Titre: Title:	Amélioration du contrôle qualité par des méthodes d'apprentissage automatique : application à un processus d'assemblage dans l'industrie du pneumatique
Auteur: Author:	Asma Ben Brahem
Date:	2025
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Ben Brahem, A. (2025). Amélioration du contrôle qualité par des méthodes d'apprentissage automatique : application à un processus d'assemblage dans l'industrie du pneumatique [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/70469/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/70469/
Directeurs de recherche: Advisors:	Camélia Dadouchi
Programme: Program:	Maîtrise recherche en génie industriel

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Amélioration du contrôle qualité par des méthodes d'apprentissage automatique : application à un processus d'assemblage dans l'industrie du pneumatique

ASMA BEN BRAHEM

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie industriel

Novembre 2025

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Amélioration du contrôle qualité par des méthodes d'apprentissage automatique : application à un processus d'assemblage dans l'industrie du pneumatique

présenté par **Asma BEN BRAHEM**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Bruno AGARD, président
Camélia DADOUCHI, membre et directrice de recherche
Daniel ALOISE, membre

DÉDICACE

À mes parents, véritables piliers de ma vie, pour les valeurs et la persévérance qu'ils m'ont transmises.

À mes frères, pour leur soutien constant tout au long de ce parcours.

À mon neveu, ma source d'énergie et de joie, ainsi qu'à mes amis précieux, merci de m'avoir accompagné dans cette belle aventure.

REMERCIEMENTS

A terme de ce projet, je tiens tout d'abord à exprimer ma profonde gratitude à **Camélia DADOUCHI**, directrice de recherche, pour son encadrement, ses conseils avisés et sa grande disponibilité. Sa patience, sa bienveillance et son soutien constant ont été une source d'inspiration et de motivation tout au long de ce travail.

Je souhaite également remercier toute l'équipe du Laboratoire en intelligence de données à Polytechnique Montréal, pour l'excellente ambiance de travail, leur esprit d'équipe et leur amitié. Vous avez fait de ce laboratoire bien plus qu'un lieu de recherche : un véritable environnement familial.

J'adresse aussi mes sincères remerciements à **Alexandre LEBLANC-RICHARD** pour son aide précieuse, sa disponibilité et son accompagnement dans ce projet, ainsi qu'à toute l'équipe du service informatique, dont le soutien et la collaboration ont grandement facilité l'avancement de mes travaux.

Enfin, je tiens à remercier mes amis à Montréal et en Tunisie, ma grande famille, et tout particulièrement ma petite famille, pour leur amour, leur écoute et leur soutien inébranlable. Votre présence m'a donné la force d'aller de l'avant et de persévérer jusqu'au bout.

RÉSUMÉ

Dans un contexte manufacturier, les produits traversent des lignes d'assemblage semi-automatisées au sein desquelles peuvent survenir différents défauts. Or parfois, l'inspection de ces défauts n'est possible qu'à la fin de la chaîne de production. Bien que ces anomalies soient rares, notamment dans des environnements fortement automatisés, elles engendrent néanmoins des coûts de production significatifs lorsqu'elles sont détectées tardivement. Dans cette optique, les entreprises cherchent à identifier ces défauts le plus tôt possible, afin d'interrompre la production ou de mettre en place des actions correctives adéquates.

Avec l'avènement de l'industrie 4.0, les industriels s'orientent vers des solutions intelligentes et temps réel, capables de répondre à ces besoins de détection précoce. La littérature scientifique recense plusieurs approches, basées soit sur des méthodes statistiques classiques, soit sur l'apprentissage automatique. Toutefois, peu d'études s'intéressent à la détection de défauts représentant une très faible proportion des données (moins de 5 %), et qui à la fois cherchent à identifier en temps réel les facteurs à l'origine de ces anomalies.

En collaboration avec une usine manufacturière de l'industrie du pneumatique au Québec, ce mémoire vise à améliorer la capacité de contrôle qualité en développant une méthode de détection en temps réel des défauts d'adhésion sur une ligne d'assemblage automatisée. Il cherche également à analyser les facteurs d'apparition de ces défauts afin de proposer des actions opérationnelles appropriées.

La méthodologie adoptée à cette fin, repose sur le cadre CRISP-DM (Cross-Industry Standard Process for Data Mining), permettant la préparation et l'analyse des données de production pertinentes relatives aux facteurs potentiels d'apparition des défauts. Une analyse exploratoire préliminaire a mis en évidence un fort déséquilibre entre les classes de données. Pour pallier ce problème, une stratégie hybride de rééchantillonnage combinant les méthodes Tomek-Links et CTGAN a été développée afin de gérer ce déséquilibre extrême.

Plusieurs techniques de classification ont ensuite été évaluées pour détecter les défauts d'adhésion, et différents algorithmes d'apprentissage automatique ont été comparés afin d'identifier le plus performant. Le modèle s'est révélé le plus adapté à la problématique étudiée, intégrant par ailleurs les principes du Contrôle Statistique des Procédés (SPC). Enfin, l'effet des variables explicatives sur la classification des défauts a été analysé à l'aide de techniques d'explicabilité locale et globale.

