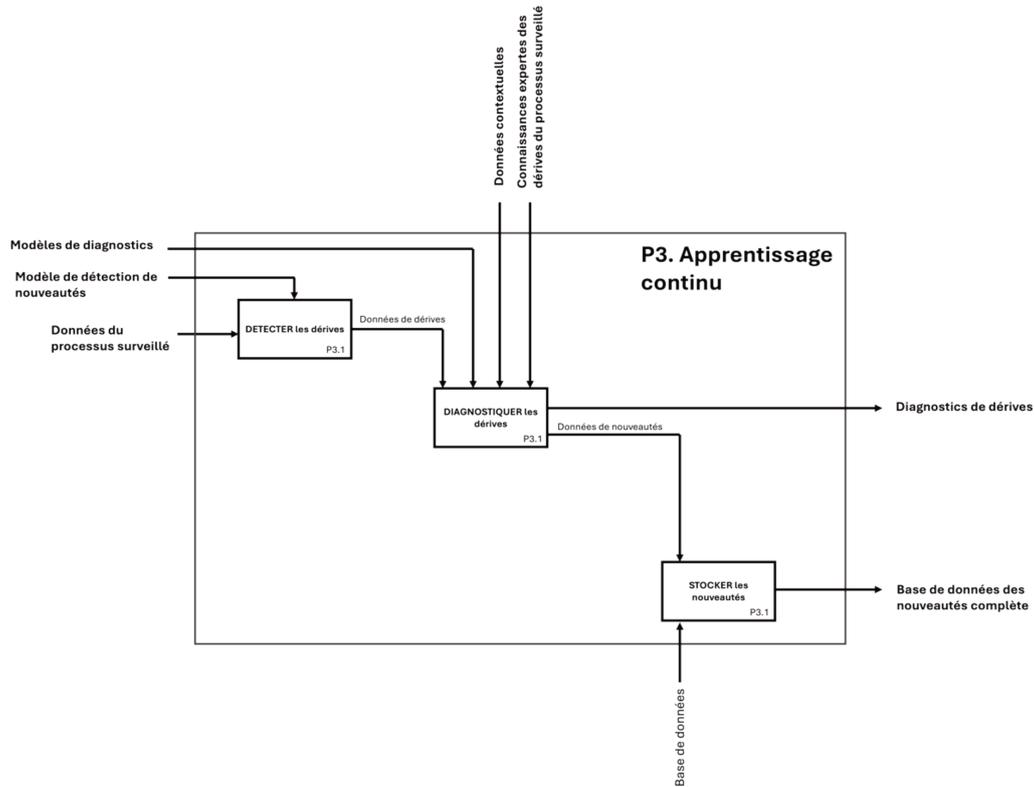


Figure 14. Illustration schématique de l'interaction entre les étapes internes de la phase P3 de l'approche méthodologique

2.4.6.1 P3.1 - Détection de dérive diagnostic et découverte de « Nouveautés »

L'étape P3.1, « Détection de dérive diagnostic et découverte de « Nouveautés » », exploite les modèles de détection de nouveautés et les ensembles de classification binaire entraînés aux phases précédentes pour effectuer l'activité de détection et de diagnostic de dérive. Lorsque le modèle de détection de nouveautés classe les nouvelles données de processus surveillées en temps réel comme étant des nouveautés, ces données sont transmises aux ensembles de classificateurs binaires pour vérifier leur correspondance avec les profils de dérive connus. Dans le cas où les données du processus sont reconnues par les ensembles de classificateurs binaires, un diagnostic est émis à destination des parties prenantes (les opérateurs machines, la maintenance, les managers, les systèmes de supervision...). Dans le cas contraire, cela signifie que les données sont inconnues et donc considérées comme une nouveauté. Ces données de nouveautés sont alors stockées dans la base de données des « nouveautés ». Lorsque la base de données des « nouveautés » accumule suffisamment de données (seuils définis manuellement par un expert ou selon un échéancier), le système déclenche un retour de l'approche méthodologique vers la phase P2 « Expertise des diagnostics ». Cette boucle d'amélioration continue de l'approche méthodologique implique l'exploitation de la base de données des nouveautés enrichie par de nouveaux profils de données par toutes les étapes de la phase 2 de l'approche méthodologique.

Il est important de relever que le rôle de l'approche méthodologique n'est pas de gérer en détail les connaissances expertes sur les causes des dérives ni de spécifier les actions de maintenance. Il se limite à identifier les dérives à l'aide d'un label, ainsi qu'à fournir une liste des éléments physiques associés au processus

concerné pour générer une alarme destinée à l'opérateur. Cette alarme se compose donc du label de la dérive et de la liste des éléments physiques impliqués comme présenter dans la Figure 15.

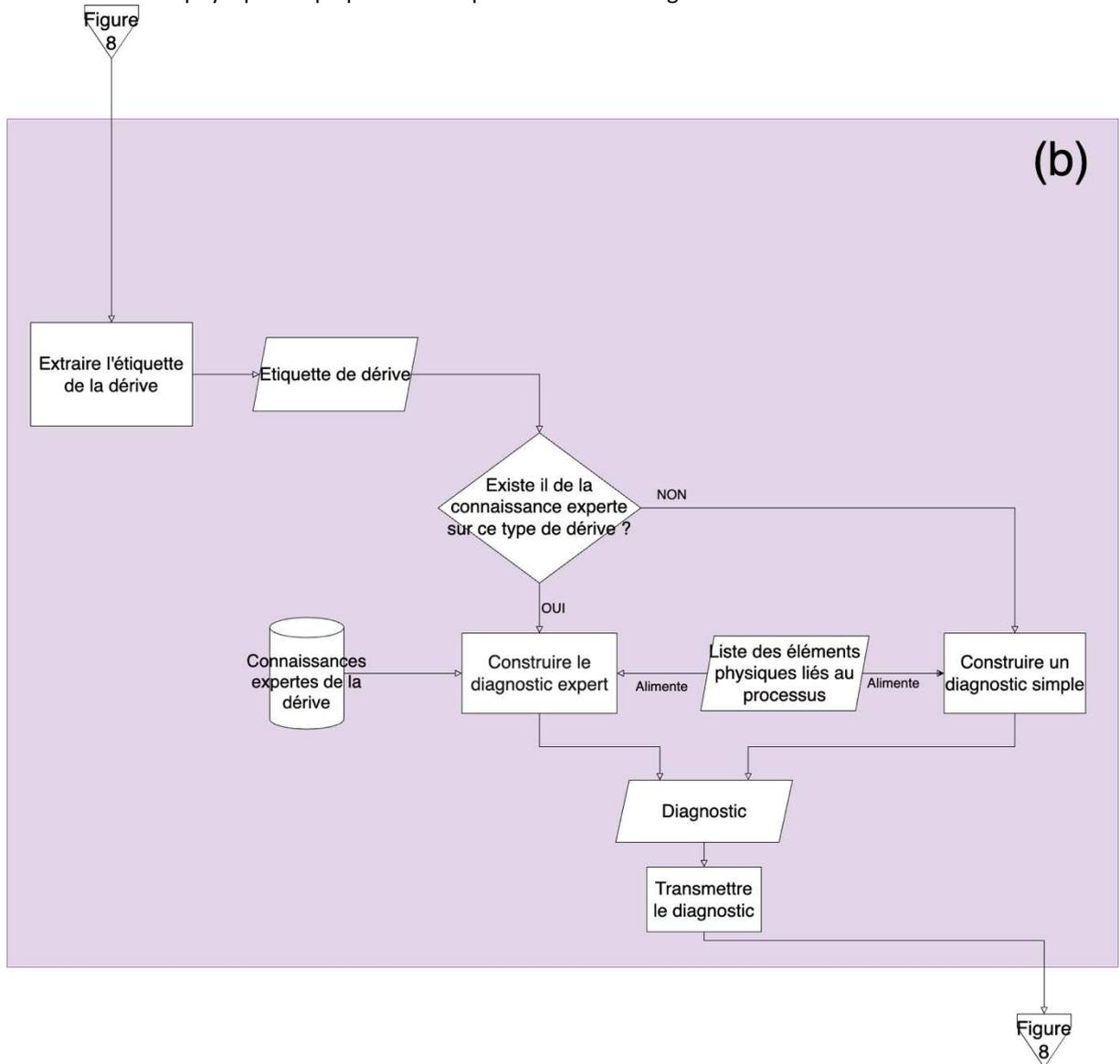


Figure 15. Représentation détaillée sous forme de diagramme de flux du bloc d'action de la Figure 8b, indiquant la gestion de la création du diagnostic dans l'approche méthodologique.

Le diagnostic de maintenance, quant à lui, doit être pris en charge par la GMAO, par exemple, qui, recevant le code de la dérive, utilise ses propres mécanismes internes, basés notamment sur les interventions passées, pour affiner le diagnostic dans le temps. Un système de retour d'expérience basé sur la pondération peut être mis en place pour améliorer le diagnostic de maintenance. Par exemple, à chaque intervention de maintenance, un retour d'information peut permettre de sélectionner les éléments physiques réellement impliqués dans la dérive et d'ajouter des remarques. Ce processus a pour objet, au fil du temps, de mettre en évidence les éléments les plus souvent concernés dans les interventions liées à un code de dérive spécifique.

2.4.7 Synthèse de la contribution scientifique à la problématique d'intégration continue des activités de détection et de diagnostic des dérives de processus de production hétérogènes

En conclusion, la troisième contribution scientifique de cette thèse sur le développement d'une approche méthodologique originale qui non seulement combine les approches de détection et de diagnostic présentées dans les sections précédentes, mais les intègre également dans un processus itératif et évolutif sous

la forme d'un enchaînement d'activités. L'approche méthodologique repose sur une architecture qui favorise l'interconnexion des systèmes de surveillance tout en prenant en compte les spécificités des environnements industriels variés, souvent caractérisés par leur non-stationnarité et leur complexité.

L'approche méthodologique proposée s'enrichit de manière continue grâce à une boucle d'apprentissage basée sur l'intégration de nouvelles données, ce qui améliore progressivement la précision des modèles de détection et de diagnostic. Ce processus permet de répondre aux défis liés à l'évolution constante des processus de production et à l'émergence de nouvelles formes de dérives, tout en minimisant les interventions humaines. En ce sens, c'est une solution évolutive capable de s'adapter aux différentes contraintes des environnements industriels hétérogènes, qu'il s'agisse de la variabilité des données, des exigences de passage à l'échelle, ou encore des besoins en personnalisation.

De plus, la méthodologie à travers cet enchaînement d'activités peut être considérée comme générique et transposable à différents contextes industriels.

2.4.7.1 Modularité de l'approche

En effet, la modularité est un des éléments fondamentaux de l'approche méthodologique proposée. La modularité se manifeste dans la structure même de l'approche méthodologique. En effet, les algorithmes de détection de dérive et d'identification des dérives utilisés dans les phases P1 et P3 peuvent être choisis librement sans remettre en cause l'infrastructure de l'approche méthodologique. Ainsi, libre à l'utilisateur de choisir les algorithmes pour traiter un type particulier de données d'entrée (séries temporelles multivariées, données textuelles, images, vidéo...). Cette capacité à intégrer de nouveaux composants sans perturber le fonctionnement global est essentielle dans des environnements où les technologies évoluent rapidement et où les exigences opérationnelles peuvent changer à court terme.

La flexibilité de l'approche méthodologique proposée réside dans sa capacité à s'adapter aux différentes architectures de systèmes de production. Que le système soit centralisé ou distribué, l'approche méthodologique » peut être configurée » pour fonctionner de manière efficace, en s'ajustant aux spécificités des données collectées et des méthodes de surveillance utilisées.

2.4.7.2 Indépendance vis-à-vis des connaissances préalables

Une des particularités de l'approche méthodologique proposée est sa capacité à fonctionner indépendamment des connaissances préalables sur les dérives. Cette caractéristique permet de déployer l'approche méthodologique dans des environnements industriels où les connaissances concernant les types de dérives sont abondantes, rares, mal documentées ou inexistantes. En effet, les systèmes de surveillance classique reposent sur des modèles prédéfinis de comportements conformes et non conformes, ce qui nécessite une connaissance approfondie et exhaustive des types de dérives possibles. Or, dans de nombreux secteurs industriels, comme celui dans lequel s'intègre ce travail de recherche, il est difficile de collecter ou de maintenir une telle base de connaissances à jour.

L'approche méthodologique contourne cette limitation en utilisant des techniques de détection de nouveauté, qui modélisent uniquement les comportements normaux des processus de production. En se basant sur des ensembles de données représentant des cycles de production conformes, les activités concernées de l'approche méthodologique sont capables de détecter des écarts significatifs sans avoir besoin de profils de dérive préalablement définis.

Cette indépendance vis-à-vis des connaissances préalables concernant les dérives rend l'approche méthodologique adaptée aux environnements industriels où les processus sont complexes, les dérives rares, et où il est difficile d'anticiper toutes les formes possibles de dérives. Cette indépendance permet également de réduire les coûts et le temps associés à la collecte et à l'étiquetage des données, tout en augmentant la réactivité et la précision du système de surveillance.

2.4.7.3 Scalabilité et gestion des données

La capacité à gérer des volumes massifs de données en temps réel est une exigence fondamentale pour tout système de surveillance industrielle moderne basée sur l'intelligence artificielle. Cette scalabilité est rendue possible par l'intégration de techniques avancées d'apprentissage automatique, comme le regroupement pour trier efficacement la masse de données amassée au fil du temps.

En outre, l'approche méthodologique est conçue pour s'adapter à l'évolution des volumes de données sans compromettre ses performances. Grâce à la modularité fournie par la boucle d'amélioration continue de

l'approche méthodologique (P3 qui déclenche un retour vers P2), les modèles de détection de dérive et de diagnostics peuvent être étendus ou ajustés pour gérer de nouveaux profils de dérive.

Cette capacité à évoluer avec les besoins de l'industrie garantit que l'approche méthodologique reste pertinente et efficace, même dans des environnements en rapide évolution technologique.

2.5 Conclusion

Dans ce chapitre trois contributions scientifiques dans le domaine de la surveillance et du diagnostic des dérives des processus de production hétérogènes sont présentées. Chacune de ces contributions se réfère à une des problématiques scientifiques identifiées dans le Chapitre 1.

Contribution 1 : Méthodologie de sélection d'algorithmes pour la détection de dérives dans les processus hétérogènes.

La première contribution porte sur l'adaptation des méthodes de détection des dérives aux spécificités des processus de production hétérogènes. Les méthodes traditionnelles de détection, qu'elles soient statistiques, physiques ou basées sur des seuils, se révèlent insuffisamment flexibles pour répondre aux exigences de ces environnements complexes. Pour résoudre cette problématique, une approche basée sur la détection de nouveauté, s'appuyant sur des techniques d'apprentissage automatique, telles que l'Isolation Forest et les Autoencodeurs, a été proposée. Cette solution permet d'identifier des dérives sans nécessiter de connaissance préalable exhaustive des modes de déviation, tout en s'adaptant aux particularités des processus hétérogènes à grande échelle.

En complément à cette orientation défendue, une méthodologie guidée pour le choix des algorithmes de détection de nouveauté en fonction des données d'entrée a été développée. Cette méthodologie repose sur une analyse des séries temporelles multivariées pour en extraire des indicateurs nécessaires pour à guider les décisions concernant la transformation des données et le choix des algorithmes de détection de nouveauté parmi ceux présentés dans le Tableau 3. Cette approche systématique illustrée par la Figure 5 permet d'optimiser les performances des algorithmes en fonction des spécificités des données, assurant ainsi une détection plus précise et adaptée aux environnements complexes.

Contribution 2 : Méthodologie combinant modèles basés sur les données et sur la connaissance pour le diagnostic des dérives et l'identification des causes racines.

La seconde contribution est centrée sur le diagnostic des dérives et leur association à des causes racines physiques. Le défi principal réside dans l'établissement d'un lien clair entre les dérives détectées et les composants physiques des systèmes industriels. Pour répondre à cette exigence, une approche hybride a été proposée, combinant les DDM et les KBM. En particulier, les DDM permettent d'identifier les dérives et les KBM de les relier à des causes racines physiques, permettant ainsi de mieux cibler les interventions de maintenance nécessaire.

Une méthodologie de sélection des algorithmes de classification binaire en fonction des données d'entrée a également été introduite. Elle guide le choix des algorithmes de classification binaire en tenant compte des dépendances temporelles, du volume des données et des relations linéaires ou non linéaires entre variables. Cette méthodologie, synthétisée par la Figure 6, permet de combiner des algorithmes complémentaires au sein d'une approche d'apprentissage ensembliste, maximisant ainsi la robustesse et l'efficacité de l'identification des dérives nécessaire au diagnostic. Le Tableau 5 qui présente la complémentarité des algorithmes est également une contribution aidant à identifier les combinaisons optimales pour renforcer les performances globales des modèles.

Contribution 3 : Approche méthodologique unifiée pour la détection et le diagnostic de dérives des processus de production hétérogènes.

La troisième contribution repose sur le développement d'une approche méthodologique pour l'intégration de la détection et du diagnostic des dérives. En ce sens, l'approche méthodologique unifie les activités de surveillance et de diagnostic dans un processus itératif, permettant un enrichissement continu des modèles grâce à l'intégration de nouvelles données et à une boucle d'apprentissage. L'approche méthodologique est conçue pour faire face à différents environnements industriels.

En résumé, ces contributions scientifiques offrent une réponse globale, à partir du questionnement de SEW USOCOME, à des défis industriels majeurs liés aux étapes de détection de dérive et de diagnostic de la maintenance prévisionnelle et du PHM. Elles apportent des solutions novatrices et évolutives pour la détection

et le diagnostic des dérives dans des processus de production hétérogènes, et posent les bases d'une gestion plus efficace des systèmes industriels dans des environnements toujours plus complexes et dynamiques.

En continuité, le prochain chapitre a pour objet de valider ces contributions en les appliquant à un cas d'usage réel sur un processus de taraudage d'une usine de SEW USOCOME. L'objectif est de démontrer l'applicabilité de ces contributions, mais aussi leur pertinence dans un contexte industriel concret.

Chapitre 3 :

Détection de dérive et diagnostic d'un processus de production hétérogène

3.1 Introduction

Le chapitre précédent a proposé des contributions aux problématiques scientifiques soulevées dans cette thèse. La première contribution concerne une méthodologie pour le choix d'algorithmes de détection de dérives adaptés aux processus de production hétérogènes. La deuxième contribution présente une méthodologie pour la sélection d'algorithmes de classification binaire et une solution de diagnostic multimodèle, permettant d'associer les dérives à des causes racines. Enfin, la troisième contribution introduit une approche méthodologique intégrée pour la détection et le diagnostic des dérives, permettant une interaction continue et complémentaire entre ces deux activités. Cette même approche méthodologique inclut une boucle d'apprentissage continue pour améliorer les performances globales d'un processus et de mieux répondre aux contraintes d'évolutivité spécifiques des environnements industriels complexes.

Le présent chapitre vise à tester l'applicabilité de ces contributions à travers un cas d'usage industriel proposé par la société SEW USOCOME. L'objectif est de démontrer, d'une part, l'application de ces contributions à un cas réel – un processus de production hétérogène, et d'autre part, leur pertinence/intérêt. En ce sens, le chapitre s'organise en deux sections principales.

La Section 3.2 présente le cas d'usage industriel en détaillant le processus de production étudié (taraudage des flasques moteurs), les spécificités techniques, ainsi que les défis rencontrés dans la surveillance de ce type de système.

La Section 3.3 vise à mettre en place une solution de détection de diagnostic des dérives du processus de taraudage des flasques moteurs sans disposer de données préalables sur les modes de dérives. Pour ce faire, en suivant l'approche méthodologique, un modèle de détection de nouveautés est entraîné à partir de données conformes uniquement. Ce modèle mis en exploitation doit permettre d'acquérir des données de nouveautés. À partir de ces données de nouveautés, une phase d'expertise permet ensuite d'isoler des profils de dérives du processus et d'entraîner des modèles de diagnostic basé sur les DD et la connaissance experte de la machine (KBM). Enfin, des boucles d'amélioration continue sont présentées pour améliorer la précision des modèles de détection de dérive et de diagnostic avec de nouvelles données et de découvrir de nouveaux modes de dérives.

Le chapitre se termine par une discussion sur les résultats obtenus, en mettant en évidence les perspectives d'amélioration et d'extension des approches proposées à d'autres contextes industriels en Section 0.