L’application de cette méthodologie au cas concret du partenaire industriel a permis d’obtenir un modèle atteignant un F1-score de 84 %, offrant un compromis optimal entre capacité de détection et réduction des fausses alertes.

Ce travail propose ainsi des pistes concrètes pour une détection proactive et une amélioration continue du contrôle qualité dans un contexte industriel automatisé.

ABSTRACT

In modern manufacturing environments, products move through semi-automated assembly lines where material adhesion defects can occasionally occur. These defects, however, are typically inspected only at the end of the production line. Although such anomalies are rare, especially in highly automated settings, they can lead to considerable production costs when detected late. Manufacturers therefore seek to identify these defects as early as possible to halt production or take corrective action before losses occur.

Driven by the rise of Industry 4.0, industrial players are increasingly adopting intelligent, real-time solutions to address this challenge. Existing research highlights a range of approaches, from classical statistical methods to advanced machine learning techniques. Yet, few studies focus on detecting very low defect rates (below 5%) while simultaneously identifying, in real time, the factors that contribute to these anomalies.

This research, conducted in collaboration with a tire manufacturing plant in Québec, aims to enhance quality control by developing a real-time method for detecting adhesion defects on an automated assembly line. It also seeks to uncover the key factors influencing defect occurrence, enabling the implementation of targeted operational improvements.

The study follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, involving systematic data preparation to extract relevant production features and potential defect indicators. An exploratory data analysis first revealed a strong class imbalance, which was addressed through a hybrid resampling strategy combining Tomek-Links and CTGAN methods.

Multiple machine learning algorithms were evaluated to identify the most effective classification model for adhesion defect detection. The AdaBoost model achieved the best performance and was further integrated with Statistical Process Control (SPC) principles to strengthen quality monitoring. Finally, the contribution of explanatory variables to defect prediction was analyzed using both global and local explainability techniques. Applied to the industrial case study, the proposed approach achieved an F1-score of 84%, offering an effective balance between accurate detection and minimal false alarms.

Overall, this research provides a practical and data-driven pathway toward proactive defect detection and continuous improvement in industrial quality control.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

ADABOOST Renforcement adaptatif

AI Intelligence artificielle

CL Limite de contrôle

CPP Paramètres critiques du procédé

CQA Attributs critiques de la qualité

CRISP-DM Processus standard intersectoriel pour l'exploration de données

CTGAN Réseau antagoniste génératif conditionnel pour données tabulaires

FN Faux négatif

FP Faux positif

KNN Méthode des k plus proches voisins

LCL Limite de contrôle inférieure

LIME Explications locales interprétables indépendantes du modèle

SHAP Explications additives de Shapley

SPC Contrôle statistique du processus

TQM Gestion totale de la qualité

TN Vrai négatif

TP Vrai positif

UCL Limite de contrôle supérieure

CHAPITRE 1 INTRODUCTION

Dans le contexte de transformation du secteur manufacturier vers des systèmes de production de nouvelle génération, la qualité s'impose comme un levier stratégique central, conditionnant la compétitivité, la pérennité et la réputation des entreprises. Celle-ci se traduit par le respect des normes ou des spécifications techniques et également la capacité de l'entreprise à répondre de manière constante aux attentes des clients, à minimiser les pertes économiques et à maintenir un haut niveau de fiabilité opérationnelle. La qualité du produit final revêt une importance cruciale, car elle représente l'aboutissement de l'ensemble du processus de fabrication et, par conséquent, influence directement les coûts liés aux retours, aux rebuts et aux opérations de reprise, tout en pouvant également avoir un impact sur l'empreinte environnementale (Gupta & Madan, 2023).

Conscientes de ces enjeux, les industries manufacturières ont progressivement adopté des démarches méthodologiques structurées telles que le Six Sigma, le Lean Manufacturing (Lal Bhaskar, 2020) ou le Total Quality Management (TQM) (Talib & Rahman, 2012) . En agissant sur la réduction de la variabilité des procédés, l'élimination des gaspillages et l'optimisation des ressources, ces approches contribuent à la stabilité et à la maîtrise des processus de fabrication. Parallèlement, l'essor de l'automatisation et des systèmes de production assistés par ordinateur a permis d'accroître la cadence et la régularité de fabrication, tout en renforçant la précision et la répétabilité des opérations (Deepika et al., 2024), éléments essentiels à l'obtention de produits conformes aux normes de qualité.

Néanmoins, bien que ces outils et méthodologies ont contribué à réduire l'occurrence des défauts, l'atteinte du paradigme du « *zéro défaut* » demeure un défi majeur pour les entreprises manufacturières (Psarommatis et al., 2020). Des défauts rares mais critiques peuvent encore survenir, qu'elles soient liées à des défaillances matérielles (Zhang et al., 2019), à des dérives progressives et difficilement perceptibles dans les procédés (Pittino et al., 2020), ou encore à des facteurs externes échappant au contrôle direct de l'opérateur (Li et al., 2025).

Dans ce contexte, les avancées récentes en apprentissage automatique (*machine learning*) offrent de nouvelles perspectives pour la détection de défauts complexes, souvent difficilement identifiables par les approches conventionnelles. Parallèlement, l'intégration croissante de capteurs intelligents au sein des équipements industriels engendre un volume massif de données hétérogènes

collectées en temps réel. L'exploitation de ces données ouvre de nouvelles pistes de recherche et d'application en matière de détection de défauts, permettant non seulement une surveillance plus proactive des procédés, mais aussi la mise en place de stratégies de maintenance corrective adaptées aux environnements industriels modernes.

Dans le contexte d'une ligne d'assemblage semi-automatisée, caractérisée par la complexité inhérente des procédés et la polyvalence des machines dédiées à la production multi-produits, l'identification des défauts générés à cette étape demeure particulièrement difficile.

En collaboration avec un manufacturier du secteur automobile, ce travail propose une méthodologie fondée sur la construction d'un modèle d'apprentissage automatique destiné à la détection de défauts dans un contexte de déséquilibre de données. L'originalité de cette approche réside dans l'intégration conjointe de méthodes classiques de contrôle et d'approches fondées sur l'apprentissage automatique, permettant ainsi d'améliorer la capacité d'identification des défauts rares. Par ailleurs, une attention particulière est accordée à l'analyse des facteurs potentiels d'apparition de défauts grâce au recours à des méthodes d'explicabilité, offrant une interprétation des décisions prises par le modèle et favorisant leur exploitation dans un cadre industriel.

Le présent mémoire est structuré en cinq chapitres. Le deuxième chapitre présente une revue de la littérature mettant en lumière les principales méthodes de détection de défauts, les approches de traitement du déséquilibre de classes ainsi que les techniques d'explicabilité des modèles de décision. Le chapitre 3 présente les objectifs de cette recherche et détaille le développement de la méthodologie adoptée, laquelle s'appuie sur le cadre CRISP-DM pour l'adaptation de la détection de défauts à un contexte marqué par un fort déséquilibre des données. Le chapitre 4 illustre l'application de cette méthodologie à un cas d'étude réel mené en collaboration avec un partenaire industriel, dont l'objectif est d'améliorer la capacité et l'efficacité du processus de contrôle qualité en temps réel au sein d'une ligne d'assemblage semi-automatique en utilisant des méthodes d'apprentissage machine. Enfin, le Chapitre 5 synthétise les contributions de cette étude, discute de ses limites et propose des perspectives d'amélioration et de recherche futures.

CHAPITRE 2 REVUE DE LA LITTÉRATURE

Dans ce chapitre, une revue de la littérature est menée pour identifier et analyser les travaux existants en lien avec la problématique et les objectifs du présent projet de recherche. Ces objectifs s'inscrivent dans une démarche visant à améliorer le processus du contrôle qualité en temps réel dans une ligne d'assemblage semi-automatisée au sein d'une industrie manufacturière, dans le but de réduire les coûts liés aux rejets et à la reproduction de produits défectueux.

Afin de structurer cette revue, la section 2.1 présente tout d'abord un ensemble de définitions clés afin de bien clarifier les concepts fondamentaux liés à la problématique du contrôle de la qualité. La section 2.2 expose le résultat pour la revue de la littérature. Celle-ci aborde trois axes de recherche, le premier portant sur les méthodes de contrôle de la qualité dans le secteur manufacturier particulièrement de détection de défauts et de détection d'anomalies, le deuxième axe s'intéresse à une problématique récurrente dans les données industrielles de qualité, à savoir le déséquilibre des classes, et présente les principales approches développées pour le gérer. Et le dernier aborde l'explicabilité des modèles en vue d'identifier les facteurs conduisant à la non-qualité dans un cadre industriel.

Enfin, une synthèse critique clôture le chapitre, en soulignant les principales limites identifiées dans la littérature et en mettant en relief la contribution originale de ce projet de recherche.

2.1 Introduction et définitions

Dans le contexte de contrôle de la qualité dans le secteur manufacturier, l'exploration de la littérature révèle une hétérogénéité terminologique. Les travaux recensés s'articulent autour de deux approches principales, ceux traitant de la détection de défauts et ceux axés sur la détection d'anomalies. La détection de défauts vise la non-qualité des produits finaux, la détection d'anomalies s'intéresse aux comportements atypiques dans les données qui sont susceptibles d'apparaître au cours du processus (Klarák et al., 2024). Elles peuvent être traitées par les mêmes ou différentes approches. En effet, une anomalie peut constituer un signal annonciateur d'un futur défaut, mais également d'autres types de dysfonctionnements tels qu'un désalignement mécanique, une dérive capteur, ou une fluctuation anormale des paramètres de procédé, susceptibles d'entraîner des arrêts de machine ou des perturbations du flux de production (Chandola et al., 2009).

Ainsi, pour répondre à l'objectif fixé, il est pertinent de regrouper et d'analyser ces deux axes de recherche afin d'offrir un éventail plus large de méthodes pour l'amélioration du contrôle de la qualité. Cette analyse permet ainsi de développer une vision plus globale et de renforcer la pertinence des solutions envisagées dans ce travail.