3.2 Présentation du cas d'usage industriel

Le cas d'usage industriel, fourni par SEW USOCOME, est un processus de taraudage automatisé. Ce cas d'usage sert de terrain d'expérimentation pour évaluer l'applicabilité, la faisabilité et l'efficacité opérationnelle de l'approche méthodologique proposée.

Le processus de taraudage se situe dans une cellule robotique automatisée, dont la mission est de tarauder les brides de moteur préperçées, de vérifier que la porosité des brides respecte les normes en vigueur, puis de nettoyer les brides afin d'éliminer tout résidu d'huile. Le processus de production utilise une machine à tarauder à multibroches, illustrée Figure 16. Cette machine, composée d'un arbre vertical et d'une pince de maintien, est conçue pour tarauder simultanément quatre trous préperçés sur une bride de moteur.

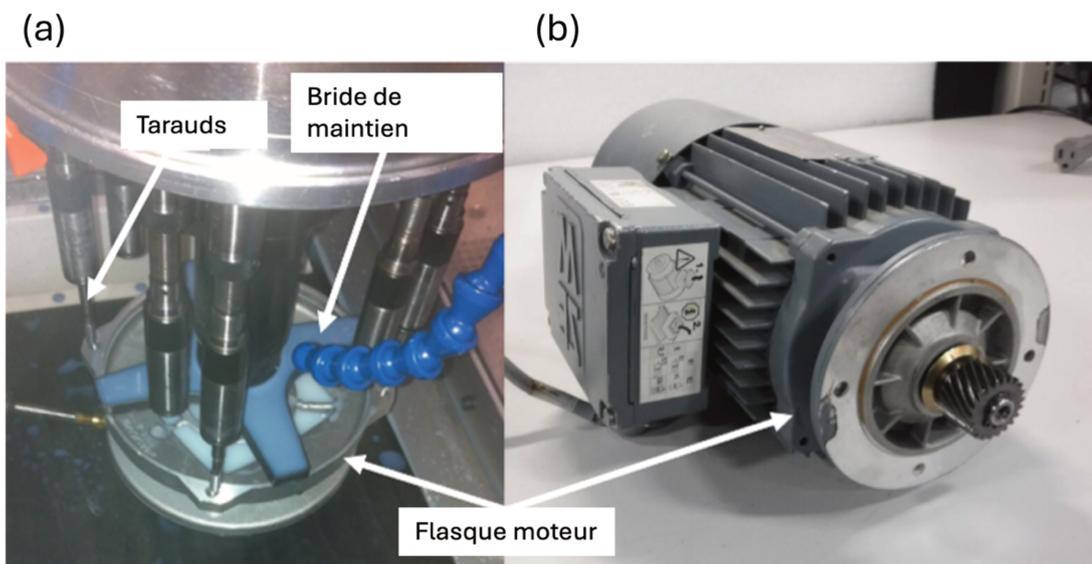


Figure 16. Taraudeuse multibroches et positionnement du flasque moteur sur le produit fini ; (a) taraudeuse multibroches, (b) moteur électrique sans réducteur.

3.2.1 Description du processus de taraudage

Le cycle de fonctionnement de cette machine se décompose en cinq phases successives :

1. La butée de maintien descend pour brider le flasque moteur, et les tarauds commencent à tourner pour atteindre leurs vitesses de consigne.
2. La broche multiple descend sur son axe vertical depuis la position haute et effectue le taraudage des trous.
3. La broche multiple s'arrête en position basse.
4. Le sens de rotation des tarauds s'inverse et la broche revient à sa position initiale haute.
5. La butée de maintien remonte pour libérer le flasque moteur.

Conformément à ce qui a été défini dans les chapitres précédents, la surveillance dans ce cas d'usage se concentre spécifiquement sur la dérive du processus de taraudage, c'est-à-dire sur sa capacité à mener à bien l'usinage de manière conforme, plutôt que sur la dégradation des actifs physiques directement liés au processus. Par exemple, l'action de verrouillage de la pièce via la pince de maintien n'est pas surveillée, car la mécanique et l'automatisation du système garantissent que l'usinage ne démarre que si la pression de serrage est correcte. Par conséquent, ces éléments n'ont aucun impact sur la surveillance de la dégradation du processus lui-même et ne sont pas pertinents pour l'évaluation de ses performances.

Ce processus de production est considéré comme hétérogène selon la définition donnée en Chapitre 1, car il implique différents types de données et des actifs physiques variés, tel que des variateurs de vitesse, un moteur électrique, une colonne de taraudage mécanique, un porte-taraud, les tarauds eux-mêmes, une bride de maintien, la pièce usinée et la bride du moteur. Cette hétérogénéité rend la surveillance du processus complexe et renforce l'importance d'une approche méthodologique bien structurée pour détecter et diagnostiquer les dérives.

L'enjeu industriel majeur de ce cas d'usage réside dans l'optimisation du processus de production des flasques moteurs. En effet, un des problèmes actuels de ce processus concerne la détection tardive des pièces non conformes, principalement due à l'usure excessive des tarauds. L'usure des tarauds n'étant pas visible à l'œil nu, elle entraîne la production continue de pièces non conformes jusqu'à ce que le problème soit détecté par un contrôle qualité fréquentiel. Chaque année, environ 1 500 pièces non conformes sont produites, nécessitant leur recyclage, ce qui représente environ 1,5 jour de production et des coûts supplémentaires pour la fonderie.

L'objectif principal de l'application de l'approche méthodologique proposée est d'anticiper et d'éviter ces défauts d'usinage en détectant les profils de dérive du processus, notamment ceux liés à l'usure des tarauds, avant qu'ils n'atteignent un seuil critique de non-conformité. Certaines dérives sont connues, telles que la casse d'un taraud ou le desserrage du porte-taraud, en revanche, aucune donnée représentative de ces dérives n'a encore été collectée. Pour mettre en œuvre l'approche méthodologique, il est essentiel d'examiner en détail les données disponibles sur le processus de taraudage.

3.2.2 Données du cas d'usage

Concernant le choix des données à surveiller, les variables à surveiller ont été définies par un expert en usinage. Ainsi, il a été décidé de surveiller les tensions entre phases et les courants de phase en amont du variateur de vitesse qui contrôle les quatre tarauds. Ces mesures de courant et de tension reflètent les efforts mécaniques que subissent les tarauds, ce qui permet de surveiller les changements dans l'état du processus. En suivant ces indicateurs, on peut détecter les signes de dysfonctionnement ou d'usure assez tôt pour intervenir avant que les problèmes liés à la qualité du produit ne s'accumulent.

Ces données sont acquises via des capteurs analogiques de courant et de tension placés en amont du variateur de vitesse. Ces capteurs sont connectés à une carte d'acquisition analogique intégrée à un automate programmable industriel (API). Le variateur de vitesse fonctionne à une fréquence d'entrée de 50 Hz. Surveiller le variateur nécessite donc une fréquence d'échantillonnage minimale de 100 Hz pour une acquisition adéquate des signaux (d'après le théorème de Shannon). Toutefois, une fréquence d'échantillonnage de 200 Hz a été choisie, représentant la limite maximale de l'API afin de garantir une meilleure précision des données recueillies. Ces données sont stockées localement sous forme de fichiers CSV sur l'API via un enregistreur de données (data logger). Les fichiers CSV contiennent sept colonnes, organisées comme présenté dans le Tableau 6.

Tableau 6. Description de l'architecture du fichier .csv de collecte des données

Colonne du fichier CSV	Type de donnée	Description
Horodatage	DATE_TIME	Date et heure de l'acquisition – format '%Y-%m-%d %H:%M:%S.%f'
CURRENT_L1	Float	Valeur du courant de phase 1 – unité : mA
CURRENT_L2	Float	Valeur du courant de phase 2 – unité : mA
CURRENT_L3	Float	Valeur du courant de phase 3 – unité : mA
VOLTAGE_L12	Float	Tension entre les phases 1 et 2 – unité : V
VOLTAGE_L23	Float	Tension entre les phases 2 et 3 – unité : V
VOLTAGE_L31	Float	Tension entre les phases 3 et 1 – unité : V

L'architecture de l'enregistreur de données fonctionne en deux modes distincts : un mode d'acquisition continue et un mode d'acquisition déclenchée. Le mode d'acquisition continue enregistre les données listées dans le tableau ci-dessus de façon continue et génère toutes les heures le fichier CSV associé. Ce mode de collecte des données produit des fichiers volumineux, chaque fichier contenant environ 4 320 000 points de mesure (6 variables mesurées 200 fois par seconde pendant 3600 secondes), pour une taille de 33 000 KB par fichier. Ces fichiers sont exportés vers un serveur FTP toutes les 10 heures pour éviter la saturation de la mémoire de l'API. Ce mode permet aux experts d'explorer les données hors ligne et de corrélérer les mesures avec d'autres systèmes comme le MES ou le GMAO afin de mener des analyses.

En parallèle, le mode d'acquisition déclenchée est activé au début de chaque cycle de taraudage. Lorsque le processus de taraudage démarre, l'enregistreur de données collecte les valeurs des capteurs durant toute la durée du cycle et le clôture à la fin du cycle de taraudage. Ainsi, un fichier CSV est généré pour chaque cycle de taraudage individuel. Chaque fichier contient entre 5 520 et 5 760 points de mesure par cycle (durée variable du cycle de taraudage), pour une taille approximative de 44 KB par fichier. Ces fichiers sont également transférés sur un serveur FTP une fois le cycle terminé.

Comme précisé précédemment, le premier mode d'acquisition sert à faire des analyses ponctuelles. Le second mode d'acquisition garantit une collecte de données à la bonne granularité pour assurer le bon fonctionnement à toutes étapes de l'approche méthodologique, de la création des modèles à la détection et le diagnostic de dérive en temps réel. La Figure 17 montre une représentation graphique des deux types de fichiers CSV générés : (a) les données historiques sur une heure de fonctionnement, et (b) les données issues d'un cycle de taraudage en comportement normal.

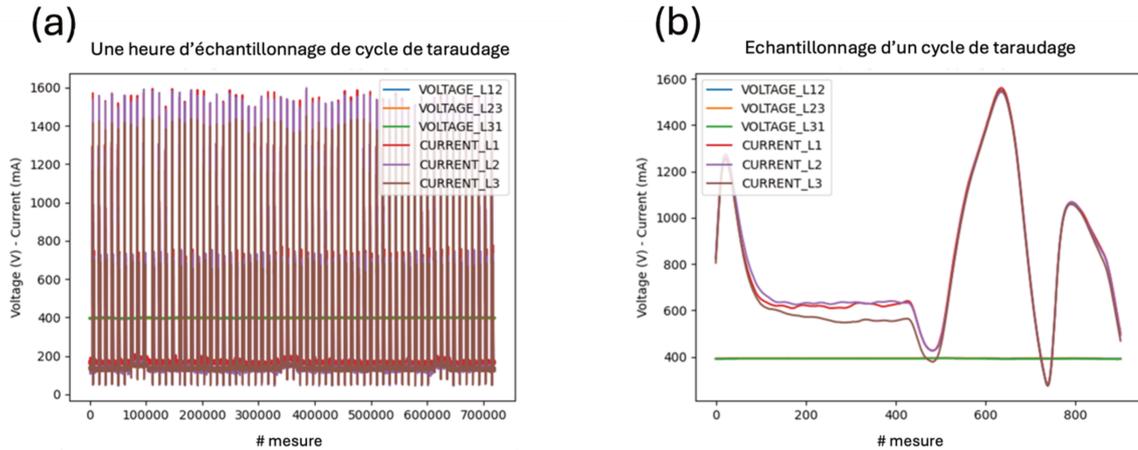


Figure 17. Représentation graphique des historiques CSV ; (a) représente un fichier historique d'une heure d'acquisition de données, (b) représente les données acquises durant un cycle de taraudage.

Enfin, la gestion des données dans ce cas d'usage est soumise à des politiques strictes de confidentialité définies par l'industriel partenaire. Toutes les données sont collectées, traitées et stockées sur le réseau interne de l'entreprise pour garantir leur sécurité et éviter tout accès non autorisé (cohérence avec les contraintes industrielles fixées au chapitre 1). Bien que les données ne soient pas disponibles publiquement en raison de leur caractère sensible, des techniques d'anonymisation sont à l'étude pour permettre un partage ultérieur sans compromettre les informations critiques du processus.

3.2.3 Outils de développement

Pour ce cas d'usage, tous les développements informatiques ont été réalisés en Python. Ce langage a été choisi pour sa simplicité de déploiement et sa flexibilité lors des tests et des mises à jour. Python est également reconnu pour son large éventail de bibliothèques dédiées à l'apprentissage automatique et à l'apprentissage profond. Tous les scripts nécessaires au cas d'usage décrit dans ce chapitre ont été développés avec la version Python 3.8. La génération des graphiques s'appuie sur la bibliothèque Matplotlib¹. La gestion des données est effectuée avec Pandas² et NumPy³. Pour les algorithmes d'apprentissage automatique, Scikit-learn⁴ a été utilisé, tandis que les réseaux de neurones ont été gérés avec Keras⁵ et PyTorch⁶. Le développement a été réalisé dans l'IDE PyCharm⁷, et le versionning a été assuré localement avec Git. Enfin, l'application finale est déployée sous forme de fichier exécutable (.exe) créé à partir des scripts Python grâce à la bibliothèque py2exe⁸. Ce fichier est exécuté sur un ordinateur industriel au pied de la machine et envoie les diagnostics de dérives l'API gérant la cellule de production via une requête Snap7⁹ afin que l'opérateur soit prévenu par une alarme via l'écran d'interface homme machine de la cellule de production.

3.3 Application de l'approche méthodologique proposée au cas d'usage industriel

Le cas industriel ayant été détaillé, cette section vise à appliquer l'approche méthodologique définie au chapitre 2. Cette approche méthodologique repose sur trois phases principales - P1 : Apprentissage de « nouveautés », P2 : Expertise des diagnostics et P3 : Apprentissage continu. Ces trois phases structurent

¹ <https://matplotlib.org/>

² <https://pandas.pydata.org/>

³ <https://numpy.org/>

⁴ <https://scikit-learn.org/>

⁵ <https://keras.io/>

⁶ <https://pytorch.org/>

⁷ <https://www.jetbrains.com/pycharm/>

⁸ <https://www.py2exe.org/>

⁹ <https://pympi.org/project/python-snap7/>

l'organisation de cette section et guident l'application concrète de la méthodologie au cas d'usage spécifique de SEW USOCOME.

3.3.1 P1 : Apprentissage de « nouveautés »

Dans cette sous-section, la phase P1 de l'approche méthodologique est appliquée au cas d'usage de SEW USOCOME. L'objectif de cette phase est d'exploiter le logigramme proposé dans la contribution scientifique 1 pour choisir un algorithme de détection de nouveautés adapté aux données récoltées. Une fois l'algorithme défini et les données collectées transformés si nécessaire (en suivant les recommandations du logigramme), le modèle de détection de nouveautés est entraîné avec des données conformes uniquement. Une fois entraîné, le modèle est mis en exploitation sur l'ordinateur industriel relié à l'API qui lui transmet les données représentatives des cycles de taraudage en temps réel. Cela permet une surveillance continue des dérives du processus vis-à-vis des données conformes ayant servi à l'entraînement du modèle. Les dérives détectées sont considérées comme des « nouveautés » et non des dérives, car elles peuvent comprendre de nouveau profil de données conformes également. Ces données de nouveautés sont ensuite stockées dans une base de données dédiée pour analyse ultérieure.

3.3.1.1 P1.1 Détection de nouveautés

3.3.1.1.1 Exploitation du logigramme de décision pour le traitement des données et le choix d'algorithme de détection de nouveautés associé.

La première étape est de calculer l'ACF des données. Pour ce faire les données de 100 cycles de taraudage ont été utilisées. L'ACF moyen des 100 profils de donnée a été utilisé comme indicateur. L'ACF permet de quantifier le degré de corrélation entre la série observée à différents décalages temporels ("lags"), pour identifier les éventuels cycles ou relations répétitives dans les données.

La procédure suivie pour la réalisation de ce test est la suivante :

1. **Définition de la série principale** : Les variables représentant la consommation de courant ont été identifiées comme la série temporelle principale à analyser.
2. **Sélection des lags** : L'ACF a été calculé par un expert en science des données pour un ensemble de lags allant de 1 à 50. Représenter l'ACF à différentes valeurs de lags permet de tester la corrélation entre la série temporelle et ses propres valeurs à différents décalages temporels.
3. **Visualisation** : Les résultats de l'ACF sont visualisables sous forme de graphique en Figure 18 et Figure 19 généré par l'expert en science des données. Chaque barre représente l'ACF à un certain lag. Les barres dépassant les bornes de signification (zone colorée) indiquent une corrélation significative à ces lags.
4. **Interprétation** : Si des pics significatifs apparaissent dans l'ACF à certains lags, cela indique une autocorrélation à ces intervalles, suggérant que les valeurs passées influencent les valeurs futures à ces distances temporelles. Un schéma de décroissance progressive de l'ACF ou un pic à des lags spécifiques peut révéler une tendance ou une saisonnalité dans la série.

Ainsi, l'analyse des Figure 18 et Figure 19 montre que l'ACF est supérieur à 0.5 jusqu'à 47 lags. Le logigramme de choix d'algorithme de détection de nouveauté recommande donc de garder les données brutes et d'utiliser un modèle de type RNN ou LSTM, car il y a une forte dépendance temporelle pour chaque variable surveillée vis-à-vis d'elle-même (autocorrélation).

Lags 1 à 5

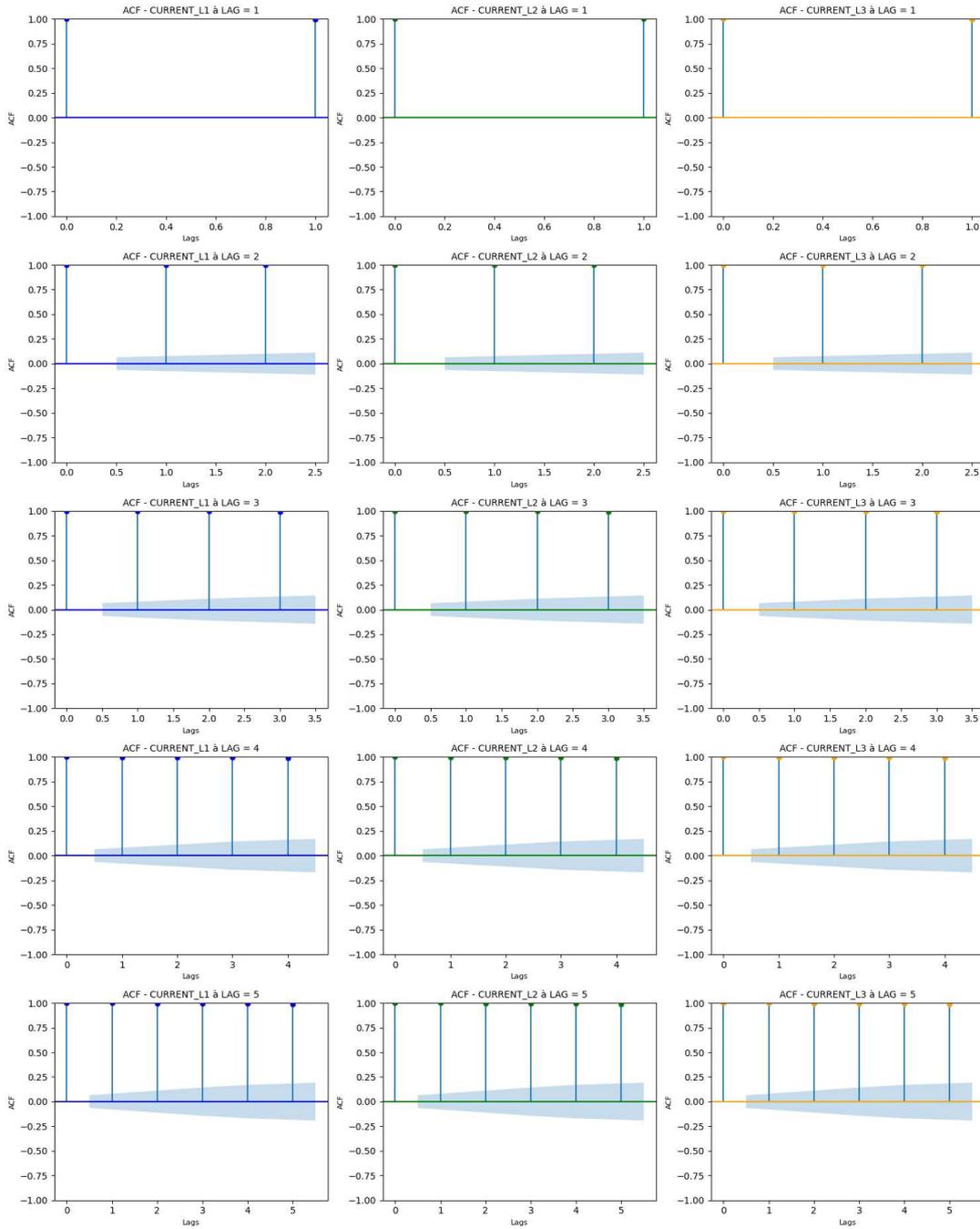


Figure 18. Représentation de l'ACF des courants L1 à L3 pour un lag allant de 1 à 5 par pas de 1.