2.2 Résultats de la revue

Ce travail de recherche adopte une revue de la littérature narrative pour fournir une synthèse des travaux existant sur l'amélioration de la capacité et l'efficacité du processus de contrôle qualité dans le secteur manufacturier. La méthode suivie dans ce travail est réalisée à partir de deux bases de données qui sont *Google Scholar* et *Engineering Village* pour l'identification et la collecte des publications scientifiques. Les sous sections suivantes portent sur quatre axes : la détection d'anomalies et la détection de défauts dans les environnements manufacturiers, le traitement du déséquilibre des données manufacturières et finalement aborde l'identification et l'analyse des facteurs responsables de l'apparition des défauts.

2.2.1 Approches de détection d'anomalies dans le secteur manufacturier

La problématique de la détection d'anomalies dans le secteur manufacturier a été abordée dans la littérature selon différentes approches. Ces travaux peuvent être regroupés en deux grandes catégories. La première regroupe les méthodes classiques, fondées principalement sur des outils statistiques, qui se concentrent sur l'étude des processus et l'identification de leurs déviations. La seconde regroupe les méthodes issues de l'apprentissage automatique, allant des techniques de classification jusqu'aux approches d'apprentissage profond utilisées pour la reconstruction de signaux et l'analyse de séries temporelles dans les processus industriels.

Les sections suivantes sont consacrées à une analyse approfondie des travaux identifiés dans la littérature sur ces différentes approches.

A. Méthodes de contrôle statistique de procédé

Les méthodes de contrôle statistique de procédés (SPC) constituent des outils fondamentaux pour le suivi de la qualité et la détection de variations anormales dans les processus industriels. Introduites par Walter A. Shewhart dans les années 1920, elles reposent sur le concept de cartes de contrôle (*control charts*) permettant de surveiller l'évolution d'un indicateur statistique et de détecter les écarts dépassant les limites de contrôle définies (Shewart 1931).

Dans un premier temps, les cartes univariées ont été proposées, telles que la carte $\bar{X}$ (pour la moyenne des sous-groupes) et la carte R (pour l'étendue des sous-groupes). Celles-ci sont destinées à suivre une seule caractéristique de qualité à la fois. Ces approches ont été largement utilisées et constituent encore aujourd'hui la base du SPC classique (Montgomery, 2009).

Cependant, les processus industriels modernes sont caractérisés par la collecte simultanée de multiples variables corrélées, rendant les cartes univariées insuffisantes. Pour pallier cette limite, des recherches ultérieures ont introduit des cartes multivariées qui permettent de prendre en compte la covariance entre les variables et d'améliorer la sensibilité à la détection des dérives multivariées (Hotelling, 1947; Lowry & Montgomery, 1995). Ces avancées ont marqué une étape importante dans l'évolution du SPC en l'adaptant aux environnements manufacturiers complexes.

Cartes de contrôle univariées

Les cartes de contrôle univariées reposent sur le calcul d'un indicateur statistique tels que la moyenne, l'étendue ou l'écart-type pour chaque lot de production et ensuite il s'agit de tracer son évolution sur un graphique. Ce graphique comprend donc une ligne centrale (CL) correspondant à la valeur attendue de l'indicateur lorsque le procédé est stable, une limite de contrôle supérieure (UCL) et une limite de contrôle inférieure (LCL) définies de manière statistique. Ces cartes servent à repérer les points qui franchissent les limites, ou si des motifs inhabituels apparaissent (tendance, cycles, regroupements). Cela indique la présence probable de causes spéciales nécessitant une investigation et une action corrective (Shewart 1931).

Les cartes univariées classiques identifiées dans cette revue se déclinent en trois types. La première est la carte $\bar{X}$ (*X-bar chart*), qui permet de suivre la moyenne des sous-groupes. La seconde est la carte R (*Range chart*), destinée à contrôler l'étendue de chaque sous-groupe afin de mesurer la dispersion. Enfin, la carte S (*Standard deviation chart*) est souvent privilégiée lorsque la taille des sous-groupes est plus importante, puisqu'elle repose sur l'écart-type (Park & Wang, 2020). Les limites de contrôle de ces cartes sont définies dans le Tableau 2.1.

Tableau 2.1. Limites de contrôle des différentes cartes de contrôle

Cartes de Contrôle	Carte $\bar{X}$	Carte R	Carte S
UCL	$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \,.\bar{R}$	$UCL_R = D_4 \,.\bar{R}$	$UCL_S = B_4 \,.\bar{S}$
CL	$CL_{\bar{X}} = \bar{\bar{X}}$	$CL_R = \bar{R}$	$CL_S = \bar{S}$
LCL	$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \,.\bar{R}$	$LCL_R = D_3 \,.\bar{R}$	$LCL_S = B_3 \,.\bar{S}$
Variables tabulées	A_2	$D_4 \; et \; D_3$	$B_4 \; et \; B_3$

Note : Les constantes A_2, D_3, D_4, B_3 et B_4 proviennent de tables normalisées en SPC et servent à déterminer les limites de contrôle en reliant l'amplitude ou l'écart-type observé aux écarts théoriques attendus, selon les équations indiquées.

Plusieurs travaux ont traité les cartes de contrôle univariées dans l'objectif de détection d'anomalies dans divers secteurs manufacturiers. Par exemple, dans le secteur automobile, où la qualité et la fiabilité sont primordiales, Teixeira et al. (2017) ont illustré l'efficacité des cartes de contrôle notamment $\bar{X}$ et R dans le suivi du couple de serrage des vis sur les lignes d'assemblage. Cette méthode statistique a offert ainsi un support fiable pour la prise de décision qualité des responsables industriels. Par ailleurs, Godina et al. (2016) se sont intéressés au suivi des processus de fabrication de pièces métalliques embouties, caractérisées par des spécifications dimensionnelles variées imposées par les clients. Ces cartes ont également été appliquées dans d'autres secteurs comme celui du pharmaceutique et l'agroalimentaire. Dans le secteur agroalimentaire, Shrestha (2021) a appliqué les cartes $\bar{X}$ et R pour contrôler le poids de lots de saucisses conditionnées. L'étude, réalisée à l'aide du logiciel Minitab, a confirmé que le procédé se trouvait sous contrôle statistique, tout en offrant des indications précieuses pour optimiser encore la régularité du poids des produits. De manière similaire, l'industrie pharmaceutique, soumise à des réglementations de qualité extrêmement rigoureuses, a bénéficié de ces outils. Riaz et Muhammad (2012) ont utilisé les cartes $\bar{X}$ et R pour surveiller des variables critiques, telles que le volume de remplissage, le pH, le pourcentage de citrate et le poids des flacons. Leur étude a montré que ces

cartes permettent d'assurer la conformité aux normes de qualité strictes du secteur. Cependant, cette étude montre par ailleurs que les cartes de contrôles univariées manquent de sensibilités face aux dérives progressives, nécessitant le recours à des outils complémentaires comme les cartes de contrôle à moyenne mobile pondérée exponentiellement (EWMA). Au-delà de cette limitation, d'autres travaux soulignent des contraintes méthodologiques inhérentes à l'approche univariée. En particulier, leur restriction à l'analyse d'une seule variable à la fois entrave la prise en compte des interdépendances complexes entre les variables et entre les observations (Bersimis et al., 2006). Afin de faire face à cette limite, une autre approche a utilisé de multiples cartes univariées pour surveiller un processus multivarié, néanmoins elle a montré que cette approche augmente significativement la probabilité de générer de fausses alarmes (Kosztyán & Katona, 2016).

Cartes de contrôle multivariées

Les limites identifiées dans les cartes de contrôle univariées ont conduit à l'émergence des cartes de contrôle multivariées, capables de traiter simultanément plusieurs variables corrélées. Dans un contexte industriel moderne, les procédés sont en effet caractérisés par la collecte de données issues de nombreux capteurs et paramètres interdépendants. L'analyse séparée de chaque variable à l'aide de cartes univariées ne permet pas toujours de détecter des dérives subtiles, alors qu'une approche multivariée peut révéler des anomalies uniquement perceptibles à travers la combinaison des variables.

L'une des premières approches multivariées proposées est la carte T^2 de Hotelling (1947). Cette méthode transforme un vecteur de variables corrélées en une statistique unique, permettant ainsi de détecter des décalages dans la moyenne multivariée. Plus tard, Lowry & Montgomery (1995) ont montré l'utilité pratique de cette technique dans des environnements de production réels, tout en soulignant certaines limites telles que la sensibilité du T^2 aux petites tailles d'échantillons et aux données à distributions non normales, pouvant conduire à des estimations peu fiables de la matrice de covariance et à une augmentation des fausses alarmes. Afin d'améliorer la sensibilité aux faibles décalages du procédé, Lowry et al. (1992) ont introduit la carte MEWMA (*Multivariate Exponentially Weighted Moving Average*), qui intègre l'information historique via un lissage exponentiel des données multivariées, la rendant particulièrement efficace pour la détection des dérives graduelles. Dans le même esprit, Pignatiello Jr & Runger (1990) ont adapté la carte CUSUM univariée au cadre multivarié en proposant la carte MCUSUM (*Multivariate Cumulative*

Sum), qui cumule les petites variations au fil du temps, ce qui la rend performante pour identifier des décalages faibles mais persistants de la moyenne du procédé. Plusieurs études comparatives ont montré que la carte MCUSUM surpasse généralement les cartes T^2 et MEWMA pour la détection des petits décalages soutenus (Kalgonda & Kulkarni, 2004; Woodall & Ncube, 1985) bien qu'elle soit plus exigeante sur le plan computationnel. Toutefois, comme le rappellent Bersimis et al. (2006), ces cartes reposent sur des hypothèses fortes de normalité multivariée et de disponibilité d'estimations fiables de la matrice de covariance, conditions rarement réunies en pratique. Plus récemment, Ramos et al. (2021) ont montré que, dans des environnements de production à forte variabilité et faible volume, les approches classiques du MSPC (Multivariate Statistical Process Control) rencontrent des difficultés face à la non-linéarité et aux interactions dynamiques des procédés. Ces constats ont favorisé l'émergence de méthodes intégrant des approches de type *data-driven* et apprentissage automatique.