Lags 10 à 50

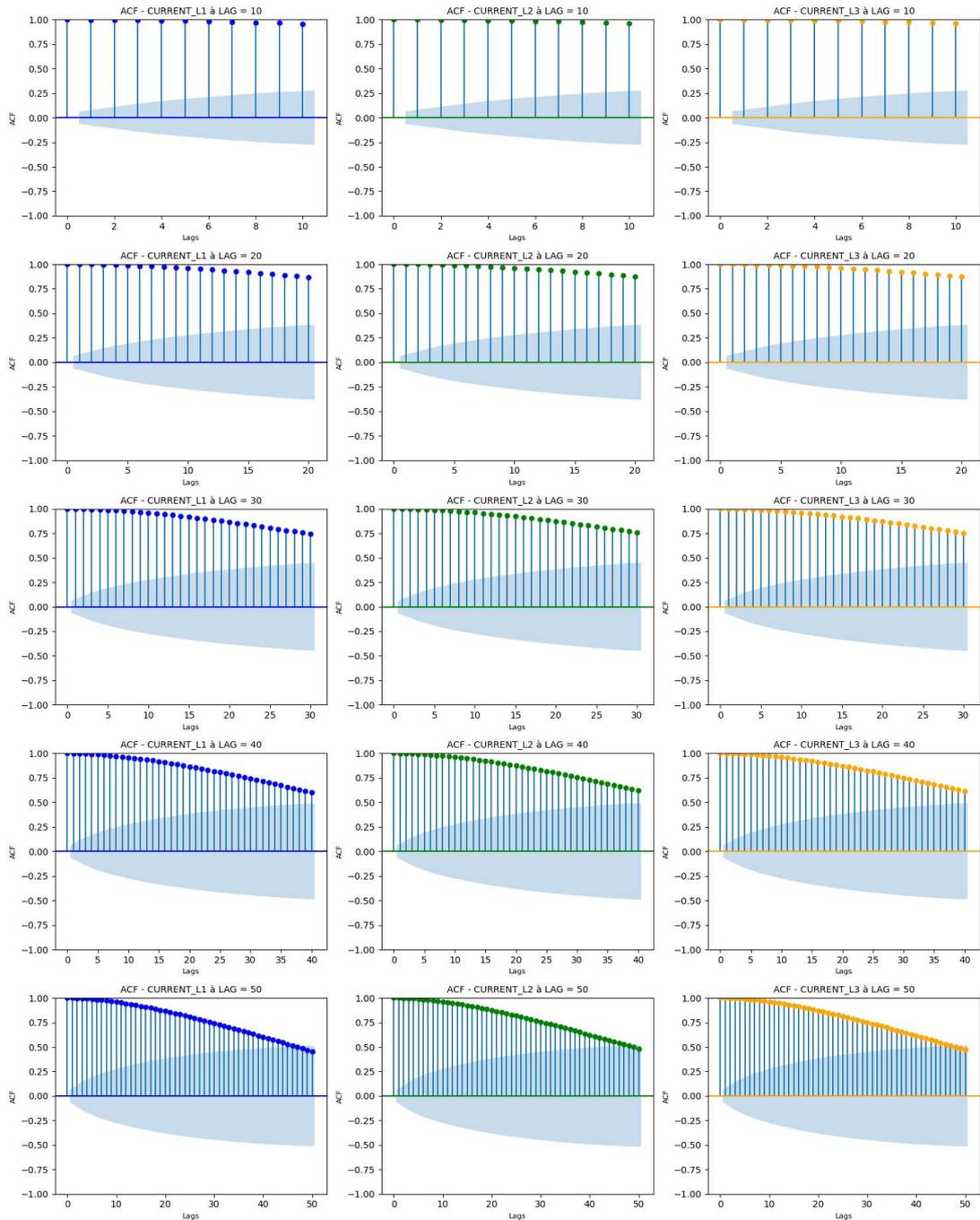


Figure 19. Représentation de l'ACF des courants L1 à L3 pour un lag allant de 10 à 50 par pas de 10.

3.3.1.1.2 Entraînement du modèle de détection de nouveautés.

Le choix s'est porté sur un LSTM, car il est mieux adapté aux séries temporelles avec des dépendances temporelles longues, comme celles du cas d'usage. En effet, les LSTM sont spécifiquement conçus pour capturer les dépendances à long terme grâce à un mécanisme de mémoire qui permet de retenir des informations sur de longues séquences temporelles. À l'inverse, les RNN souffrent du problème de disparition du gradient, ce qui ne permet pas de retenir des relations temporelles à long terme. Ainsi, pour les données du cas d'usage où la dépendance temporelle est étendue jusqu'à 47 lags, LSTM est le choix le plus adaptés pour entraîner un modèle de détection de nouveauté.

En ce sens, un modèle de détection de nouveauté basé sur les LSTM a été entraîné avec l'équivalent d'une journée de données conformes correspondant à 600 fichiers CSV selon une architecture de **classifieur à une seule classe** (*one-class classifier*). Cette approche permet au modèle d'apprendre uniquement à partir des données conformes, en identifiant les patterns normaux au sein de celles-ci. Une fois le modèle entraîné, toute observation qui dévie significativement de ces patterns est détectée comme étant une nouveauté.

Une fois entraîné et validé, le modèle de détection de nouveautés a été déployé sur un ordinateur industriel au pied de la machine. Ce même ordinateur industriel effectue la fonction de serveur FTP et récolte les fichiers CSV représentant les cycles de taraudage. Ces cycles de taraudage unitaire sont ensuite évalués par le modèle de détection de nouveautés. Lorsqu'un cycle est considéré comme une nouveauté, les données sont envoyées sur un serveur distant. La Figure 20 représente l'architecture d'exploitation de la solution de détection de nouveautés.

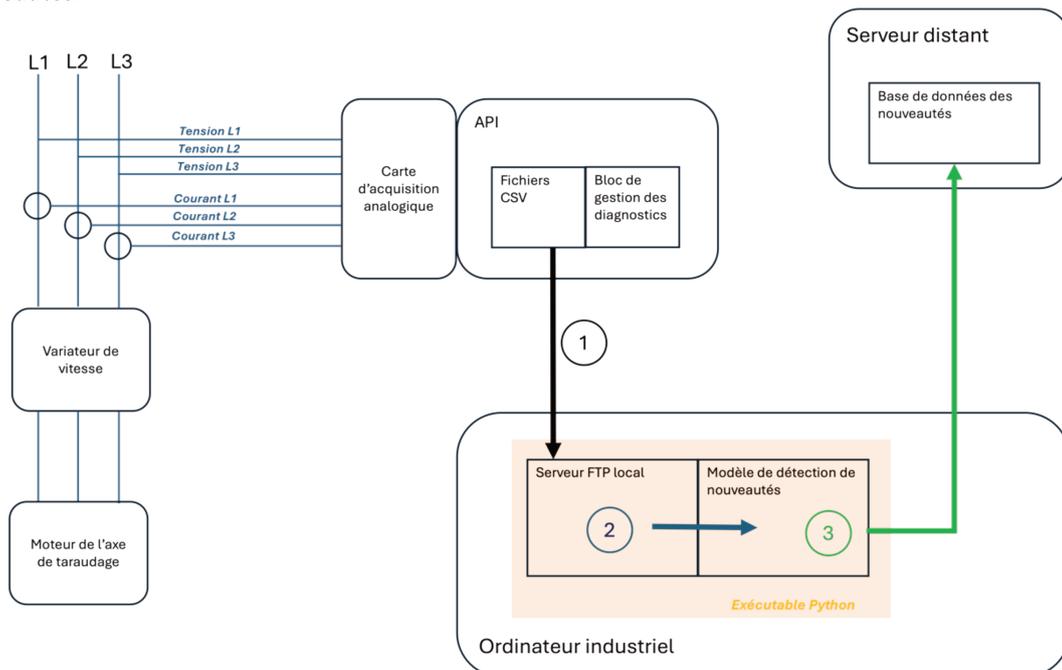


Figure 20. Représentation de l'architecture d'exploitation du modèle de détection de nouveautés.

Le flux d'information de la figure précédente suit la logique suivante :

1. L'API transfère en temps réel les fichiers CSV représentant les cycles de processus unitaire.
2. L'exécutable récupère les fichiers CSV dès réception et transfère les données au modèle de détection de nouveautés.
3. Si les données étiquetées sont étiquetées comme nouveautés par le modèle de détection de nouveautés, elles sont transmises au modèle de diagnostic

3.3.1.2 P1.2 Enrichissement de la base de données de « nouveautés »

Une fois le modèle de détection de nouveauté déployé, l'identification des déviations s'exécute de manière autonome pour collecter de nouveaux profils de processus déviant de la conformité établie par les données d'entraînement du modèle. Ces profils déviants sont, appelés « nouveautés ». Toutes les données considérées comme des nouveautés sont stockées dans une base de données de « nouveautés ».

Avec cette architecture déployée et fonctionnant de manière autonome pendant six jours, 3 000 fichiers CSV, parmi 3 600, ont été considérés comme étant des nouveautés. À ce stade, il a été jugé nécessaire de déclencher la phase P1.3 et de passer à l'étape P2 d'expertise des diagnostics.

3.3.2 P2 : Expertise des diagnostics

L'objectif de cette phase est d'analyser les « nouveautés » détectées lors de la phase P1 afin de diagnostiquer leur origine. Pour ce faire, il est nécessaire de recourir aux techniques de diagnostic proposées dans la contribution scientifique 2, en intégrant à la fois des modèles basés sur les données et des modèles basés sur les connaissances du domaine. Dans un premier temps les données de nouveautés sont analysées afin d'identifier des profils de données correspondant à des dérives ou des nouveaux comportements conformes. Ensuite des ensembles de classificateurs binaires sont entraînés pour chaque profil de donnée de dérive et un

autre est entraîné pour les profils conformes. Les étiquettes de ces ensembles de classificateurs sont utilisées comme base de diagnostics et recoupées avec les connaissances expertes pour proposer des diagnostics de dérive pertinents. Enfin, le modèle de détection de nouveautés est amélioré en le réentraînant avec les nouvelles données conformes afin d'améliorer ces performances.

3.3.2.1 P2.1 Regroupement et analyse experte de la base de données de « nouveautés »

L'objectif de cette phase est de regrouper les données de nouveautés par similarité dans des « clusters » de données afin de pouvoir étiqueter les données en phase P2.2. Pour ce faire, des algorithmes d'apprentissage non supervisé sont utilisés.

Le processus de regroupement des données ne s'appuie pas sur les données brutes, mais sur des caractéristiques statistiques dérivées des séries temporelles. Cela inclut la moyenne, l'écart-type, les valeurs minimales et maximales, ainsi que les percentiles (25 %, 50 %, et 75 %). Ces attributs ont été sélectionnés après plusieurs itérations, car ils se sont révélés être les plus discriminants pour ce contexte industriel spécifique et permettent d'isoler efficacement les cycles conformes, qui se distinguent nettement des dérives. Travailler avec ces caractéristiques statistiques simplifie la détermination des comportements atypiques, en offrant un gain de rapidité et d'efficacité, sans la complexité ni la charge de calcul des données brutes.

En revanche, pour la détection de dérive elle-même, les données brutes sont indispensables, car elles permettent de capter les changements subtils et progressifs dans le comportement au fil du temps, des variations qui peuvent être masquées par les statistiques globales. L'intervention de l'expert est également décisive à cette étape, car son expertise permet d'affiner le choix des caractéristiques pertinentes pour le clustering et d'interpréter les résultats des regroupements avec précision.

L'intervention de l'expert est également décisive à cette étape, car son expertise permet d'affiner le choix des caractéristiques pertinentes et d'interpréter les résultats des regroupements avec précision.

Les algorithmes utilisés pour cette tâche sont le DBSCAN et le K-means. La stratégie de regroupement des données appliquée se déroule en deux étapes. La première étape utilise l'algorithme DBSCAN pour effectuer un regroupement « grossier ». DBSCAN est choisi pour sa capacité à identifier des groupes de données disparates, ce qui permet de distinguer les déviations liées à des événements de dérives qui se situent très loin des données conformes du point de vue de la densité des données (voir Figure 22a). Concernant le paramétrage du DBSCAN, le choix de la mesure de distance est crucial pour la performance du regroupement. Plusieurs mesures ont été testées, comme la distance euclidienne et la distance de Manhattan. Finalement, la distance euclidienne a été retenue en raison de sa stabilité dans le contexte spécifique du cas d'usage.

La seconde étape utilise l'algorithme K-means pour diviser les données du profil « central » en plusieurs groupes distincts (voir Figure 23a). Cela permet d'évaluer plus précisément le niveau de dégradation des tarauds. DBSCAN, bien qu'efficace pour le regroupement initial, ne permet pas cette finesse de segmentation, comme l'ont montré certains tests.

Une fois les données de nouveautés regroupées en cluster, l'approche méthodologique peut être utilisée pour passer à l'étape suivante P2.2 afin qu'un expert étiquette les clusters de données.

3.3.2.2 P2.2 Expertise basée sur les connaissances des clusters de nouveautés

L'objectif de cette phase est d'étiqueter les nouveautés en les classant dans des catégories nominales, de dérive, ou de bruit. Pour ce faire, l'intervention d'un expert du processus est cruciale, car son expertise permet de guider les choix méthodologiques et d'interpréter les résultats issus des analyses. Le regroupement des nouveaux profils repose sur des méthodes non supervisées, qui regroupent les données selon leurs similarités.

Grâce aux clusters de données générés à l'étape précédente, et à l'historique de maintenance de la machine récupéré via le service de maintenance, l'expert a pu identifier trois profils de dérives supplémentaires, en plus du profil de données conformes. Ces profils correspondent à des cycles où les tarauds étaient trop usés pour faire de la qualité (étiquette : « tarauds très usés »), ou au moins un des tarauds était cassé (étiquette : « taraud cassé ») ou un défaut inconnu s'est manifesté (étiquette : « dérive inconnue »). Ces profils de dérive, ainsi que le profil nominal, sont représentés dans la Figure 21.

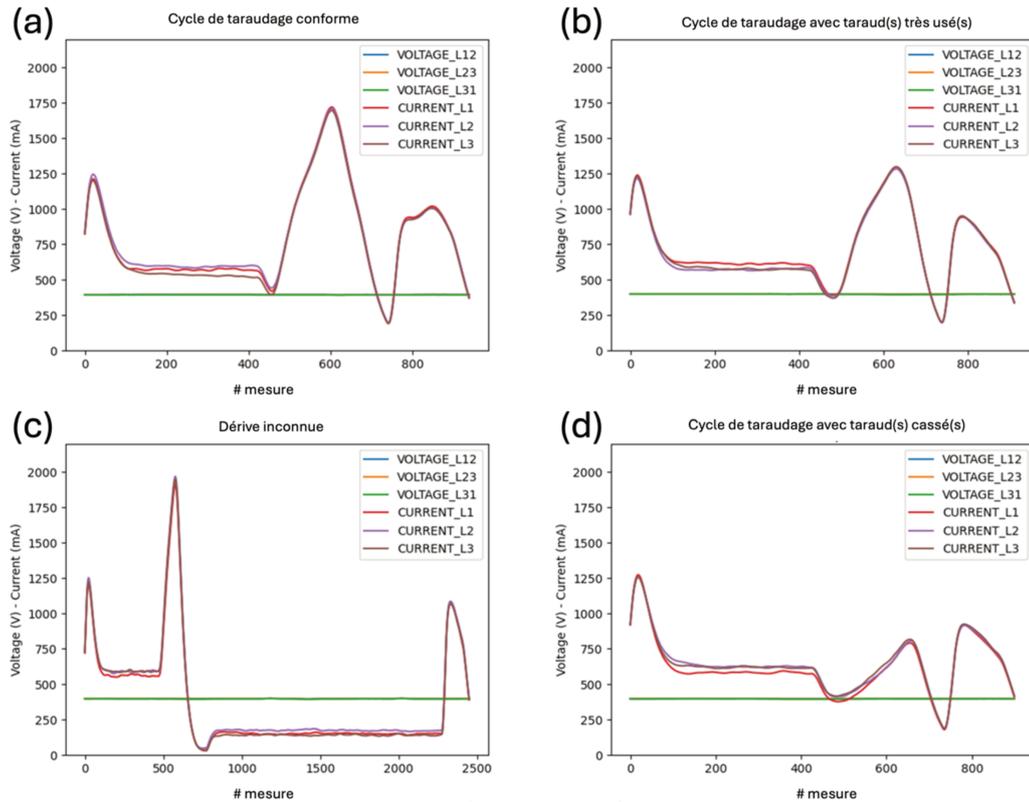


Figure 21. Présentation des données du comportement conforme et des profils de dérive identifiés ; (a) représentation des données du comportement conforme, (b) représentation des données de dérive représentant des taraud(s) très usé(s), (c) représentation des données de dérive inconnue, et (d) représentation des données de dérive de taraud(s) cassé(s).

La Figure 22b illustre l'étiquetage des clusters de données de dérives isolé via DBSCAN. On retrouve différents types de dérives étiquetées : celles liées aux dérives de taraud(s) très usée(s), de taraud(s) cassé(s), celles dues au bruit ou encore celles inconnues.

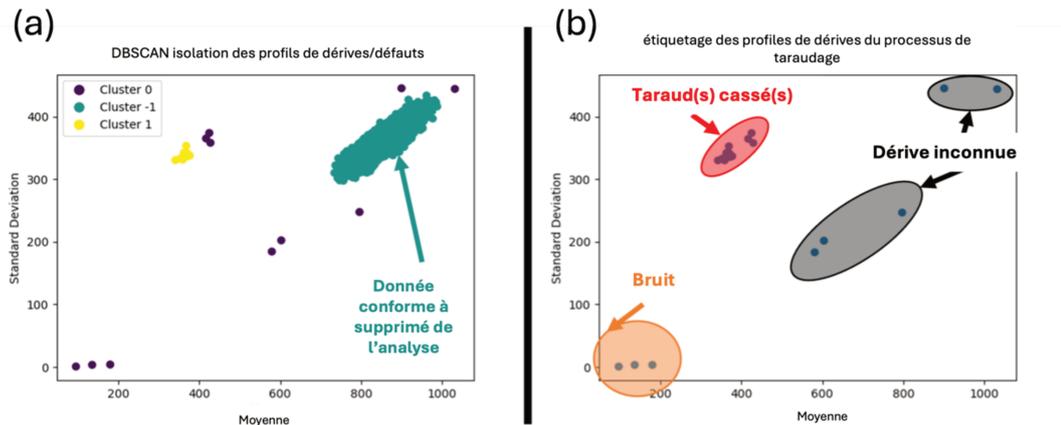


Figure 22. Processus d'isolation des dérives et des défauts ; (a) procédure d'isolation des dérives et des défauts avec DBSCAN, (b) étiquetage par l'expert des dérives des défauts associés aux clusters.

La Figure 23 (a) présente le regroupement effectué avec l'algorithme K-means sur les données de processus qui représente les données conformes et l'usure normale des tarauds. Deux variantes de K-means l'une basée sur l'algorithme de Lloyd et l'autre sur celui d'Elkan, ont été envisagées. L'algorithme de Lloyd a été retenu pour ses meilleures performances sur cet ensemble de données. La valeur initiale de k (nombre de clusters) a été déterminée à l'aide de la méthode dite « du coude », qui a fourni une estimation du nombre de clusters optimale. Cette estimation a ensuite été affinée avec le score de silhouette, permettant une meilleure

séparation des clusters. La Figure 23 (b), quant à elle, illustre comment l'expert a étiqueté chaque cluster à la suite de son analyse.