B. Méthodes basées sur l'apprentissage automatique

Récemment, les études récurrentes dans la littérature pour l'amélioration du contrôle de la qualité sont celles traitants la détection d'anomalies à travers des méthodes d'apprentissage automatique non supervisées. Celles-ci consistent à identifier des comportements atypiques dans les données, tels que des dysfonctionnements de capteurs, des dérives de processus industriels ou encore des événements rares pouvant compromettre la qualité et la fiabilité de la production. Ces méthodes sont souvent étudiées dans des contextes où les inspections de qualité ne sont pas systématiquement réalisées à la fin de chaque étape de production, ce qui fait que les étiquettes de conformité ou de défaut sont généralement inexistantes. Cette situation reflète d'ailleurs la réalité courante des environnements industriels.

Ces méthodes se distinguent en deux approches, une qui repose sur la reconstruction des données normales et l'autre repose sur la définition des frontières de décision.

Méthodes de reconstruction

Les méthodes de reconstruction reposent sur l'idée qu'un modèle entraîné uniquement sur des données normales peut les reproduire correctement, mais échouera face à des observations atypiques (Huang et al., 2025). L'erreur de reconstruction entre les données originales et leur reproduction sert alors d'indicateur d'anomalie. Parmi les techniques récemment utilisées dans le secteur manufacturier, les auto-encodeurs profonds qui offrent une grande capacité de modélisation

de structures complexes (Givnan et al., 2022). Par exemple, dans une étude menée par Tsai et Jen (2021) un auto-encodeur convolutionnel régularisé a démontré une capacité à identifier les anomalies sur les surfaces texturés ou à motifs industriel. De même, Bergmann et al. (2019), à travers le benchmark MVTec-AD, ont démontré l'efficacité des auto-encodeurs et modèles de type VAE pour la détection des anomalies variées (rayures, contaminations, fissures) sur des images d'objets industriels. Pang et al. (2022) ont souligné à travers une revue sur les modèles de détection d'anomalies que les modèles de reconstruction basés sur des architectures profondes multi-échelles permettent non seulement de détecter les anomalies, mais aussi de les localiser avec précision grâce aux cartes d'erreur, ce qui constitue un atout pour le diagnostic en production. De leur côté Gao et al. (2023) ont proposé une approche pour l'identification de défauts de collage basée sur des représentations latentes apprises par un Stacked Denoising Autoencoder (SDAE), qui consiste à empiler plusieurs auto-encodeurs successifs afin d'extraire des caractéristiques hiérarchiques de plus en plus abstraites. Toutefois, ces méthodes peinent à généraliser à des anomalies hors distribution ou inédites (Schlüter et al., 2022). Ruff et al. (2021) évoquent aussi la nécessité d'un calibrage fin des seuils de détection étant donné que ces méthodes peuvent souffrir de la « sur-reconstruction » de certains défauts subtils, ce qui risque de masquer leur présence.

Méthodes basées sur la définition de frontière

Les méthodes de frontière définissent une région de normalité dans l'espace des variables et considèrent comme anomalies les points qui s'en écartent (Khan & Madden, 2014). Dans le secteur manufacturier, ces méthodes ont été mises en œuvre principalement à travers le One-Class SVM (OCSVM) et l'Isolation Forest (iForest). Dans le cas des robots industriels, l'isolation Forest a été évalué auprès de SVM et DBSCAN et a montré une robustesse pour la détection des anomalies collectives à grande échelle (Lu et al. 2022). Dans une étude de surveillance des anomalies dans des systèmes industriels multivariables comme les réacteurs et les réservoirs, Zope et al. (2019) ont aussi comparé OCSVM et iForest à des approches statistiques et profondes. Les résultats ont montré que ces deux modèles ont des performances comparables aux modèles profonds tout en restant plus légers. Mattera et al. (2025) ont souligné leur efficacité dans l'analyse de données multisensorielles issues du smart manufacturing, à condition de recourir à une réduction de dimension préalable. Néanmoins, Lu (2025) montre que l'intégration en production de ces

méthodes reste contraignante du fait de la nécessité d'un recalibrage régulier et de leur dépendance à la nature des anomalies.

2.2.2 Approches de détection de défauts dans le secteur manufacturier

Dans le secteur manufacturier, la détection de défauts a principalement été abordée à travers des approches de classification. La littérature distingue à cet égard les méthodes classiques, reposant sur la détection manuelle ou visuelle, et les approches récentes, fondées sur l'apprentissage automatique, visant à automatiser et à optimiser les processus d'inspection.