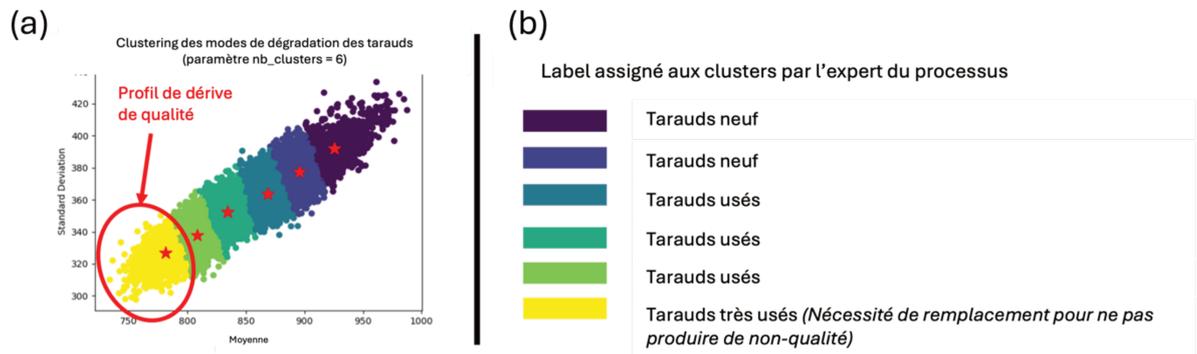


Figure 23. Procédure d'identification experte des dérives du processus de taraudage ; (a) création des clusters, (b) étiquettes associées.

Ce processus d'expertise des clusters de données a permis de distinguer plusieurs clusters de données considérées comme conformes (étiquettes : « tarauts neufs » et « tarauts usés ») et trois clusters de données considérés comme des dérives (étiquettes : « Tarauts très usés », « Tarauts cassés » et « Dérive inconnue »). Le cluster de données « bruits » a quant à lui été ignoré, car il contenait des données non représentatives. Ainsi, l'intervention de l'expert a été centrale pour interpréter et étiqueter ces résultats avec précision.

Concernant l'intégration de la connaissance via les KBM, ce processus n'impliquant à ce stade de connaissance qu'un seul composant, à savoir les tarauds, il a été jugé pour le moment inutile de faire un arbre de décision. De ce fait, les étiquettes retournées par les ensembles de classificateurs votant constituent l'unique source de diagnostic. Les arbres de décision seront intégrés a posteriori si nécessaire.

Une fois les clusters de nouveautés étiquetés, il est possible de passer à la l'étape P2.3 de l'approche méthodologique.

3.3.2.3 P2.3 (Ré-) Entraînement des modèles de diagnostics

À ce stade du cas d'usage, c'est la première fois que l'approche méthodologique entre dans cette étape. L'objectif ici est d'exploiter les clusters de données étiquetées par les experts et d'élaborer les modèles de classification binaire basée sur de l'apprentissage ensembliste supervisé pour chaque cluster de données (profil de dérives et profil conforme).

Pour ce faire le logigramme de décision défini pour le choix des algorithmes de classification binaire d'identification des dérives a été exploité. Ce logigramme se base sur la méthodologie guidée pour l'optimisation du traitement des données et la sélection des algorithmes de détection de nouveauté et l'étend avec le Tableau 5 afin de pouvoir construire un ensemble de classificateurs binaires adaptés aux données.

Ainsi, en se basant sur cette méthodologie étendue, l'algorithme principal recommandé pour créer les classificateurs binaires selon le logigramme proposé en Figure 6 est soit les RNN ou soit les LSTM. Pour les mêmes raisons que pour le choix de l'algorithme de détection de nouveautés, le choix en regard du cas d'usage s'est porté sur le LSTM.

Une fois l'algorithme de classification principal choisit, l'approche méthodologique impose de choisir des algorithmes complémentaires à partir du Tableau 5. Le tableau montre que le LSTM nécessite des algorithmes complémentaires pour : gérer les grands volumes de données, capturer les relations non linéaires, et traiter les données bruitées. En ce sens, le premier algorithme complémentaire choisi est RF. En effet d'après le tableau, le RF est complémentaire au LSTM sur les aspects de la gestion des grands volumes de données ainsi que pour la gestion du bruit. Ensuite, pour la capture des relations non linéaires, le SVM a été retenu comme algorithme complémentaire au LSTM pour la capture des relations non linéaires selon le tableau. Ainsi, l'utilisation conjointe du LSTM, du RF et du SVM permet de couvrir l'ensemble des besoins.

Concernant la méthode d'apprentissage des ensembles de classificateurs binaires, le bagging avec vote à la majorité a été retenu pour sa simplicité de mise en œuvre. Le choix peut être remis en cause si cette approche venait à ne pas être efficace en exploitation. Ainsi, à partir de ces choix, les ensembles de classificateurs binaires peuvent être entraînés pour chaque classe de données identifiée dans la phase P2.2.

En combinant ces différentes approches, les classificateurs peuvent équilibrer leurs faiblesses individuelles. Cet équilibre est atteint en combinant leurs prédictions dans un ensemble de votes, qui moyenne les probabilités prédites par chaque modèle pour améliorer la précision globale des prédictions. De plus, la

conception des modèles ensemblistes pour résoudre le problème d'identification des dérives ne concerne pas seulement l'intégration de différents modèles, mais également la répartition des poids entre eux. À ce stade du cas d'usage, chaque modèle de l'ensemble a le même poids en raison du manque de connaissance lié à l'efficacité des modèles en exploitation. La question de la répartition des poids est une perspective d'amélioration possible, lorsque plus d'expérience sera emmagasinée sur le cas d'usage.

De plus, évaluer la performance des modèles de classification binaire est crucial pour valider l'efficacité des modèles entraînés. À cet égard, les algorithmes sélectionnés ont été comparés individuellement et en collaboration dans un ensemble de classificateurs votant. Les performances de ces modèles ont été évaluées sur deux types de jeux de données d'entraînement (équilibrés et déséquilibrés) afin d'évaluer le comportement des modèles dans des scénarios réels. Les métriques utilisées pour évaluer les performances des cinq classificateurs sont présentées dans le Tableau 7.

Tableau 7. Métriques d'évaluation des performances.

Nom de la métrique	Description et objectif de la métrique
Justesse (Accuracy)	Mesure le pourcentage de prédictions correctes parmi toutes les prédictions du modèle ; donne un aperçu des performances globales du modèle.
Précision	Mesure la proportion de prédictions positives correctes parmi toutes les prédictions positives faites par le modèle ; utile pour minimiser les faux positifs.
Rappel (Recall)	Indique la proportion de véritables instances positives correctement identifiées par un modèle parmi toutes les instances positives réelles d'un jeu de données ; important pour minimiser les faux négatifs.
F1-score	Se réfèrent à la combinaison harmonique de la précision et du rappel ; utile pour évaluer l'équilibre entre ces deux aspects dans les prédictions.
AUC-ROC	Évalue la capacité du modèle à discriminer entre les classes ; plus l'AUC-ROC est proche de 1, meilleure est la capacité du modèle à distinguer entre les classes.
AUC-PR	Mesure la précision du modèle en fonction du rappel, notamment dans les situations où les classes sont déséquilibrées.
Perte (Loss)	Représente la valeur de la fonction de perte utilisée lors de l'entraînement ; indique comment le modèle ajuste ses poids pour minimiser cette fonction.
Temps d'entraînement	Indique le temps nécessaire pour entraîner le modèle sur un jeu de données spécifique.
Temps d'exécution	Donne une indication du temps nécessaire au modèle pour générer des prédictions sur de nouvelles données ; utile pour évaluer la vitesse opérationnelle du modèle.

D'un point de vue plus détaillé, la justesse mesure le pourcentage de prédictions correctes, tandis que le F1-score combine justesse et rappel pour évaluer l'équilibre entre les faux positifs et les faux négatifs. L'AUC-ROC évalue la capacité du modèle à discriminer entre les classes, tandis que l'AUC-PR offre un aperçu de la précision dans les situations où l'exactitude des prédictions positives est cruciale. La perte est utilisée pour comprendre l'apprentissage du modèle, tandis que les temps d'entraînement et d'exécution permettent d'évaluer l'efficacité opérationnelle. Toutes ces métriques permettent une évaluation approfondie des performances des classificateurs sur un jeu de données équilibré, en tenant compte des différents aspects de la qualité des prédictions et de l'efficacité du modèle.

En ce qui concerne le protocole de test, une stratégie de validation croisée à 5 blocs (k-fold cross-validation (Ait tchakoucht *et al.*, 2024)) été sélectionnée dans le but d'estimer la capacité du modèle à être généralisé ainsi que la cohérence des performances du modèle. Dans ce protocole, le jeu de données principal est divisé en cinq sous-ensembles principaux, en veillant à ce que les données présentes dans ces sous-ensembles

soient équilibrées selon les besoins du test. Ainsi, le jeu de données peut être équilibré (par exemple, 50 % de données nominales et 50 % de données de dérive) ou déséquilibré (par exemple, 70 % de données nominales et 30 % de données de dérive). Le modèle est ensuite testé cinq fois. À chaque cycle d'évaluation, quatre sous-ensembles sont utilisés pour l'entraînement et un pour le test. À chaque nouvelle itération, un sous-ensemble différent est utilisé pour le test, tandis que l'ancien sous-ensemble de test est réintégré dans l'entraînement. Cette rotation garantit que chaque échantillon est utilisé une fois pour le test et quatre fois pour l'entraînement. Les métriques présentées dans le tableau 1 sont mesurées à chaque itération. À la fin de chaque cycle d'évaluation, les métriques sont moyennées et la variance est calculée.

Le Tableau 8 présente les résultats de performance des classificateurs binaires entraînés avec un jeu de données équilibré composé de 500 profils nominaux et de 500 profils de dérive ("Tarauds très usés"). (ratio 50/50 entre le profil nominal et le profil de dérives)

Tableau 8. Évaluation des performances des classificateurs binaires avec un jeu de données équilibré.

Type de Classificateur	Justesse moyenne (Variance)	F1-score moyen (Variance)	AUC-ROC moyen (Variance)	AUC-PR moyen (Variance)	Perte moyenne (Variance)	Temps d'entraînement (ms)	Temps d'exécution (ms)
SVM	0.989 (0.002)	0.981 (0.0025)	1 (0.0)	1 (0.0)	0.3807 (0.005)	4	5
RF	1 (0.0)	1 (0.0)	1 (0.0)	1 (0.0)	0 (0.0)	71	3
LSTM	0.996 (0.001)	0.997 (0.001)	0.999 (0.0005)	0.999 (0.0005)	0.1269 (0.002)	3583	341
Ensemble de classificateurs votants	0.992 (0.0015)	0.989 (0.0018)	1 (0.0)	1 (0.0)	0.2538 (0.003)	3672	350

Plusieurs observations peuvent être faites à partir des résultats de comparaison des performances présentés dans le Tableau 8. Tous les modèles dérivés des algorithmes testés fournissent des résultats satisfaisants pour l'application au cas d'usage. De plus, la faible variance confirme les résultats en termes de cohérence et de généralisation du modèle. Le SVM montre des résultats très satisfaisants en termes de précision et de F1-score, dépassant 0,98 pour ces métriques. Le RF se distingue de tous les autres modèles avec des performances excellentes (justesse et F1-score à 1). De son côté, le LSTM montre des performances légèrement meilleures que SVM, mais a les moins bonnes performances en termes de temps d'entraînement et d'exécution. Enfin, l'ensemble de classificateurs basé sur le vote présente également de très bonnes performances avec une précision et un F1-score élevés, proches des performances du meilleur modèle individuel (RF). Ces résultats suggèrent que RF est la meilleure option pour ce cas d'usage, mais l'ensemble de classificateurs pourrait être une alternative intéressante dans des contextes où RF montrerait des performances moins élevées.

Les résultats du Tableau 8 représentent un cas où le jeu de données d'entraînement est équilibré, ce qui ne correspond pas à la réalité dans tous les scénarios. Dans un contexte où le déséquilibre des classes est important au sein du jeu de données d'entraînement, la gestion des faux positifs devient cruciale. Les faux positifs surviennent lorsque le modèle a mal classé une observation comme appartenant à une classe alors qu'en réalité, elle appartient à une autre. Dans un jeu de données déséquilibré, une classe est majoritaire par rapport à l'autre. Les modèles prédisent constamment la classe la plus dominante pour garantir un minimum d'erreurs globales, mais cela se fait au détriment des cas correctement identifiés de la classe minoritaire. Dans le cas du processus de taraudage, une grande quantité de données correspond au mode de fonctionnement normal du processus, tandis qu'une quantité limitée de données représente les dérives. En conséquence, apprendre avec un jeu de données déséquilibré, dans lequel les dérives sont sous-représentées, peut induire un risque d'erreur et les dérives du processus pourraient être masquées. Minimiser les faux positifs devient donc essentiel pour garantir que les prédictions positives faites par le modèle correspondent effectivement à la classe cible (par exemple, la dérive pour laquelle le classificateur a été entraîné). Ainsi, pour évaluer la performance du modèle sur des jeux de données déséquilibrés, la justesse a été remplacée par la précision et le rappel. La précision mesure la fiabilité des prédictions positives en minimisant les faux positifs importants pour la classe minoritaire ; à l'inverse, le rappel évalue la capacité à détecter autant de vrais positifs que possible en réduisant les faux négatifs critiques lorsque les observations de la classe minoritaire sont rares. Le Tableau 9 présente les résultats de performance

des classificateurs binaires entraînés selon le protocole de test défini précédemment et avec un jeu de données déséquilibré composé de 750 profils nominaux et de 250 profils de dérive (“Tarauds très usés”) (ratio 75/25 entre le profil nominal et le profil de dérive).

Tableau 9. Évaluation des performances des classificateurs binaires avec un jeu de données déséquilibré

Type de Classificateur	Précision moyenne (Variance)	Rappel moyen (Variance)	F1-score moyen (Variance)	AUC-ROC moyen (Variance)	AUC-PR moyen (Variance)	Perte moyenne (Variance)	Temps d'entraînement (ms)	Temps d'exécution (ms)
SVM	0.992 (0.002)	0.986 (0.002)	0.989 (0.002)	0.999 (0.0005)	0.999 (0.0005)	0.3807 (0.005)	7	6
Random Forest (RF)	1 (0.0)	1 (0.0)	1 (0.0)	1 (0.0)	1 (0.0)	0 (0.0)	118	5
LSTM	0.985 (0.0025)	0.983 (0.003)	0.981 (0.003)	0.998 (0.001)	0.998 (0.001)	0.0910 (0.002)	4740	328
Ensemble de classificateurs votants	0.999 (0.001)	0.992 (0.002)	0.997 (0.0015)	1 (0.0)	1 (0.0)	0.2255 (0.003)	4866	340

Plusieurs observations peuvent être faites à partir des résultats présentés dans le Tableau 9. Le RF montre à nouveau des résultats excellents pour toutes les mesures évaluées (précision, rappel, F1-score, AUC-ROC et AUC-PR), suggérant une grande capacité à classifier correctement les données sans erreur. Encore une fois, la faible variance confirme les résultats en termes d'adaptabilité aux données et de capacité à généraliser de manière optimale. L'ensemble de classificateurs votant présente également des performances proches de ce résultat global, avec des scores élevés pour toutes les métriques, bien qu'ils soient légèrement inférieurs à ceux du RF. Les modèles SVM et LSTM affichent également de bonnes performances, mais légèrement inférieures à celles du RF et de l'ensemble de votes. Cependant, le temps d'entraînement du RF est encore plus élevé que celui des modèles d'apprentissage automatique comme SVM, tandis que LSTM nécessite le temps d'entraînement le plus long parmi les quatre algorithmes individuels. En résumé, le RF se distingue dans cette comparaison en raison de ses performances excellentes pour toutes les mesures ; néanmoins, l'ensemble de votes reste une alternative efficace avec des performances très proches de celles du RF.

Dans l'ensemble, la comparaison entre les performances des modèles sur des ensembles de données équilibrés et déséquilibrés met en évidence des différences mineures dans le cadre de l'application du modèle à ce cas d'usage. Théoriquement, le choix entre des données équilibrées et déséquilibrées a un impact sur la performance du modèle. Les données équilibrées permettent une évaluation plus précise de la capacité prédictive à travers toutes les classes, réduisant le risque de biais en faveur de la classe majoritaire, ce qui est crucial dans des domaines sensibles, comme la détection de fraudes médicales ou de santé. Cependant, cela n'est pas vérifié dans ce cas d'usage, probablement en raison de la taille du jeu de données d'entraînement, qui reste relativement petite pour l'instant. De plus, la création de jeux de données d'entraînement équilibrés pour ce cas d'usage s'avère complexe, car l'occurrence de dérives et de défauts est plus faible à l'échelle du processus, dans lequel des milliers de cycles sont effectués chaque semaine, et un seul cycle est requis pour un cas d'usage donné. Le choix entre des ensembles de données équilibrés et déséquilibrés dépend donc des objectifs du problème et des coûts associés à la collecte ou au prétraitement. La prise en compte de ces facteurs est cruciale pour des performances de modèle fiables et appropriées.

L'**ensemble de classificateurs votant**, bien que légèrement moins performant que le **RF** avec un score excellent, reste pertinent. L'ensemble de classificateurs votant combine plusieurs modèles pour une performance robuste et équilibrée, atténuant les biais d'un seul modèle et offrant une meilleure généralisation face à des situations changeantes. En ce qui concerne le choix des algorithmes pour les classificateurs binaire, **RF**

seul peut être très satisfaisant dans ce cas d'usage particulier et avec cet ensemble de données, mais RF seul ne garantit pas une généralisation à des profils de donnée conformes et déviants encore inconnue. L'ensemble de classificateurs votant, en agrégeant plusieurs modèles, gère mieux la variabilité des données, offrant une stabilité sur divers ensembles de données ou contextes évolutifs. Maintenir l'ensemble de classificateurs votant constitue une approche adaptable face aux variations des données, soulignant ainsi l'importance de ne pas être enfermé dans un modèle particulier. Les approches ensemblistes, telles que l'ensemble de classificateurs votant, assurent une résilience et une généralisation robuste dans une variété de scénarios.

Ainsi, pour ce cas d'usage, un ensemble de classificateurs votant composé de modèles LSTM, SVM et RF a été entraîné à partir du jeu de données d'entraînement équilibré. Une fois le modèle entraîné l'approche méthodologique peut passer à sa phase P2.4.

3.3.2.4 P2.4 Amélioration du modèle de détection de dérive

L'objectif de cette étape est d'améliorer les performances du modèle de détection de nouveautés afin qu'il intègre les données de cycles étiquetés « conforme » à l'étape P2.2 et précédemment détectés comme nouveautés. En effet, le premier modèle de détection de nouveautés a été entraîné avec des données considérées comme conformes et récoltées sur un court laps de temps. De ce fait, le modèle actuel de détection de nouveautés considère des cycles conformes comme des nouveautés, les données ayant servi à son entraînement étant limitées.

Pour ce faire, les données ayant été identifiées comme conformes à l'étape P2.1 sont utilisées pour « augmenter » le dataset d'entraînement utilisé pour l'entraînement de l'algorithme de détection de nouveautés. Le dataset ainsi augmenté va servir à réentraîner le modèle de détection de nouveautés pour qu'ils soient plus efficaces dans son exploitation et reconnaissent moins de cycle conforme comme des nouveautés.

Une fois le modèle de détection de nouveautés ainsi mis à jour, cette étape peut prendre fin et il est possible de passer à sa phase de vie suivante P3 : Apprentissage continu de l'approche méthodologique.

3.3.3 P3 : Apprentissage continu

La phase P3 de l'approche méthodologique a pour objectif de mettre en exploitation le modèle de détection de nouveautés amélioré et les modèles de diagnostics développés en phase P2. Cette mise en exploitation suit l'architecture représentée en Figure 24.