A. Méthodes manuelles et visuelles

Dans plusieurs secteurs notamment automobile, pharmaceutique, électronique, la conformité du produit final est encore très largement validée par inspection visuelle humaine. Dans ce cadre, des inspecteurs examinent les pièces ou les sous-assemblages pour repérer des défauts visibles comme des rayures, bulles, zones de mauvaise adhésion, manque de matière, délamination, déformation et contamination de surface (Melchore, 2011). Kujawińska et Vogt (2015) ont montré que cette méthode reste ancrée dans les pratiques industrielles parce qu'elle est simple à déployer, ne nécessite pas toujours d'investissement matériel lourd, et mobilise l'expertise tacite des opérateurs, capables d'identifier des défauts subtils ou contextuels que le système automatisé ne connaît pas encore. Des études en inspection visuelle indiquent que la variabilité entre inspecteurs peut conduire à des incohérences de détection et à des « échappements qualité », c'est-à-dire des défauts qui passent malgré le contrôle. See (2012) à travers une revue montre que cette dépendance à l'humain transforme souvent l'inspection visuelle en goulot d'étranglement pour les lignes à haut débit, et limite l'échelle de la surveillance quand le volume de production augmente. Les analyses récentes comparant inspection manuelle et inspection automatisée soulignent notamment que l'erreur humaine s'accroît sous l'effet de la fatigue, et que la vitesse d'inspection humaine est difficilement compatible avec des cadences industrielles élevées.

B. Méthodes basées sur l'apprentissage automatique

Ces méthodes reposent sur des algorithmes capables d'apprendre à partir de variables explicatives afin de prédire l'appartenance d'une observation à la classe "conforme" ou "défectueuse". La littérature couvre un large éventail de modèles, allant des méthodes classiques aux modèles ensemblistes, tels que les forêts aléatoires et les algorithmes de boosting. Ces derniers combinent

plusieurs classifieurs de base afin de réduire l'erreur de prédiction et d'améliorer la performance globale.

Les approches traditionnelles, telles que la régression logistique, les Support Vector Machines (SVM), le k-Nearest Neighbour (k-NN) ou les arbres de décision, ont servi de premières bases grâce à leur simplicité et leur interprétabilité. Par exemple, Kausik et al. (2025) ont utilisé des SVMs ainsi que les arbres de décision dans le cadre de l'inspection qualité dans l'industrie textile, pour repérer les défauts de tissage susceptibles d'affecter la qualité des produits. Une étude récente, menée par Herzog et al. (2024), qui s'est concentrée à la surveillance de procédé en fabrication additive, vise à développer un cadre de détection automatique de défauts à partir des données issues du suivi en temps réel du processus. Les auteurs ont appliqué plusieurs approches classiques, dont le Decision Tree (C5.0), le SVM afin de classifier les couches produites lors du procédé de fusion laser sur lit de poudre (Laser Powder Bed Fusion – LPBF) et d'identifier des défauts de fabrication tels que les pores, les manques de fusion ou les zones de surchauffe. Le SVM, entraîné avec un noyau radial de base (RBF), a montré des performances stables et précises, avec un F1-score de 0.89 et un AUC supérieur à 0.9, détectant particulièrement bien les défauts localisés associés à une signature thermique caractéristique. De son côté, le modèle d'arbre de décision C5.0, choisi pour sa structure simple et sa forte capacité d'interprétation, s'est révélé pertinent pour la détection de défaillances récurrentes et a permis de relier l'occurrence des défauts aux paramètres physiques du procédé.

Cependant ces méthodes montrent rapidement leurs limites face à un déséquilibre où la classe minoritaire représente moins de 20 % des observations et à la forte dimensionnalité des données industrielles (Fan et al., 2022; Mirzaei et al., 2023). Pour pallier ces faiblesses, les méthodes ensemblistes par bagging, comme les forêts aléatoires et les Extra Trees, ont été largement mobilisées afin de réduire la variance et accroître la robustesse (He et al., 2023; Muntasir Nishat et al., 2022). Dans le cas de l'inspection de soudure sur PCB et de la détection de défauts de roulements, He et al. (2023) montrent que les forêts aléatoires équilibrées atteignent un G-mean supérieur à 0.85, alors que les modèles simples présentent un rappel inférieur à 0.4. De même, Muntasir Nishat et al. (2022) constatent que le Random Forest, couplé à SMOTE-ENN et à l'optimisation d'hyperparamètres, surpasse l'ensemble des autres classifieurs avec une précision de 0.9 sur un jeu de données médical déséquilibré. D'autres méthodes de boosting, notamment AdaBoost, Gradient Boosting, XGBoost et LightGBM, se sont imposées récemment comme la

référence pour améliorer le rappel, le F1-score et le G-mean sur la classe minoritaire, en particulier lorsqu'elles sont associées à un ratio de déséquilibre de 1:10 de données soit 9% d'observations représentant des défauts (de Giorgio et al. 2023). Malgré leurs performances, les méthodes ensemblistes présentent des limites. Imani et al. (2025) montrent que les forêts aléatoires peuvent rester biaisées face aux déséquilibres sévères. De même, Zhang et al. (2022) soulignent que la nature séquentielle des méthodes de boosting, les rend plus sensibles au bruit et aux valeurs aberrantes dans les données.