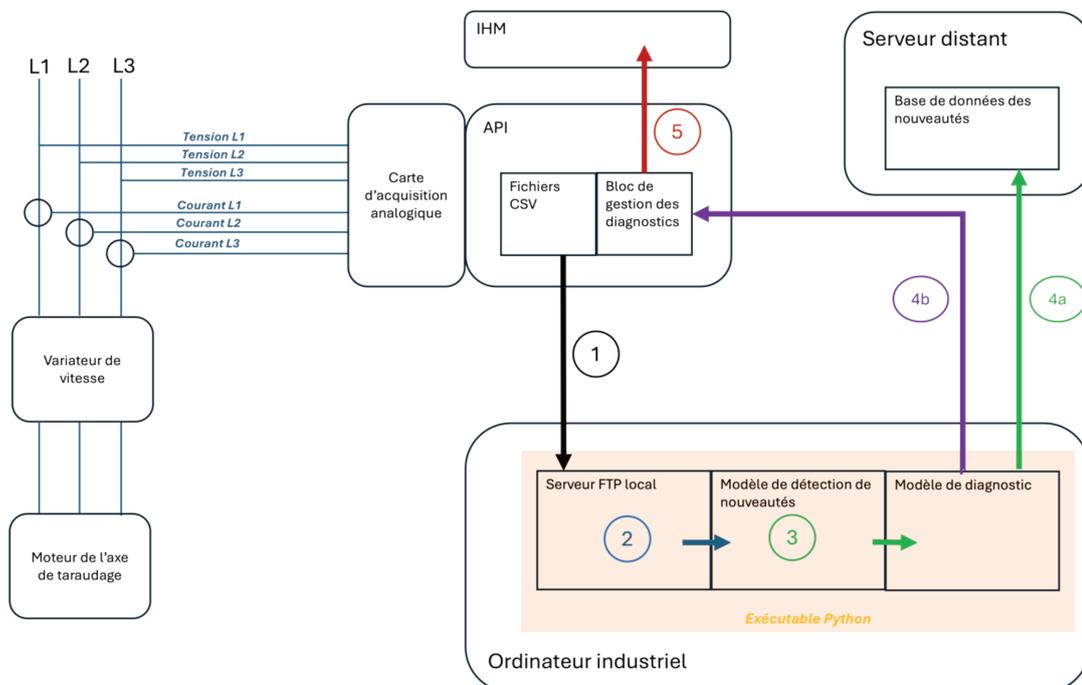


Figure 24. Représentation de l'architecture d'exploitation des modèles durant la phase 3 de l'approche méthodologique.

Concernant le flux d'information représenté dans la figure précédente, il suit la logique suivante :

1. L'API transfère en temps réel les fichiers CSV représentant les cycles de processus unitaire.
2. L'exécutable récupère les fichiers CSV dès réception et transfère les données au modèle de détection de nouveautés.
3. Si les données étiquetées sont étiquetées comme nouveautés par le modèle de détection de nouveautés, elles sont transmises au modèle de diagnostic
- 4a. Si les données de nouveautés ne sont pas reconnues par les modèles de diagnostic, alors elles sont envoyées dans la base de données des nouveautés sur un serveur distant au format CSV
- 4b. En revanche, si les données sont reconnues par un des modèles de diagnostic, alors le diagnostic est construit et envoyé à l'API via protocole Snap7
- 5 L'API peut alors mettre à disposition le diagnostic aux opérateurs en générant un message sur les IHM.

Dans cette architecture le modèle de détection de nouveautés sert à identifier en temps réel les dérives par rapport au comportement conforme. Lorsque des données déviantes sont détectées, elles sont ensuite soumises aux ensembles de classificateurs binaires. Si l'un des classificateurs reconnaît ces données, un diagnostic est automatiquement généré et envoyé à l'API qui transmettra un message aux opérateurs en conséquence. Dans le cas où aucun classificateur ne reconnaît les données, cela signifie qu'elles correspondent à des nouveautés. Ces données sont alors stockées dans une base de données dédiée aux nouveautés. Ce stockage permet, à terme, de repasser ces données dans la phase P2, lorsque suffisamment d'exemples sont accumulés, afin d'améliorer les modèles existants et potentiellement d'ajouter de nouveaux profils de diagnostic correspondant à des dérives inédites.

3.3.3.1 P3.1 Détection de dérive diagnostic et découverte de « Nouveautés »

L'objectif de cette étape est de déployer les activités de détection et de diagnostics des dérives de processus de production et de découvrir de nouveaux profils de données (nouveautés). Pour ce faire, le modèle de détection de nouveautés mis à jour et les modèles de diagnostics développés en phase P2 sont déployés en exploitation via la mise à disposition des modèles entraînés au format joblib sur le PC industriel (utilisation des exécutables). Le processus de fonctionnement suit la même logique qu'en étape P1.1. À chaque cycle de taraudage, l'API génère un fichier CSV contenant les données du cycle. Ce fichier est immédiatement envoyé au PC industriel via son serveur FTP intégré. Le script Python convertit en exécutable va pouvoir exploiter les modèles de détection de nouveautés et de diagnostics mis à disposition et traiter les données représentatives des derniers cycles de taraudage en temps réel.

Si le cycle n'est pas considéré comme une nouveauté par le modèle de détection de nouveauté, le traitement des données s'arrête là, et aucune autre action n'est entreprise. En revanche, si les données sont jugées comme une nouveauté par le modèle, elles sont alors présentées aux modèles de diagnostic. Si l'un des modèles de diagnostic reconnaît le profil de données, un diagnostic est généré et transmis à l'API gérant la cellule de production selon l'étiquette de dérive retenue. Si aucun modèle de diagnostic ne parvient à identifier les données, elles sont alors qualifiées de "nouveautés" et stockées dans la base de données dédiée sur un serveur distant de l'entreprise. La base de données de nouveautés servira à faire des analyses pour améliorer les modèles existants.

En ce sens, une fois la base de données ayant atteint un certain nombre de données ou selon un échéancier, l'approche méthodologique effectue une boucle d'apprentissage continue en revenant en phase P2.

Pour ce cas d'usage, une première boucle d'amélioration continue a été lancée après 12 jours en phase P3.2. Durant ces jours, 40 nouveautés ont été détectées, ne contenant pas de nouveau profil de données inédites pour le moment, mais présentant de nouvelles données pour la classe de défauts « tarauds cassés ».

Cette première phase d'apprentissage continu a permis de détecter l'usure des tarauds qui peut conduire prématurément à la non-conformité, évitant ainsi jusqu'à 60 pièces non conformes en l'espace de deux semaines. Cela a également eu un impact financier positif en réduisant les coûts de recyclage des pièces (les rebuts) et les coûts logistiques, tout en permettant de ne pas prendre de retard sur les commandes clients. À l'échelle d'un seul processus de taraudage, les résultats ne représentent pas encore un gain financier énorme. Cependant, à l'échelle des sites de production français, et sur une année de production, l'accumulation de ces petits gains pourrait potentiellement atteindre un montant de plusieurs centaines de milliers d'euros.

Le cadre méthodologique a été maintenu en phase P3.2 pendant deux mois avant de lancer une deuxième boucle d'apprentissage continu. Au cours de cette phase, un nouveau profil de taraudage conforme a été identifié. Ce nouveau profil correspondait à une altération du processus induite par l'usage d'un lot de production exceptionnel et non planifié dans lequel l'outil et la pièce usinée différaient de ceux habituellement

utilisés sur cette machine. Cette situation a permis de valider, directement sur le terrain, l'efficacité des algorithmes de détection des dérives ainsi que la robustesse de la boucle d'amélioration continue, en identifiant un nouveau type de profil conforme.

À ce stade l'approche méthodologique doit permettre de conduire des phases d'amélioration continue en alternant entre les phases P3 et P2 de manière périodique.

3.3.4 Résultats

Les résultats initiaux du déploiement de l'approche méthodologique sur le processus de taraudage sont très intéressants. Ils montrent un potentiel avéré pour la détection et le diagnostic des dérives dans un processus de production hétérogène. En effet, lors de la phase de détection de nouveautés, la méthodologie guidée pour l'optimisation du traitement des données et la sélection des algorithmes de détection de nouveauté a permis de caractériser les données du cas d'usage et de sélectionner l'algorithme le plus adapté pour entraîner un modèle de détection de nouveautés. Le modèle basé sur les LSTM a permis d'isoler 3 000 fichiers CSV comme étant des nouveautés sur une période de six jours.

Ces données ont permis dans la phase d'expertise des diagnostics d'identifier des profils de données correspondants à de nouveaux profils de données conformes ainsi qu'à des profils de données correspondants à des dérives telles que « taraud cassé », « dérive inconnue » ou encore « tarauds très usés ». À partir de ces données étiquetées, il a été possible d'entraîner des DDM capables d'identifier des dérives. Les DDM entraînés reposent sur une approche d'apprentissage ensembliste de classificateurs binaire prenant les décisions par vote pondéré par leurs degrés de certitude. Cette méthode permet de combiner plusieurs modèles de manière à renforcer la fiabilité et la précision des classifications. Elle atténue les faiblesses des modèles individuels, tout en capitalisant sur leurs forces respectives. Les algorithmes formant l'ensemble de classificateurs binaires ont été sélectionnés en exploitant la méthodologie guidée pour le choix des algorithmes de classification binaire d'identification des dérives. Ainsi des ensembles de classificateurs binaires composés de LSTM, SVM et RF ont été entraînés pour chaque profil de données disponible. De plus, le modèle de détection de nouveautés a été mis à jour avec les nouvelles données conformes afin d'améliorer ces performances en termes de détection de nouveauté. Une fois le modèle de détection de nouveautés amélioré et les DDM entraînés, ils ont été déployés en exploitation et la troisième phase d'apprentissage continu de l'approche méthodologique a été enclenchée.

Ainsi, durant les cycles d'amélioration continue réalisés par le rebouclage de la phase d'apprentissage continu vers la phase d'expertise des diagnostics, le modèle de détection de nouveautés et les DDM ont gagné en robustesse via l'intégration de nouvelles données. Également, de nouveaux profils de données conformes et de dérives ont pu être découverts.

En termes de performances industrielles, les résultats pratiques sont multiples. Tout d'abord, la détection de la casse des outils a atteint un taux de confiance suffisamment élevé pour éviter la production de séries de pièces non conformes dans 90% des cas de casse de tarauds. De plus, la détection des profils de dérive liés à l'usure avancée des tarauds permet d'anticiper les non-conformités, évitant ainsi jusqu'à 60 pièces non conformes en seulement deux semaines. Bien que l'impact réel de cette anticipation ne puisse pas encore être pleinement évalué, l'approche méthodologique pourrait à terme permettre de remplacer les interventions de maintenance systématique par une maintenance prévisionnelle basée sur l'évaluation continue de l'usure des tarauds.

Pour **SEW USOCOME**, cette anticipation des non-conformités peut augmenter la production de 1,5 %, soit environ 30 pièces supplémentaires par jour. Ces premiers résultats mettent en avant non seulement les bénéfices économiques et opérationnels, mais aussi la flexibilité de l'approche méthodologique, qui semble apte à être déployée dans d'autres applications industrielles.

Cependant, pour exploiter pleinement le potentiel de cette approche méthodologique, il est nécessaire d'effectuer des analyses approfondies de chaque processus pendant la phase initiale, ainsi qu'une collaboration étroite avec des experts du domaine pour affiner l'approche méthodologique. Des analyses préliminaires, comme l'analyse des modes de défaillance, de leurs effets et de leur criticité, ou encore l'analyse des dangers et de l'opérabilité, peuvent être menées afin de garantir que les données collectées sont représentatives du processus industriel.

Lors de la phase de diagnostic, l'implication des experts du domaine est cruciale. Ils permettent de labelliser correctement les clusters de données issues des nouveautés et d'établir des modèles de diagnostic représentatifs. Leur connaissance approfondie des processus garantit que les modèles développés sont non seulement précis techniquement, mais également pertinents sur le plan opérationnel. Cette expertise est essentielle pour s'assurer que l'approche méthodologique offre des informations exploitables et adaptées aux réalités industrielles.

En complément de l'expertise métier, des compétences en apprentissage automatique et en apprentissage profond sont nécessaires pour le choix, l'entraînement, le déploiement et la maintenance des algorithmes aux différentes étapes de l'apprentissage des nouveautés et du diagnostic. La collaboration entre les experts du domaine et les spécialistes en apprentissage automatique est donc essentielle pour ajuster les activités de l'approche méthodologique en fonction des spécificités du contexte industriel. Cela garantit que les modèles restent performants face à l'apparition de nouvelles données et de nouveaux profils de dérive.

La capacité de l'approche méthodologique à être généralisée réside dans son approche structurée et sa capacité d'adaptation. Son succès repose sur une méthodologie claire : une analyse détaillée des processus, l'implication des experts pour assurer un bon labelling des données et l'utilisation de techniques avancées d'apprentissage automatique. Bien que des ajustements spécifiques soient nécessaires pour chaque nouvelle application en raison des particularités opérationnelles et des modes de défaillance, les principes de base de l'approche méthodologique restent vrais.

3.4 Conclusion

Ce chapitre a démontré l'applicabilité et la pertinence des contributions scientifiques développées dans les chapitres précédents en les appliquant à un cas d'usage industriel réel fourni par SEW USOCOME, à savoir un processus de taraudage automatisé. L'application de l'approche méthodologique proposée a permis de valider sa faisabilité dans un environnement de production complexe et hétérogène, en s'adaptant aux spécificités du processus et aux défis liés à la variabilité des données et à la détection des dérives.

La phase de détection des dérives basée sur l'apprentissage de nouveautés a permis de capturer efficacement des anomalies dans les cycles de taraudage, notamment en identifiant des dérives telles que l'usure avancée des tarauds ou la casse de ces derniers. Cette étape, combinée à l'expertise humaine pour étiqueter les profils détectés, a permis de construire une base de connaissances solide pour améliorer le diagnostic des dérives. En outre, l'intégration des DDM via des **ensembles de classificateurs votant**, associant plusieurs modèles d'apprentissage automatique, a renforcé la capacité à identifier les dérives tout en atténuant les biais liés à l'utilisation d'un seul modèle.

L'enrichissement continu des modèles grâce à la phase d'apprentissage continu a également montré des résultats prometteurs, avec une réduction notable des pièces non conformes et une anticipation des dérives avant qu'elles n'entraînent des défauts majeurs dans le processus de production. L'application de l'approche méthodologique permet ainsi de réduire les pertes liées aux non-conformités et d'améliorer l'efficacité globale du processus.

En conclusion, l'application de l'approche méthodologique au cas d'usage industriel de SEW USOCOME a permis de valider les contributions scientifiques dans un contexte réel. Cette expérimentation a permis de démontrer que l'approche méthodologique est robuste et adaptable, et qu'elle peut également être étendue à d'autres processus industriels représentés par des séries temporelles multivariées.

Chapitre 4 :

Conclusion et perspectives

4.1 Conclusion

Cette thèse qui résulte d'un partenariat de recherche entre SEW USOCOME et les laboratoires ICUBE et CRAN, s'inscrit dans un contexte d'amélioration des pratiques de maintenance prévisionnelle au sein d'environnements industriels complexes. En réponse aux problématiques scientifiques identifiées, ce travail de recherche a permis le développement de trois contributions méthodologiques originales du domaine du PHM visant à détecter et diagnostiquer les dérives des processus de production hétérogènes, tout en investiguant, un déploiement généralisé de ces solutions à grande échelle (et non pas uniquement dans le contexte SEW USOCOME).

La **première problématique scientifique** abordée concerne le développement de méthodes de détection de dérives pour les processus de production, sans nécessiter de données étiquetées comme dérives. En ce sens, la **première contribution scientifique** propose une méthodologie guidée pour le choix d'algorithmes de détection de dérives. Cette méthodologie s'articule spécifiquement sur des algorithmes de détection de nouveautés, qui ne requièrent que des données conformes (fonctionnement normal) pour être entraînés. En s'appuyant sur des indicateurs caractérisant les données, la méthodologie permet, à partir d'un logigramme de décision, de sélectionner l'algorithme le plus adapté au processus surveillé.

La **deuxième problématique scientifique** porte sur la conception d'un système de diagnostic des dérives dans les processus de production hétérogènes, capable de combiner des méthodes basées sur les données et des méthodes basées sur les connaissances expertes. Pour y répondre, la **deuxième contribution scientifique** propose des outils méthodologiques permettant de construire un modèle basé sur les données pour identifier les dérives. À partir de cela, une méthodologie guidée pour le choix des algorithmes de classification binaire est définie. Elle s'appuie sur un tableau de complémentarité d'algorithmes de classification qui permet de construire des ensembles de classificateurs binaires pour chaque profil de dérive identifié. Ensuite, une proposition de modèle de diagnostic des dérives identifiées reposant sur l'intégration de la connaissance experte est faite, pour orienter le diagnostic vers les composants physiques en cause. Ceci implique de fournir, par un expert, une liste des composants physiques associés au processus surveillé, qui constitue la base à partir de laquelle un arbre de décision est élaboré. Cet arbre est structuré en fonction de la connaissance disponible sur le processus, et permet d'associer les dérives identifiées par les modèles basés sur les données aux composants physiques potentiellement en cause. L'arbre de décision doit pouvoir évoluer avec l'expérience opérationnelle. Ainsi, lorsqu'une dérive est détectée, les opérateurs peuvent intervenir et fournir un retour d'information sur les composants responsables. Ce retour enrichit progressivement l'arbre de décision : à chaque occurrence de dérive, des actions de maintenance sont documentées, et les composants en cause sont intégrés au modèle sous forme de règles décisionnelles. Au fil du temps, cette accumulation de retours d'expérience permet de cibler de plus en plus précisément les éléments physiques responsables des dérives identifiées, rendant le diagnostic des dérives toujours plus fiable.

Enfin, la **troisième problématique scientifique** vise l'élaboration d'une approche méthodologique proposant intégrant des étapes de détection et de diagnostic des dérives pour les processus de production hétérogènes. La **troisième contribution scientifique** élabore cette dite approche méthodologique, divisée en trois phases et intégrant les deux premières contributions scientifiques.

La **première phase** de l'approche méthodologique exploite la première contribution scientifique pour choisir un algorithme, et entraîner un modèle de détection de dérives. Ce modèle, une fois mis en exploitation, détecte les dérives dont les données représentatives sont stockées dans une base de données dédiée. Ce stockage permet d'accumuler les données qui ne correspondent pas à la conformité du processus surveillé à l'instant t.

La **deuxième phase** exploite la seconde contribution scientifique et la base de données des dérives pour créer les modèles de diagnostic de dérive. Dans un premier temps, un expert regroupe avec des algorithmes d'apprentissage automatique, la base de données des dérives afin de regrouper les données par cluster (similarité). Dans un second temps, un expert examine les clusters afin d'étiqueter les profils de données identifiés. En exploitant les apports méthodologiques de la seconde contribution scientifique, des modèles de diagnostic sont créés pour chaque profil de données identifié.

Enfin, la **troisième phase** assure la continuité de l'apprentissage dans l'approche méthodologique en intégrant de nouvelles données de dérives, et des données conformes au fur et à mesure de leur acquisition. Cette phase permet la mise à jour continue des modèles de détection et de diagnostic, garantissant leurs adaptabilités aux évolutions des processus de production. Cette contribution méthodologique permet donc de déployer un système de détection et de diagnostic de dérives des processus de production avec un minimum de données afin de faciliter une généralisation des solutions de maintenance prévisionnelle à l'échelle d'une « usine de production ».

Une première étape de validation des contributions scientifiques de cette thèse a été réalisée à travers un cas d'usage industriel fourni par SEW USOCOME. Ceci a permis, dans un cas d'usage spécifique et restreint (donc limité), de démontrer l'applicabilité et l'intérêt des approches proposées au cas d'un processus de production réel. Cette application a mis en évidence que, par les contributions proposées, à la fois les problématiques scientifiques ont été « résolues », mais aussi les verrous industriels identifiés dans le chapitre 1 et référant aux contraintes industrielles définies. Les résultats obtenus confirment que le modèle de détection de dérives, issu de la première contribution, a pu identifier avec précision les écarts de performance tout en minimisant le besoin de données étiquetées. Cette capacité de détection a permis de constituer une base de données de dérives fiables, qui a enrichi les activités de diagnostic.

En appliquant la deuxième contribution, les modèles de diagnostic développés se sont révélés capables de lier les dérives identifiées à des causes racines spécifiques, impactant ainsi de façon positive l'efficacité et l'efficacité des interventions de maintenance sur site. En effet, le recours à une architecture hybride combinant classification binaire et connaissance experte a permis de cibler les composants en cause avec rapidité et donc réduire les temps d'arrêt, tout en optimisant les ressources engagées.

Les résultats montrent également que la troisième contribution, qui intègre l'apprentissage continu, a permis de maintenir les performances des modèles de détection et de diagnostic en s'adaptant aux nouvelles configurations et variations du processus au fil du temps. Cette adaptabilité démontre la robustesse et la viabilité des solutions proposées pour un déploiement à grande échelle dans les usines de SEW USOCOME.

En conclusion, les présentes contributions scientifiques apportent des contributions méthodologiques pour les activités de détection de dérive et de diagnostic dans un contexte PHM et répondent aux problématiques scientifique et industrielles de la thèse.

4.2 Perspectives

Malgré les résultats prometteurs obtenus, certaines limitations subsistent en regard du déploiement, à l'échelle, de la méthodologie, mais aussi de la crédibilisation des contributions dans un contexte plus général. Ce constat soulève par conséquent plusieurs perspectives à envisager.

Tout d'abord, bien que le l'approche méthodologique ait été validée sur un cas d'usage industriel, sa généralité reste à démontrer dans d'autres environnements de production. Des phases de validation supplémentaires sur une plus grande échelle sont nécessaires pour confirmer son adaptabilité à divers processus. Chez SEW USOCOME, différents processus, tels que les lignes de fonderie ou les procédés d'usinage, où les données de dérive peuvent varier considérablement, sont des expérimentations pertinentes à envisager pour valider de manière plus pointue, dans des contextes différents, les contributions scientifiques (à des fins de généralisation). De plus, l'utilisation de données simulées pour ces phases de validation est inapplicable en raison de la complexité liée à la génération de données qui reflète la diversité des processus de production réels, avec toutes leurs incertitudes.

Concernant les modèles d'identification des dérives basés sur des ensembles de classificateurs binaires, une piste d'amélioration réside dans l'attribution de poids aux différents modèles de classification. En donnant plus de poids aux modèles les plus précis, l'identification des dérives doit être améliorée.

Cet aspect n'a pas pu être pris en compte dans cette thèse en raison de limitations dans la disponibilité des données. Si cette contrainte de données venait à être levée, une voie de recherche prometteuse consisterait à développer des algorithmes d'optimisation de l'attribution des poids pour chaque classificateur. De telles méthodes permettraient d'améliorer significativement la précision globale des ensembles de classificateurs binaires, en exploitant plus efficacement les performances spécifiques de chaque algorithme. De plus, dans la

phase de tri des données de l'approche méthodologique, des travaux sont à mener pour explorer les techniques d'apprentissage semi-supervisé et de regroupement avec auto-étiquetage pour réduire la dépendance à l'expertise humaine dans la fusion et l'affinement des clusters. Des approches d'apprentissage actif, où le modèle sollicite l'expert uniquement dans les cas ambigus ou incertains, sont une solution pour alléger la charge de travail des dits experts. Cette méthode hybride est prometteuse, car elle rationalise le fonctionnement du système. Cependant, l'apprentissage actif présente des défis liés à la complexité des données de production et à la variabilité des schémas de dérive selon les équipements et les conditions opérationnelles.

En outre, l'utilisation des techniques d'auto-étiquetage et d'apprentissage par transfert pourrait aider le modèle à se généraliser plus efficacement à divers processus, avec une intervention humaine minimale, surtout lorsque les données étiquetées sur les dérives sont limitées. Le développement d'algorithmes capables d'adapter automatiquement les paramètres de regroupement au fur et à mesure de l'apparition de nouvelles données et profils opérationnels améliorerait la scalabilité et la précision de l'approche méthodologique. Ces avancées renforceraient la capacité de l'approche méthodologique à s'adapter aux environnements de production en évolution constante, tout en limitant la dépendance aux interventions d'experts (permettant ainsi une meilleure efficacité de ce dernier).

Une autre perspective scientifique a pour objet l'intégration de la phase de pronostic à l'approche méthodologique, permettant de prévoir la durée de vie restante des processus de production, et de recommander des actions de maintenance en conséquence. Cet ajout requiert de proposer de nouvelles contributions méthodologiques en support du choix des algorithmes adaptés de pronostic, en aval des phases de détection et de diagnostic de dérive.

Sur un axe plus métier, une perspective intéressante à ce travail se dessine autour des Systèmes de Production Reconfigurables (RMS), qui sont au cœur de l'industrie 4.0. L'idée est de rendre les lignes de production plus adaptables et modulaires, quasiment par elles-mêmes. Ce faisant, les lignes de production doivent devenir capables de s'auto-surveiller et de s'autodiagnostiquer, cela pour leur conférer une grande autonomie, et donc une grande capacité à se reconfigurer au besoin.

Chez le partenaire SEW USOCOME, cet objectif s'inscrit pleinement dans la feuille de route en matière d'innovation. L'entreprise s'efforce de développer des lignes de production RMS pour accroître sa flexibilité et sa réactivité, des qualités essentielles pour répondre aux fluctuations de la demande et aux exigences croissantes de personnalisation des produits. Toutefois, atteindre un tel niveau d'autonomie demeure un défi de taille, partagé à l'échelle de la communauté du PHM.

Ce besoin accru d'intelligence et de réactivité dans les systèmes RMS a conduit l'Europe à soutenir des initiatives comme le projet MODAPTO¹⁰ dans lesquels est impliqué SEW USOCOME, qui explore le potentiel des jumeaux numériques modulaires, conçus pour rendre les systèmes RMS encore plus agiles et adaptables aux changements.

Dans ce contexte, les contributions scientifiques de cette thèse, axées sur la détection et le diagnostic des dérives dans des processus de production industrielle hétérogènes, apportent des avancées concrètes pour répondre aux défis de flexibilité et d'adaptabilité dans l'industrie moderne, et s'intègrent parfaitement dans cette dynamique.

¹⁰ <https://modapto.eu>

Bibliographie de l'auteur

Publication dans un journal

Julien Chapelin, Alexandre Voisin, Bertrand Rose, Benoît Iung, Lionel Steck, Ludovic Chaves, Mathieu Lauer, Olivier Jotz, *Data-driven drift detection and diagnosis framework for predictive maintenance of heterogeneous production processes: Application to a multiple tapping process*, **Engineering Applications of Artificial Intelligence**, Volume 139, Part A, 2025, <https://doi.org/10.1016/j.engappai.2024.109552>.

Conférence internationale avec comité de lecture et actes

Julien CHAPELIN, Alexandre VOISIN, Bertrand ROSE, Benoît IUNG, Lionel STECK, Mathieu LAUER, Ludovic CHAVES, Olivier JOTZ, *Data-driven drift detection and diagnostic for heterogeneous production process*, Proceeding in 22nd IFAC World Congress, 9-14 Juillet 2023, Yokohama, Japan, <https://doi.org/10.1016/j.ifacol.2023.10.1112>.

Julien CHAPELIN, Lionel STECK, Alexandre VOISIN, Benoît IUNG, Bertrand ROSE, *Digital continuity to improve the performance of the Industry 4.0*, Proceeding in 10th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2022, 22-24 juin 2022, Nantes, France, <https://doi.org/10.1016/j.ifacol.2022.09.501>.

Conférence nationale avec comité de lecture et actes

Julien CHAPELIN, Bertrand ROSE, Benoît IUNG, Lionel STECK, Mathieu LAUER, *Intégration du jumeau numérique dans le cycle de vie d'un projet industriel*, Proceeding in 17ème colloque national S-mart (AIP-PRIMECA), 31 mars -2 avril 2021, Valenciennes, France

Julien CHAPELIN, *Développement d'une interface logicielle orientée utilisateur*, Proceeding in eCONFERE 2020 - 27 ème colloque des Sciences de la conception et de l'innovation, 2 juillet 2020, online

Principales interventions de diffusion (présentations)

Julien CHAPELIN, *L'Intelligence Artificielle au service de la maintenance et de la durabilité : le cas de SEW Usocome*, Innova maintenance 2024, 24 octobre 2024, Strasbourg, France.

Julien CHAPELIN, *Mettre l'intelligence artificielle au service de la maintenance*, Salon SEPÉM 2024, 4 juin 2024, Colmar, France.

Julien CHAPELIN, *Détection et diagnostic basé sur les données pour la maintenance prévisionnelle des processus de production hétérogènes*, Journée partenaires Industrie du Futur et Usine Intelligente (IFUI) 2024, 23 janvier 2024, Illkirch, France.

Julien CHAPELIN, *La place de l'intelligence artificielle dans la maintenance prévisionnelle : application à un cas d'usage industriel*, L'after des managers SEW USOCOME, 10 juillet 2023, Brumath, France.

Julien CHAPELIN, *La place de l'intelligence artificielle dans la maintenance prévisionnelle : application à un cas d'usage industriel*, Réseau Résilian : L'intelligence artificielle, un outil pour la sobriété énergétique, la production et l'emploi, 25 mai 2023, Haguenau, France.

Julien CHAPELIN, *Predictive maintenance of industrial production lines based on analysis of heterogeneous production processes*, Atelier doctoral "Predictive maintenance & PHM: How AI is contributing?", 25 novembre 2022, Nancy, France.

Julien CHAPELIN, *Predictive maintenance of industrial production lines based on analysis of heterogeneous production processes*, AI and Industry 4.0 Forum, 3 octobre 2022, Nancy, France.

Julien CHAPELIN, *Continuité numérique des processus d'automatisation, capitalisation des données et standardisation pour l'industrie 4.0*, Assemblée générale SAGIP, 23-25 mai 2022, Bidart, France.

Bibliographie

- Abedi, M., Aghamohammadi, M.R. and Ameli, M.T. (2022) 'SVM based intelligent predictor for identifying critical lines with potential for cascading failures using pre-outage operating data', *International Journal of Electrical Power & Energy Systems*, 136, p. 107608. Available at: <https://doi.org/10.1016/J.IJEPES.2021.107608>.
- Adamczyk, B., Szejka, A. and Canciglieri Junior, O. (2019) 'Knowledge-based expert system to support the semantic interoperability in smart manufacturing', *Computers in Industry*, 115, pp. 1–16. Available at: <https://doi.org/10.1016/j.compind.2019.103161>.
- Aghabozorgi, S., Seyed Shirkhorshidi, A. and Ying Wah, T. (2015) 'Time-series clustering – A decade review', *Information Systems*, 53, pp. 16–38. Available at: <https://doi.org/10.1016/J.IS.2015.04.007>.
- Aimiyekagbon, O.K. et al. (2021) 'Rule-based Diagnostics of a Production Line', *Phme*, pp. 527–536.
- Ait tchakoucht, T. et al. (2024) 'Random forest with feature selection and K-fold cross validation for predicting the electrical and thermal efficiencies of air based photovoltaic-thermal systems', *Energy Reports*, 12, pp. 988–999. Available at: <https://doi.org/10.1016/J.EGYR.2024.07.002>.
- Aivaliotis, P. et al. (2021) 'Degradation curves integration in physics-based models: Towards the predictive maintenance of industrial robots', *Robotics and Computer-Integrated Manufacturing*, 71, p. 102177. Available at: <https://doi.org/10.1016/j.rcim.2021.102177>.
- Alaswad, S. and Xiang, Y. (2017) 'A review on condition-based maintenance optimization models for stochastically deteriorating system', *Reliability Engineering and System Safety*, 157, pp. 54–63. Available at: <https://doi.org/10.1016/J.RESS.2016.08.009>.
- Aldeoes Yasser N. and Gokhale, P. and S.S.Y. (2023) 'A Review of Predictive Maintenance of Bearing Failures in Rotary Machines by Predictive Analytics Using Machine-Learning Techniques', in P.N. and J.P. and S.G.R. Neustein Amy and Mahalle (ed.) *AI, IoT, Big Data and Cloud Computing for Industry 4.0*. Cham: Springer International Publishing, pp. 115–138. Available at: https://doi.org/10.1007/978-3-031-29713-7_6.
- Al-Dulaimi, A. et al. (2019) 'A multimodal and hybrid deep neural network model for Remaining Useful Life estimation', *Computers in Industry*, 108, pp. 186–196. Available at: <https://doi.org/10.1016/J.COMPIND.2019.02.004>.
- Al-Selwi, S.M. et al. (2024) 'RNN-LSTM: From applications to modeling techniques and beyond— Systematic review', *Journal of King Saud University - Computer and Information Sciences*, 36(5), p. 102068. Available at: <https://doi.org/10.1016/J.JKSUCI.2024.102068>.
- Alshorman, O. et al. (2020) 'A Review of Artificial Intelligence Methods for Condition Monitoring and Fault Diagnosis of Rolling Element Bearings for Induction Motor', *Shock and Vibration*, 2020(Cm). Available at: <https://doi.org/10.1155/2020/8843759>.
- Amini, A., Wah, T. and Saboohi, H. (2014) 'On Density-Based Data Streams Clustering Algorithms: A Survey', *Journal of Computer Science and Technology*, 29, pp. 116–141. Available at: <https://doi.org/10.1007/s11390-013-1416-3>.
- Aragon, Y. (2011) 'Modèles de base en séries temporelles', in Y. Aragon (ed.) *Séries temporelles avec R: Méthodes et cas*. Paris: Springer Paris, pp. 57–95. Available at: https://doi.org/10.1007/978-2-8178-0208-4_4.
- Atoui, M.A. and Cohen, A. (2021) 'Coupling data-driven and model-based methods to improve fault diagnosis', *Computers in Industry*, 128, p. 103401. Available at: <https://doi.org/10.1016/J.COMPIND.2021.103401>.
- Aydemir, G. et al. (2021) 'An Ensemble of LSTM Networks for Fault Detection, Classification, and Root Cause Identification in Quality Control Line', *PHM Society European Conference*, 6, pp. 537–542.
- Bagheri, B., Siegel, D., et al. (2015) 'A stochastic asset life prediction method for large fleet datasets in big data environment', *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)*, 14–2015. Available at: <https://doi.org/10.1115/IMECE2015-52458>.
- Bagheri, B., Yang, S., et al. (2015) 'Cyber-physical systems architecture for self-aware machines in industry 4.0 environment', in *IFAC-PapersOnLine*, pp. 1622–1627. Available at: <https://doi.org/10.1016/j.ifacol.2015.06.318>.
- Bender, A. and Sextro, W. (2021) 'Hybrid Prognosis Method for Remaining Useful Lifetime Estimation Considering Uncertainties', *PHM Society European Conference*, 6, pp. 42–52.
- Bohte, S., Poutré, H. and Kok, J. (2002) 'Unsupervised clustering with spiking neurons by sparse temporal coding and multilayer RBF networks', *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 13, pp. 426–435. Available at: <https://doi.org/10.1109/72.991428>.

- Bourgana, T. *et al.* (2021) 'Wavelet Scattering Network Based Bearing Fault Detection', *PHM Society European Conference*, pp. 80–87.
- Breiman, L. (2001) 'Random Forests', *Machine learning*, 45, pp. 5–32.
- Campos, G. *et al.* (2016) 'On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study', *Data Mining and Knowledge Discovery*, 30. Available at: <https://doi.org/10.1007/s10618-015-0444-8>.
- Canizo, M. *et al.* (2019) 'Multi-Head CNN-RNN for Multi-Time Series Anomaly Detection: An industrial case study', *Neurocomputing*, 363, pp. 246–260.
- Cao, S. *et al.* (2021) 'Research on fault diagnosis technology of centrifugal pump blade crack based on PCA and GMM', *Measurement*, 173, p. 108558. Available at: <https://doi.org/10.1016/J.MEASUREMENT.2020.108558>.
- Castellano-Quero, M. *et al.* (2023) 'A multidimensional Bayesian architecture for real-time anomaly detection and recovery in mobile robot sensory systems', *Engineering Applications of Artificial Intelligence*, 125, p. 106673. Available at: <https://doi.org/10.1016/J.ENGAPPAI.2023.106673>.
- Chalapathy, R. and Chawla, S. (2019) 'Deep Learning for Anomaly Detection: A Survey', *ArXiv*, abs/1901.03407. Available at: <https://api.semanticscholar.org/CorpusID:57825713>.
- Chan, J., Han, T. and Pan, E. (2024) 'Variational autoencoder-driven adversarial SVDD for power battery anomaly detection on real industrial data', *Journal of Energy Storage*, 103, p. 114267. Available at: <https://doi.org/10.1016/J.EST.2024.114267>.
- Chapelin, J. (2021) 'Intégration du jumeau numérique dans le cycle de vie d'un projet industriel.', in *S-MART 2021 : 17ème colloque national S-mart (AIP-PRIMECA)*, pp. 156–163. Available at: https://s-mart2021.sciencesconf.org/data/pages/Actes17ColloqueS_mart_5.pdf.
- Chapelin, J. *et al.* (2022) 'Digital continuity to improve performance of the Industry 4.0 Digital continuity to improve the performance of the Industry 4.0', *IFAC PapersOnLine*, 55(10), pp. 761–766. Available at: <https://doi.org/10.1016/j.ifacol.2022.09.501>.
- CHAPELIN, J. *et al.* (2023) 'Data-driven drift detection and diagnostic for heterogeneous production process', *IFAC-PapersOnLine*, 56(2), pp. 2102–2107. Available at: <https://doi.org/10.1016/j.ifacol.2023.10.1112>.
- Chen, J. *et al.* (2019) 'Anomaly detection for wind turbines based on the reconstruction of condition parameters using stacked denoising autoencoders', *Renewable Energy*, 147. Available at: <https://doi.org/10.1016/j.renene.2019.09.041>.
- Chen, X. *et al.* (2023) *Deep Transfer Learning for Bearing Fault Diagnosis: A Systematic Review Since 2016*. Available at: <https://www.ieee.org/publications/rights/rights-policies.html>.
- Chen, Y. *et al.* (2024) 'K-means clustering method based on nearest-neighbor density matrix for customer electricity behavior analysis', *International Journal of Electrical Power & Energy Systems*, 161, p. 110165. Available at: <https://doi.org/10.1016/J.IJEPES.2024.110165>.
- Chiachío, J. *et al.* (2019) 'A knowledge-based prognostics framework for railway track geometry degradation', *Reliability Engineering and System Safety*, 181, pp. 127–141. Available at: <https://doi.org/10.1016/J.RESS.2018.07.004>.
- Cholette, M.E. *et al.* (2019) 'Degradation modeling and condition-based maintenance of boiler heat exchangers using gamma processes', *Reliability Engineering & System Safety*, 183, pp. 184–196. Available at: <https://doi.org/10.1016/J.RESS.2018.11.023>.
- Climente-Alarcon, V. *et al.* (2017) 'Combined Model for Simulating the Effect of Transients on a Damaged Rotor Cage', *IEEE Transactions on Industry Applications*, 53(4), pp. 3528–3537. Available at: <https://doi.org/10.1109/TIA.2017.2691001>.
- Dhakar, A., Singh, B. and Gupta, P. (2023) 'Fault diagnosis of air compressor set-up using decision tree based J48 classification algorithm', *Journal of Engineering Research* [Preprint]. Available at: <https://doi.org/10.1016/J.JER.2023.09.028>.
- Ding, Z. and Fei, M. (2013) 'An Anomaly Detection Approach Based on Isolation Forest Algorithm for Streaming Data using Sliding Window', *IFAC Proceedings Volumes*, 46(20), pp. 12–17. Available at: <https://doi.org/https://doi.org/10.3182/20130902-3-CN-3020.00044>.
- Dinh, D.H., Do, P. and lung, B. (2020) 'Maintenance optimisation for multi-component system with structural dependence: Application to machine tool sub-system', *CIRP Annals*, 69(1), pp. 417–420. Available at: <https://doi.org/10.1016/J.CIRP.2020.04.004>.
- Dong, D., Li, X.Y. and Sun, F.Q. (2017) 'Life prediction of jet engines based on LSTM-recurrent neural networks', *2017 Prognostics and System Health Management Conference, PHM-Harbin 2017 - Proceedings* [Preprint]. Available at: <https://doi.org/10.1109/PHM.2017.8079264>.

- Downey, A. *et al.* (2019) 'Physics-based prognostics of lithium-ion battery using non-linear least squares with dynamic bounds', *Reliability Engineering & System Safety*, 182, pp. 1–12. Available at: <https://doi.org/https://doi.org/10.1016/j.res.2018.09.018>.
- Einabadi, B. *et al.* (2023) 'Dynamic predictive and preventive maintenance planning with failure risk and opportunistic grouping considerations: A case study in the automotive industry', *Journal of Manufacturing Systems*, 69, pp. 292–310. Available at: <https://doi.org/10.1016/J.JMSY.2023.06.012>.
- Es-sakali, N. *et al.* (2022) 'Review of predictive maintenance algorithms applied to HVAC systems', *Energy Reports*, 8(May), pp. 1003–1012. Available at: <https://doi.org/10.1016/j.egy.2022.07.130>.
- Fahad, A. *et al.* (2014) 'A Survey of Clustering Algorithms for Big Data: Taxonomy and Empirical Analysis', *Emerging Topics in Computing, IEEE Transactions on*, 2, pp. 267–279. Available at: <https://doi.org/10.1109/TETC.2014.2330519>.
- Fahle, S., Prinz, C. and Kuhlenkötter, B. (2020a) 'Systematic review on machine learning (ML) methods for manufacturing processes - Identifying artificial intelligence (AI) methods for field application', in *Procedia CIRP*. Elsevier B.V., pp. 413–418. Available at: <https://doi.org/10.1016/j.procir.2020.04.109>.
- Fahle, S., Prinz, C. and Kuhlenkötter, B. (2020b) 'Systematic review on machine learning (ML) methods for manufacturing processes - Identifying artificial intelligence (AI) methods for field application', in *Procedia CIRP*. Elsevier B.V., pp. 413–418. Available at: <https://doi.org/10.1016/j.procir.2020.04.109>.
- Fassois, S.D. and Kopsaftopoulos, F.P. (2013a) 'Statistical Time Series Methods for Vibration Based Structural Health Monitoring', in *CISM International Centre for Mechanical Sciences, Courses and Lectures*. Springer International Publishing, pp. 209–264. Available at: https://doi.org/10.1007/978-3-7091-1390-5_4.
- Fassois, S.D. and Kopsaftopoulos, F.P. (2013b) 'Statistical Time Series Methods for Vibration Based Structural Health Monitoring', in *CISM International Centre for Mechanical Sciences, Courses and Lectures*. Springer International Publishing, pp. 209–264. Available at: https://doi.org/10.1007/978-3-7091-1390-5_4.
- Fejjari, A., Saheb Ettabaa, K. and Korbaa, O. (2021) 'Feature Extraction Techniques for Hyperspectral Images Classification', in V.E. Balas *et al.* (eds) *Soft Computing Applications*. Cham: Springer International Publishing, pp. 174–188.
- Foube, L. (2021) 'Power Devices Health Condition Monitoring: A Review of Recent Papers', pp. 136–150.
- Freyermuth, B. (1991) 'Knowledge based Incipient Fault Diagnosis of Industrial Robots', *IFAC Proceedings Volumes*, 24(6), pp. 369–375. Available at: [https://doi.org/10.1016/S1474-6670\(17\)51169-6](https://doi.org/10.1016/S1474-6670(17)51169-6).
- Galar, M. *et al.* (2011) 'An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes', *Pattern Recognition*, 44(8), pp. 1761–1776. Available at: <https://doi.org/10.1016/j.patcog.2011.01.017>.
- Gao, Y. *et al.* (2023) 'Drift-aware Anomaly Detection for Non-stationary Time Series', in *Proceedings - 2023 IEEE International Conference on Big Data, BigData 2023*. Institute of Electrical and Electronics Engineers Inc., pp. 1095–1100. Available at: <https://doi.org/10.1109/BigData59044.2023.10386160>.
- Glaessgen, E.H. and Stargel, D.S. (2012) 'The digital twin paradigm for future NASA and U.S. Air force vehicles', *Collection of Technical Papers - AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, pp. 1–14. Available at: <https://doi.org/10.2514/6.2012-1818>.
- Goswami, P. and Nandan Rai, R. (2023) 'A systematic review on failure modes and proposed methodology to artificially seed faults for promoting PHM studies in laboratory environment for an industrial gearbox', *Engineering Failure Analysis*. Elsevier Ltd. Available at: <https://doi.org/10.1016/j.engfailanal.2023.107076>.
- Guillén Antonio J. and González-Prida, V. and G.J.F. and C.A. (2016) 'Standards as Reference to Build a PHM-Based Solution', in H. and A.J. and U.T. and K.K. and M.J. and L.J. Koskinen Kari T. and Kortelainen (ed.) *Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015)*. Cham: Springer International Publishing, pp. 207–214.
- Hajihosseini, M., Maghsoudi, A. and Ghezelbash, R. (2024) 'Intelligent mapping of geochemical anomalies: Adaptation of DBSCAN and mean-shift clustering approaches', *Journal of Geochemical Exploration*, 258, p. 107393. Available at: <https://doi.org/10.1016/J.GEXPLO.2024.107393>.
- Hamidi, A.A., Robertson, B. and Ilow, J. (2023) 'A new approach for ECG artifact detection using fine-KNN classification and wavelet scattering features in vital health applications', *Procedia Computer Science*, 224, pp. 60–67. Available at: <https://doi.org/10.1016/J.PROCS.2023.09.011>.
- He, Y. *et al.* (2020) 'An ontology-based method of knowledge modelling for remanufacturing process planning', *Journal of Cleaner Production*, 258, p. 120952. Available at: <https://doi.org/10.1016/j.jclepro.2020.120952>.
- Heaton, J. (2018) 'Deep learning', *Genetic Programming and Evolvable Machines*, 19(1), pp. 305–307. Available at: <https://doi.org/10.1007/s10710-017-9314-z>.

- Hiruta, T. *et al.* (2021) 'Unsupervised Learning Based Diagnosis Model for Anomaly Detection of Motor Bearing with Current Data', in *Procedia CIRP*. Elsevier B.V., pp. 336–341. Available at: <https://doi.org/10.1016/j.procir.2021.01.113>.
- Hossen, M.J. *et al.* (2024) 'Unsupervised novelty detection for time series using a deep learning approach', *Heliyon*, 10(3), p. e25394. Available at: <https://doi.org/10.1016/J.HELIYON.2024.E25394>.
- Hoteit, I. and Köhl, A. (2006) 'Efficiency of reduced-order, time-dependent adjoint data assimilation approaches', *Journal of Oceanography*, 62(4), pp. 539–550. Available at: <https://doi.org/10.1007/s10872-006-0074-2>.
- Huang, C. *et al.* (2024) 'Prognostics and health management for predictive maintenance: A review', *Journal of Manufacturing Systems*, 75, pp. 78–101. Available at: <https://doi.org/10.1016/J.JMSY.2024.05.021>.
- Huang, K. *et al.* (2024) 'IPORF: A combined improved parrot optimizer algorithm and random forest for fault diagnosis in AUV', *Ocean Engineering*, 313, p. 119665. Available at: <https://doi.org/10.1016/J.OCEANENG.2024.119665>.
- Ibs, I. *et al.* (2024) 'From human explanations to explainable AI: Insights from constrained optimization', *Cognitive Systems Research*, 88, p. 101297. Available at: <https://doi.org/10.1016/J.COGSYS.2024.101297>.
- Islam, Md.S., Pears, R. and Bacic, B. (2018) 'A wavelet approach for precursor pattern detection in time series', *Journal of Electrical Systems and Information Technology*, 5(3), pp. 337–348. Available at: <https://doi.org/10.1016/J.JESIT.2018.03.003>.
- ISO (2003) 'ISO 13374-1, Condition monitoring and diagnostics of machines — Data processing, communication and presentation — Part 1: General guidelines, International Standards Organization', *International Standards Organization* [Preprint].
- Jain, A.K., Murty, M.N. and Flynn, P.J. (1999) 'Data clustering: a review', *ACM Comput. Surv.*, 31, pp. 264–323. Available at: <https://api.semanticscholar.org/CorpusID:12744045>.
- Jardine, A.K.S., Lin, D. and Banjevic, D. (2006) 'A review on machinery diagnostics and prognostics implementing condition-based maintenance', *Mechanical Systems and Signal Processing*, 20(7), pp. 1483–1510. Available at: <https://doi.org/10.1016/J.YMSSP.2005.09.012>.
- Javed, K., Gouriveau, R. and Zerhouni, N. (2017) 'State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels', *Mechanical Systems and Signal Processing*, 94, pp. 214–236. Available at: <https://doi.org/10.1016/J.YMSSP.2017.01.050>.
- Jha, A., Dave, M. and Madan, S. (2019) *Comparison of Binary Class and Multi-Class Classifier Using Different Data Mining Classification Techniques*. Available at: <https://ssrn.com/abstract=3464211>.
- Jimenez, V.J., Bouhmala, N. and Gausdal, A.H. (2020) 'Developing a predictive maintenance model for vessel machinery', *Journal of Ocean Engineering and Science*, 5(4), pp. 358–386. Available at: <https://doi.org/10.1016/j.joes.2020.03.003>.
- Jin, T. *et al.* (2024) 'Double-classifier adversarial learning for fault diagnosis of rotating machinery considering cross domains', *Mechanical Systems and Signal Processing*, 216, p. 111490. Available at: <https://doi.org/10.1016/J.YMSSP.2024.111490>.
- Jin, X. and Ni, J. (2019) 'Physics-based Gaussian process for the health monitoring for a rolling bearing', *Acta astronautica*, 154, pp. 133–139.
- Jing, T. *et al.* (2022) 'Transformer-based hierarchical latent space VAE for interpretable remaining useful life prediction', *Advanced Engineering Informatics*, 54, p. 101781. Available at: <https://doi.org/10.1016/J.AEI.2022.101781>.
- Kamm, S., Jazdi, N. and Weyrich, M. (2021) 'Knowledge Discovery in Heterogeneous and Unstructured Data of Industry 4.0 Systems: Challenges and Approaches', *Procedia CIRP*, 104, pp. 975–980. Available at: <https://doi.org/10.1016/j.procir.2021.11.164>.
- Kammerer, C. *et al.* (2020) 'Comparison of predictive maintenance methods for thermal systems', *Procedia Computer Science*, 176, pp. 166–174. Available at: <https://doi.org/10.1016/j.procs.2020.08.018>.
- Kans, M. and Campos, J. (2024) 'Digital capabilities driving industry 4.0 and 5.0 transformation: Insights from an interview study in the maintenance domain', *Journal of Open Innovation: Technology, Market, and Complexity*, 10(4), p. 100384. Available at: <https://doi.org/10.1016/J.JOITMC.2024.100384>.
- Kashpruk, N., Piskor-Ignatowicz, C. and Baranowski, J. (2023) 'Time Series Prediction in Industry 4.0: A Comprehensive Review and Prospects for Future Advancements', *Applied Sciences (Switzerland)*. Multidisciplinary Digital Publishing Institute (MDPI). Available at: <https://doi.org/10.3390/app132212374>.
- Khan, A.A., Chaudhari, O. and Chandra, R. (2024) 'A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation', *Expert Systems with Applications*, 244, p. 122778. Available at: <https://doi.org/10.1016/j.eswa.2023.122778>.

- Khan, S. and Madden, M. (2009) *A Survey of Recent Trends in One Class Classification, Artificial Intelligence and Cognitive Science, Lecture Notes in Computer Science, vol. 6206*. Available at: https://doi.org/10.1007/978-3-642-17080-5_21.
- Khanzadeh, M. et al. (2018) 'Porosity prediction: Supervised-learning of thermal history for direct laser deposition', *Journal of Manufacturing Systems*, 47(April), pp. 69–82. Available at: <https://doi.org/10.1016/j.jmsy.2018.04.001>.
- Kürs,at, K. et al. (2021) 'Fault Detection and Classification for Robotic Test-bench: À Data Challenge', *PHM Society European Conference*, 6, pp. 552–558.
- Kusiak, A. (2019) 'Fundamentals of smart manufacturing: A multi-thread perspective', *Annual Reviews in Control*, 47(xxxx), pp. 214–220. Available at: <https://doi.org/10.1016/j.arcontrol.2019.02.001>.
- Lee, J., Ghaffari, M. and Elmeligy, S. (2011) 'Self-maintenance and engineering immune systems: Towards smarter machines and manufacturing systems', *Annual Reviews in Control*, 35(1), pp. 111–122. Available at: <https://doi.org/10.1016/j.arcontrol.2011.03.007>.
- Lei, Y. et al. (2018) 'Machinery health prognostics: A systematic review from data acquisition to RUL prediction', *Mechanical Systems and Signal Processing*, 104, pp. 799–834. Available at: <https://doi.org/10.1016/J.YMSSP.2017.11.016>.
- Lei, Y. et al. (2020) 'Applications of machine learning to machine fault diagnosis: A review and roadmap', *Mechanical Systems and Signal Processing*, 138, p. 106587. Available at: <https://doi.org/10.1016/j.ymsp.2019.106587>.
- Li, C. et al. (2024) 'Small data challenges for intelligent prognostics and health management: a review', *Artificial Intelligence Review*, 57(8). Available at: <https://doi.org/10.1007/s10462-024-10820-4>.
- Li, D. et al. (2024) 'Sensor-aware CapsNet: Towards trustworthy multisensory fusion for remaining useful life prediction', *Journal of Manufacturing Systems*, 72, pp. 26–37. Available at: <https://doi.org/10.1016/J.JMSY.2023.11.009>.
- Li, J. et al. (2018) 'Three-dimensional Simulation and Prediction of Solenoid Valve Failure Mechanism Based on Finite Element Model', *IOP Conference Series: Earth and Environmental Science*, 108(2). Available at: <https://doi.org/10.1088/1755-1315/108/2/022035>.
- Li, M. et al. (2024) 'A Fault Detection Method for Electric Vehicle Battery System Based on Bayesian Optimization SVDD Considering a Few Faulty Samples', *Energy Engineering*, 121(9), pp. 2543–2568. Available at: <https://doi.org/10.32604/EE.2024.051231>.
- Li, S. et al. (2024) 'An intelligent framework of upgraded CapsNets with massive transmissibility data for identifying damage in bridges', *Applied Soft Computing*, 155, p. 111459. Available at: <https://doi.org/10.1016/J.ASOC.2024.111459>.
- Liu, J. et al. (2019) 'Anomaly detection for time series using temporal convolutional networks and Gaussian mixture model', in *Journal of Physics: Conference Series*. Institute of Physics Publishing. Available at: <https://doi.org/10.1088/1742-6596/1187/4/042111>.
- Luo, H., Huang, M. and Zhou, Z. (2018) 'Integration of Multi-Gaussian fitting and LSTM neural networks for health monitoring of an automotive suspension component', *Journal of Sound and Vibration*, 428, pp. 87–103. Available at: <https://doi.org/10.1016/J.JSV.2018.05.007>.
- Luo, Y. et al. (2023) 'Adaptive online optimization of alarm thresholds using multilayer Bayesian networks and active transfer entropy', *Control Engineering Practice*, 137(December 2022), p. 105534. Available at: <https://doi.org/10.1016/j.conengprac.2023.105534>.
- Lyu, F. et al. (2024) '3D in-situ stress prediction for shale reservoirs based on the CapsNet-BiLSTM hybrid model', *International Journal of Rock Mechanics and Mining Sciences*, 183, p. 105937. Available at: <https://doi.org/10.1016/J.IJRMMS.2024.105937>.
- Ma, J. and Perkins, S. (2003) 'Time-series Novelty Detection Using One-class Support Vector Machines', *Proceedings of the International Joint Conference on Neural Networks*, 3(June), pp. 1741–1745. Available at: <https://doi.org/10.1109/ijcnn.2003.1223670>.
- Mahfoud Hassana and Moutaoukil, O. and T.B.M. and L.A. (2024) 'Real-Time Predictive Maintenance-Based Process Parameters: Towards an Industrial Sustainability Improvement', in J. and B.V.E. Ezziyyani Mostafa and Kacprzyk (ed.) *International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023)*. Cham: Springer Nature Switzerland, pp. 18–34.
- Majstorović, V.D. and Milačić, V.R. (1990) 'Expert systems for maintenance in the CIM concept', *Computers in Industry*, 15(1–2), pp. 83–93. Available at: [https://doi.org/10.1016/0166-3615\(90\)90086-5](https://doi.org/10.1016/0166-3615(90)90086-5).
- Mao, S. et al. (2020) 'Dis-AE-LSTM: Generative Adversarial Networks for Anomaly Detection of Time Series Data', in *Proceedings - 2020 International Conference on Artificial Intelligence and Computer Engineering*,

- ICAICE 2020. Institute of Electrical and Electronics Engineers Inc., pp. 330–336. Available at: <https://doi.org/10.1109/ICAICE51518.2020.00070>.
- Mateus, B. *et al.* (2022) 'Data Analysis for Predictive Maintenance Using Time Series and Deep Learning Models—A Case Study in a Pulp Paper Industry', in, pp. 11–25. Available at: https://doi.org/10.1007/978-3-030-99075-6_2.
- El Mawas, Z., Cappelle, C. and El Badaoui El Najjar, M. (2023) 'Decision Tree based diagnosis for hybrid model-based/data-driven fault detection and exclusion of a decentralized multi-vehicle cooperative localization system*', *IFAC-PapersOnLine*, 56(2), pp. 7740–7745. Available at: <https://doi.org/10.1016/J.IFACOL.2023.10.1179>.
- Medina-Oliva, G. *et al.* (2014) 'Predictive diagnosis based on a fleet-wide ontology approach', *Knowledge-Based Systems*, 68, pp. 40–57. Available at: <https://doi.org/10.1016/j.knosys.2013.12.020>.
- Mejri, N. *et al.* (2024) 'Unsupervised anomaly detection in time-series: An extensive evaluation and analysis of state-of-the-art methods', *Expert Systems with Applications*, 256, p. 124922. Available at: <https://doi.org/10.1016/J.ESWA.2024.124922>.
- Mian, Z. *et al.* (2024) 'A literature review of fault diagnosis based on ensemble learning', *Engineering Applications of Artificial Intelligence*, 127. Available at: <https://doi.org/10.1016/j.engappai.2023.107357>.
- Mohammadi, A. *et al.* (2023) 'Analysis of Numerical Integration in RNN-Based Residuals for Fault Diagnosis of Dynamic Systems', *IFAC-PapersOnLine*, 56(2), pp. 2909–2914. Available at: <https://doi.org/10.1016/J.IFACOL.2023.10.1411>.
- Monostori, L. *et al.* (2016) 'Cyber-physical systems in manufacturing', *CIRP Annals*, 65(2), pp. 621–641. Available at: <https://doi.org/10.1016/j.cirp.2016.06.005>.
- Montero Jimenez, J.J. *et al.* (2020) 'Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics', *Journal of Manufacturing Systems*, 56(March), pp. 539–557. Available at: <https://doi.org/10.1016/j.jmsy.2020.07.008>.
- Morgan, J. *et al.* (2021) 'Industry 4.0 smart reconfigurable manufacturing machines', *Journal of Manufacturing Systems*, 59(November 2020), pp. 481–506. Available at: <https://doi.org/10.1016/j.jmsy.2021.03.001>.
- Mosallam, A., Medjaher, K. and Zerhouni, N. (2016) 'Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction', *Journal of Intelligent Manufacturing*, 27(5), pp. 1037–1048. Available at: <https://doi.org/10.1007/S10845-014-0933-4>.
- Mühlbauer, M. *et al.* (2022) 'Deviation Detection in Production Processes based on Video Data using Unsupervised Machine Learning Approaches', *Procedia CIRP*, 112, pp. 162–167. Available at: <https://doi.org/10.1016/j.procir.2022.09.066>.
- Munir, M. *et al.* (2019) 'DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series', *IEEE Access*, 7, pp. 1991–2005. Available at: <https://doi.org/10.1109/ACCESS.2018.2886457>.
- Nacchia, M. *et al.* (2021) 'A systematic mapping of the advancing use of machine learning techniques for predictive maintenance in the manufacturing sector', *Applied Sciences (Switzerland)*. MDPI AG. Available at: <https://doi.org/10.3390/app11062546>.
- Nasution, F.P., Sævik, S. and Gjøsteen, J.K.Ø. (2012) 'Fatigue analysis of copper conductor for offshore wind turbines by experimental and FE method', in *Energy Procedia*. Elsevier Ltd, pp. 271–280. Available at: <https://doi.org/10.1016/j.egypro.2012.06.109>.
- Ngo, G., Beard, R. and Chandra, R. (2022) 'Evolutionary bagging for ensemble learning', *Neurocomputing*, 510, pp. 1–14. Available at: <https://doi.org/10.1016/j.neucom.2022.08.055>.
- Nunes, P., Santos, J. and Rocha, E. (2023) 'Challenges in predictive maintenance – A review', *CIRP Journal of Manufacturing Science and Technology*. Elsevier Ltd, pp. 53–67. Available at: <https://doi.org/10.1016/j.cirpj.2022.11.004>.
- Núñez, D.L. and Borsato, M. (2017) 'An ontology-based model for prognostics and health management of machines', *Journal of Industrial Information Integration*, 6, pp. 33–46. Available at: <https://doi.org/10.1016/j.jii.2017.02.006>.
- Olsson, T. *et al.* (2021) 'A data-driven approach for predicting long-term degradation of a fleet of micro gas turbines', *Energy and AI*, 4, p. 100064. Available at: <https://doi.org/10.1016/j.egyai.2021.100064>.
- Panthi, M. (2020) *Anomaly Detection in Smart Grids using Machine Learning Techniques*. Available at: <https://doi.org/10.1109/ICPC2T48082.2020.9071434>.
- Peng, A. *et al.* (2023) 'Diagnosis, Feedback, Adaptation: A Human-in-the-Loop Framework for Test-Time Policy Adaptation', in A. Krause *et al.* (eds) *Proceedings of the 40th International Conference on Machine Learning*. PMLR (Proceedings of Machine Learning Research), pp. 27630–27641. Available at: <https://proceedings.mlr.press/v202/peng23c.html>.

- Pimentel, M.A.F. *et al.* (2014) 'A review of novelty detection', *Signal Processing*, pp. 215–249. Available at: <https://doi.org/10.1016/j.sigpro.2013.12.026>.
- Potthoff Jean Paul and Tienken, C. and F.T. (2023) 'Business Models for Commercializing Predictive Maintenance in the Context of Industrial Sustainability', in S. and B.T. Meierhofer Jürg and West (ed.) *Smart Services Summit*. Cham: Springer Nature Switzerland, pp. 163–170.
- Qi, J., Chu, Y. and He, L. (2018) 'Iterative anomaly detection algorithm based on time series analysis', in *Proceedings - 15th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2018*. Institute of Electrical and Electronics Engineers Inc., pp. 548–552. Available at: <https://doi.org/10.1109/MASS.2018.00085>.
- Qi, J., Gryllias, K. and Mauricio, A. (2021) 'Multiple-Model Estimation-based Prognostics for Rotating Machinery', *PHM Society European Conference*, 6(1), pp. 11–11. Available at: <https://papers.phmsociety.org/index.php/phme/article/view/2956>.
- Qiao, G. and Weiss, Brian A (2018) 'Quick health assessment for industrial robot health degradation and the supporting advanced sensing development', *Journal of Manufacturing Systems*, 48, pp. 51–59. Available at: <https://doi.org/https://doi.org/10.1016/j.jmsy.2018.04.004>.
- Qiao, G. and Weiss, Brian A. (2018) 'Quick health assessment for industrial robot health degradation and the supporting advanced sensing development', *Journal of Manufacturing Systems*, 48, pp. 51–59. Available at: <https://doi.org/10.1016/J.JMSY.2018.04.004>.
- Raja, M. *et al.* (2024) 'Membership determination in open clusters using the DBSCAN Clustering Algorithm', *Astronomy and Computing*, 47, p. 100826. Available at: <https://doi.org/10.1016/J.ASCOM.2024.100826>.
- Ran, Y. *et al.* (2019) 'A Survey of Predictive Maintenance: Systems, Purposes and Approaches'. Available at: <http://arxiv.org/abs/1912.07383>.
- Rehab, A. *et al.* (2021) 'Bearings Fault Detection Using Hidden Markov Models and Principal Component Analysis Enhanced Features', *PHM Society European Conference*, 6(1), pp. 11–11. Available at: <http://papers.phmsociety.org/index.php/phme/article/view/2947>.
- Ren, L. *et al.* (2018) 'Bearing remaining useful life prediction based on deep autoencoder and deep neural networks', *Journal of Manufacturing Systems*, 48, pp. 71–77. Available at: <https://doi.org/10.1016/J.JMSY.2018.04.008>.
- Ruan, D. *et al.* (2023) 'CNN parameter design based on fault signal analysis and its application in bearing fault diagnosis', *Advanced Engineering Informatics*, 55, p. 101877. Available at: <https://doi.org/10.1016/J.AEI.2023.101877>.
- Ruff, L. *et al.* (2018) *Deep One-Class Classification*.
- Safari, A., Sabahi, M. and Oshnoei, A. (2024) 'ResFaultyMan: An intelligent fault detection predictive model in power electronics systems using unsupervised learning isolation forest', *Heliyon*, 10(15), p. e35243. Available at: <https://doi.org/10.1016/J.HELIVON.2024.E35243>.
- Sahli, A., Evans, R. and Manohar, A. (2021) 'Predictive Maintenance in Industry 4.0: Current Themes', *Procedia CIRP*, 104, pp. 1948–1953. Available at: <https://doi.org/10.1016/j.procir.2021.11.329>.
- Van De Sand, R. and Eng, M.A.Y. (2020) *A Predictive Maintenance Model for Heterogeneous Industrial Refrigeration Systems*. Thesis. University of Rome.
- Sankararaman, S., Daigle, M.J. and Goebel, K. (2014) 'Uncertainty Quantification in Remaining Useful Life Prediction Using First-Order Reliability Methods', *IEEE Transactions on Reliability*, 63(2), pp. 603–619. Available at: <https://doi.org/10.1109/TR.2014.2313801>.
- Sarmadi, H. and Karamodin, A. (2020) 'A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects', *Mechanical Systems and Signal Processing*, 140, p. 106495. Available at: <https://doi.org/10.1016/J.YMSSP.2019.106495>.
- Satoła, A. and Satoła, K. (2024) 'Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking, and super-learner ensembles versus single machine learning models', *Journal of Dairy Science*, 107(6), pp. 3959–3972. Available at: <https://doi.org/10.3168/JDS.2023-24243>.
- Saxena, A. *et al.* (2017) 'A review of clustering techniques and developments', *Neurocomputing*, 267, pp. 664–681. Available at: <https://doi.org/https://doi.org/10.1016/j.neucom.2017.06.053>.
- Schmidl, S., Wenig, P. and Papenbrock, T. (2022) 'Anomaly Detection in Time Series: A Comprehensive Evaluation', in *Proceedings of the VLDB Endowment*. VLDB Endowment, pp. 1779–1797. Available at: <https://doi.org/10.14778/3538598.3538602>.

- Schwartz, S. *et al.* (2020) 'A fault mode identification methodology based on self-organizing map', *Neural Computing and Applications*, 32(17), pp. 13405–13423. Available at: <https://doi.org/10.1007/s00521-019-04692-x>.
- Serras, J.L., Vinga, S. and Carvalho, A.M. (2021) 'Outlier detection for multivariate time series using dynamic bayesian networks', *Applied Sciences (Switzerland)*, 11(4), pp. 1–22. Available at: <https://doi.org/10.3390/app11041955>.
- Sikorska, J.Z., Hodkiewicz, M. and Ma, L. (2011) 'Prognostic modelling options for remaining useful life estimation by industry', *Mechanical Systems and Signal Processing*, 25(5), pp. 1803–1836. Available at: <https://doi.org/10.1016/J.YMSSP.2010.11.018>.
- Silvestri, L. *et al.* (2020) 'Maintenance transformation through Industry 4.0 technologies: A systematic literature review', *Computers in Industry*, 123, p. 103335. Available at: <https://doi.org/10.1016/J.COMPIND.2020.103335>.
- Siraskar, R. *et al.* (2023) 'Reinforcement learning for predictive maintenance: a systematic technical review', *Artificial Intelligence Review*, 56(11), pp. 12885–12947. Available at: <https://doi.org/10.1007/s10462-023-10468-6>.
- Sisode, M. and Devare, M. (2023) 'A Review on Machine Learning Techniques for Predictive Maintenance in Industry 4.0', in, pp. 774–783. Available at: https://doi.org/10.2991/978-94-6463-136-4_67.
- Soualhi, A. *et al.* (2014) 'Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System', *IEEE Transactions on Industrial Electronics*, 61(6), pp. 2864–2874. Available at: <https://doi.org/10.1109/TIE.2013.2274415>.
- Studer, R., Benjamins, V.R. and Fensel, D. (1998) 'Knowledge engineering: Principles and methods', *Data & Knowledge Engineering*, 25(1–2), pp. 161–197. Available at: [https://doi.org/10.1016/S0169-023X\(97\)00056-6](https://doi.org/10.1016/S0169-023X(97)00056-6).
- Sun, B. *et al.* (2023) 'Ensemble learning based on stacking and blending predicts glass forming ability', *Materials Today Communications*, 37. Available at: <https://doi.org/10.1016/j.mtcomm.2023.107385>.
- Sürmeli, B.G. and Tümer, M.B. (2020) 'Multivariate Time Series Clustering and its Application in Industrial Systems', *Cybernetics and Systems*, 51(3), pp. 315–334. Available at: <https://doi.org/10.1080/01969722.2019.1691851>.
- Swain, R., Khilar, P. and Bhoi, S. (2017) 'Heterogeneous Fault Diagnosis for Wireless Sensor Networks', *Ad Hoc Networks*, 69. Available at: <https://doi.org/10.1016/j.adhoc.2017.10.012>.
- Teixeira, H.N., Lopes, I. and Braga, A.C. (2020a) 'Condition-based maintenance implementation: A literature review', in *Procedia Manufacturing*. Elsevier B.V., pp. 228–235. Available at: <https://doi.org/10.1016/j.promfg.2020.10.033>.
- Teixeira, H.N., Lopes, I. and Braga, A.C. (2020b) 'Condition-based maintenance implementation: a literature review', *Procedia Manufacturing*, 51, pp. 228–235. Available at: <https://doi.org/10.1016/J.PROMFG.2020.10.033>.
- Thomas, L. *et al.* (2021) 'Ontology-based representation of tolerancing and design knowledge for an automated tolerance specification of product concepts', *Procedia CIRP*, 52(2019), pp. 194–199. Available at: <https://doi.org/10.1016/j.net.2020.03.028>.
- Tso, G.K.F. and Yau, K.K.W. (2007) 'Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks', *Energy*, 32(9), pp. 1761–1768. Available at: <https://doi.org/10.1016/j.energy.2006.11.010>.
- Vachtsevanos, G. *et al.* (2007) 'Intelligent Fault Diagnosis and Prognosis for Engineering Systems', *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, pp. 1–434. Available at: <https://doi.org/10.1002/9780470117842>.
- Vafaei, N., Ribeiro, R.A. and Camarinha-Matos, L.M. (2019) 'Fuzzy early warning systems for condition based maintenance', *Computers and Industrial Engineering*, 128, pp. 736–746. Available at: <https://doi.org/10.1016/J.CIE.2018.12.056>.
- Vanek, B. *et al.* (2012) 'Fault detection of electrical flight control system actuators using parameter dependent estimation', *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 8(PART 1), pp. 1358–1363. Available at: <https://doi.org/10.3182/20120829-3-MX-2028.00218>.
- Verbert, K., Babuška, R. and De Schutter, B. (2017) 'Bayesian and Dempster–Shafer reasoning for knowledge-based fault diagnosis—A comparative study', *Engineering Applications of Artificial Intelligence*, 60, pp. 136–150. Available at: <https://doi.org/10.1016/J.ENGAPAI.2017.01.011>.
- Verbert, K., De Schutter, B. and Babuška, R. (2015) 'Reasoning under uncertainty for knowledge-based fault diagnosis: A comparative study', *IFAC-PapersOnLine*, 28(21), pp. 422–427. Available at: <https://doi.org/10.1016/J.IFACOL.2015.09.563>.

Vingerhoeds, R.A. *et al.* (1995) 'Enhancing off-line and on-line condition monitoring and fault diagnosis', *Control Engineering Practice*, 3(11), pp. 1515–1528. Available at: [https://doi.org/10.1016/0967-0661\(95\)00162-N](https://doi.org/10.1016/0967-0661(95)00162-N).

Vogl, G.W., Weiss, B.A. and Donmez, M.A. (2015) 'A sensor-based method for diagnostics of machine tool linear axes', *Proceedings of the Annual Conference of the Prognostics and Health Management Society, PHM*, pp. 408–417.

Wang, B. *et al.* (2024) 'A rolling bearing fault diagnosis technique based on recurrence quantification analysis and Bayesian optimization SVM', *Applied Soft Computing*, 156, p. 111506. Available at: <https://doi.org/10.1016/J.ASOC.2024.111506>.

Wang, H. *et al.* (2024) 'A novel deep-learning model for detecting small-scale anomaly temperature zones in RDTs based on attention mechanism and K-Means clustering', *Optical Fiber Technology*, 88, p. 103969. Available at: <https://doi.org/10.1016/J.YOFTE.2024.103969>.

Wang, Youming and Wang, Yue (2023) 'A denoising semi-supervised deep learning model for remaining useful life prediction of turbofan engine degradation', *Applied Intelligence*, 53(19), pp. 22682–22699. Available at: <https://doi.org/10.1007/s10489-023-04777-0>.

Weisburd, D. *et al.* (2020) 'Measuring Association for Scaled Data: Pearson's Correlation Coefficient', in D. Weisburd *et al.* (eds) *Basic Statistics in Criminology and Criminal Justice*. Cham: Springer International Publishing, pp. 479–530. Available at: https://doi.org/10.1007/978-3-030-47967-1_14.

Weiss, B.A. *et al.* (2016) 'Use Case Development to Advance Monitoring, Diagnostics, and Prognostics in Manufacturing Operations', *IFAC-PapersOnLine*, 49(31), pp. 13–18. Available at: <https://doi.org/10.1016/J.IFACOL.2016.12.154>.

Wescoat, E., Krugh, M. and Mears, L. (2021) 'Random forest regression for predicting an anomalous condition on a UR10 cobot end-effector from purposeful failure data', *Procedia Manufacturing*, 53, pp. 644–655. Available at: <https://doi.org/10.1016/J.PROMFG.2021.06.064>.

Wilhelm Kirch (2008) 'Pearson's Correlation Coefficient', in W. Kirch (ed.) *Encyclopedia of Public Health*. Dordrecht: Springer Netherlands, pp. 1090–1091. Available at: https://doi.org/10.1007/978-1-4020-5614-7_2569.

Wu, R. (2024) 'Behavioral analysis of electricity consumption characteristics for customer groups using the k-means algorithm', *Systems and Soft Computing*, 6, p. 200143. Available at: <https://doi.org/10.1016/J.SASC.2024.200143>.

Xiao, Y. *et al.* (2016) 'Robust one-class SVM for fault detection', *Chemometrics and Intelligent Laboratory Systems*, 151, pp. 15–25. Available at: <https://doi.org/10.1016/J.CHEMOLAB.2015.11.010>.

Xie, S., Li, L. and Zhu, Y. (2024) 'Anomaly detection for multivariate time series in IoT using discrete wavelet decomposition and dual graph attention networks', *Computers & Security*, 146, p. 104075. Available at: <https://doi.org/10.1016/J.COSE.2024.104075>.

Xu, W. *et al.* (2022) 'RUL prediction for rolling bearings based on Convolutional Autoencoder and status degradation model', *Applied Soft Computing*, 130, p. 109686. Available at: <https://doi.org/10.1016/j.asoc.2022.109686>.

Yan, Y. *et al.* (2024) 'An unknown fault diagnosis Scheme: A novel random deep forest for fault diagnosis of HVACs', *Energy and Buildings*, 321, p. 114660. Available at: <https://doi.org/10.1016/J.ENBUILD.2024.114660>.

Yang, Y., Yu, B. and Wang, W. (2021) 'Research on equipment health prediction technology based on edge computing and VAE-TCN', *Procedia Computer Science*, 183, pp. 100–106. Available at: <https://doi.org/10.1016/J.PROCS.2021.02.036>.

Ye, Y. *et al.* (2023) 'Multivariate Time Series Anomaly Detection with Fourier Time Series Transformer', in *2023 IEEE 12th International Conference on Cloud Networking, CloudNet 2023*. Institute of Electrical and Electronics Engineers Inc., pp. 381–388. Available at: <https://doi.org/10.1109/CloudNet59005.2023.10490086>.

Yıldız, A. and Kalayci, A. (2024) *Gradient Boosting Decision Trees on Medical Diagnosis over Tabular Data*. Available at: <https://doi.org/10.48550/arXiv.2410.03705>.

Yu, H. and Li, H. (2021) 'A conditional factor VAE model for pump degradation assessment under varying conditions', *Applied Soft Computing*, 100, p. 106992. Available at: <https://doi.org/10.1016/J.ASOC.2020.106992>.

Zhang, C. and Yang, T. (2023) 'Anomaly Detection for Wind Turbines Using Long Short-Term Memory-Based Variational Autoencoder Wasserstein Generation Adversarial Network under Semi-Supervised Training', *Energies*, 16, p. 7008. Available at: <https://doi.org/10.3390/en16197008>.

Zhang, J., Liang, J. and Liu, J. (2024) 'Fault diagnosis in reactor coolant pump with an automatic CNN-based mixed model', *Progress in Nuclear Energy*, 175, p. 105294. Available at: <https://doi.org/10.1016/J.PNUCENE.2024.105294>.

Zhang, M. and Yan, J. (2021) 'A data-driven method for optimizing the energy consumption of industrial robots', *Journal of Cleaner Production*, 285, p. 124862. Available at: <https://doi.org/10.1016/j.jclepro.2020.124862>.

Zhao, B.Y. *et al.* (2024) 'A deep belief rule base-based fault diagnosis method for complex systems', *ISA Transactions*, 150, pp. 77–91. Available at: <https://doi.org/10.1016/J.ISATRA.2024.05.019>.

Zhao, R. *et al.* (2019) 'Deep learning and its applications to machine health monitoring', *Mechanical Systems and Signal Processing*, 115, pp. 213–237. Available at: <https://doi.org/https://doi.org/10.1016/j.ymssp.2018.05.050>.

Zhou, A., Yu, D. and Zhang, W. (2015) 'A research on intelligent fault diagnosis of wind turbines based on ontology and FMECA', *Advanced Engineering Informatics*, 29(1), pp. 115–125. Available at: <https://doi.org/10.1016/J.AEI.2014.10.001>.

Zimek, A., Campello, R. and Sander, J. (2014) 'Ensembles for unsupervised outlier detection: challenges and research questions. A position paper', *ACM SIGKDD Explorations Newsletter*, 15, pp. 11–22. Available at: <https://doi.org/10.1145/2594473.2594476>.

Julien CHAPELIN

Détection et diagnostic de dérives de processus de production hétérogènes et complexes : proposition d'une approche générique d'intelligence artificielle basée sur l'apprentissage continu

Résumé

Cette thèse s'inscrit dans le contexte de la maintenance prévisionnelle des processus de production hétérogènes, en ciblant trois problématiques scientifiques : la détection des dérives sans besoin de données étiquetées, le diagnostic des dérives à l'aide de méthodes hybrides, et l'intégration de ces activités dans une approche méthodologique adaptable. La première contribution propose, via un logigramme décisionnel, une méthode guidée pour sélectionner les algorithmes d'apprentissage automatique les mieux adaptés aux données via le calcul d'indicateurs. La deuxième contribution présente un diagnostic hybride, combinant classificateurs binaires et arbres de décision enrichis par des connaissances expertes, pour identifier les causes racine des dérives. Enfin, la troisième contribution élabore une approche méthodologique intégrée de détection et diagnostic des dérives de processus intégrant des étapes d'apprentissage continu, pour permettre l'adaptation aux évolutions des processus. Validée sur un processus industriel chez SEW USOCOME, cette approche montre des résultats prometteurs et confirme la pertinence des contributions.

Mots clés : Maintenance prévisionnelle ; Processus de production hétérogènes ; Détection de dérive ; Diagnostic de dérive ; Cadre méthodologique ; Connaissances expertes ; Apprentissage d'ensemble ; Détection des nouveautés ; Apprentissage continu

Abstract

This thesis is set in the context of predictive maintenance for heterogeneous production processes, targeting three scientific issues: drift detection without the need for labeled data, drift diagnosis using hybrid methods, and the integration of these activities into an adaptable methodological approach. The first contribution proposes, via a decision logigram, a guided method for selecting the machine learning algorithms best suited to the data via the calculation of indicators. The second contribution presents a hybrid diagnosis, combining binary classifiers and decision trees enriched with expert knowledge, to identify the root causes of drift. Finally, the third contribution develops an integrated methodological approach to process drift detection and diagnosis, incorporating continuous learning steps to enable adaptation to process evolutions. Validated on an industrial process at SEW USOCOME, this approach shows promising results and confirms the relevance of the contributions.

Keywords : Predictive maintenance; Heterogeneous production processes; Drift detection; Drift diagnosis; Methodological framework; Expert knowledge; Ensemble learning; Novelty detection; Continuous learning