La suite de cette section examine en détail le fonctionnement de AdaBoost et son utilisation dans un contexte de déséquilibre de classes. Ce modèle a été retenu dans le cadre de ce travail en raison de sa capacité intrinsèque à pondérer les instances difficiles à classer, ce qui lui permet de mieux gérer les déséquilibres de données. De plus, sa robustesse face à la forte dimensionnalité et sa capacité de généralisation aux défauts hors distribution en font une approche qui peut être adaptée aux problématiques industrielles de détection de défauts (de Giorgio et al. 2023).

Explication du processus de fonctionnement du modèle AdaBoost

Le modèle de classification AdaBoost (Adaptive Boosting) est une méthode d'apprentissage ensembliste de la famille du boosting. L'algorithme, introduit par Schapire (2003), fonctionne de manière itérative en construisant séquentiellement des classifieurs faibles, où chaque nouveau classifieur est conçu pour corriger les erreurs de classification commises par les précédents. Cette approche vise à améliorer progressivement la performance prédictive du modèle global.

Nous supposons que l'ensemble de données total contient n échantillons, répartis entre n_{train}, correspondant au nombre d'échantillons dédiés à l'apprentissage du modèle, et n_{test}, correspondant au nombre d'échantillons réservés au test, où $n = n_{train} + n_{test}$. On considère l'ensemble d'apprentissage $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\}$, où x_i désigne le vecteur de caractéristiques associé au *i-ème* échantillon, et $y_i \in \{0,1\}$ représente la variable cible, où 0 pour défectueux et 1 pour conforme. Le processus de construction du modèle débute par une initialisation des poids des échantillons, puis se poursuit par des étapes successives d'ajustement des classifieurs faibles.

Initialisation des poids des échantillons

La première étape de l'algorithme consiste à attribuer à chaque échantillon de l'ensemble d'apprentissage un poids uniforme. Cette pondération uniforme, notée $w_i^{(1)}$ pour l'échantillon i à

la première itération, garantit qu'aucune observation n'est privilégiée au démarrage du processus. La valeur attribuée à chaque poids est définie comme suit :

$$w_i^{(1)} = \frac{1}{n_{train}}, i = 1, \dots, n_{train} \quad (2.1)$$

Amélioration itérative du modèle

Cette étape présente le cœur du processus d'apprentissage, où le modèle ajoute de manière itérative de nouveaux apprenants faibles pour affiner ses prédictions. Pour *m* = 1, 2, ..., *M*, où *M* est le nombre total d'apprenants faibles, les étapes suivantes sont répétées.

Tout d'abord, un classificateur $h_m(x)$ est entraîné sur les données d'apprentissage en tenant compte de la distribution des poids $\left\{w_i^{(m)}\right\} ou \sum_{i=1}^{n\_train} w_i = 1$. L'objectif est de déterminer le classifieur qui minimise l'erreur de classification pondérée, notée ε_m . Cette erreur est calculée comme la somme des poids des échantillons mal classés :

$$\varepsilon_m = \sum_{i=1}^{n_{train}} w_i^{(m)} I\{h_m(x_i) \neq y_i\} \quad (2.2)$$

où *I* désigne la fonction indicatrice, égale à 1 si la prédiction est incorrecte et 0 sinon.

Ensuite, l'importance de ce classifieur,$\propto_m$ est déterminée. Ce coefficient reflète la performance de l'apprenant faible, un classifieur plus précis aura un poids plus important dans la décision finale. Il est calculé comme suit :

$$\propto_m = \frac{1}{2} \ln\left(\frac{1-\varepsilon_m}{\varepsilon_m}\right) \quad (2.3)$$

Cette valeur est positive tant que l'erreur $\varepsilon_m < 0.5$, ce qui signifie que le classifieur est plus performant qu'une classification aléatoire.

Enfin, les poids des échantillons sont mis à jour pour préparer l'itération suivante, *m+1*. Le poids des échantillons mal classés est augmenté, forçant le prochain classifieur à se concentrer sur ces cas difficiles, tandis que le poids des échantillons bien classés est diminué. La formule de mise à jour est la suivante :

$$w_i^{(m+1)} = \frac{w_i^{(m)}}{Z_m} \exp(- \propto_m y_i h_m(x_i)) \tag{2.4}$$

où Z_m est un facteur de normalisation qui assure que la somme des nouveaux poids reste égale à 1. Cette actualisation adaptative des poids est la caractéristique fondamentale de l'algorithme AdaBoost.

Prédiction de l'état du défaut

Après *M* itérations, le modèle AdaBoost final, $H(x)$, est obtenu en agrégeant les prédictions de tous les apprenants faibles, pondérées par leur importance respective $\propto_m$. Pour un nouvel échantillon *x*, la prédiction est calculée en appliquant la fonction signe à cette somme pondérée :

$$H(x) = sign\left(\sum_{m=1}^{M} \propto_m \ h_m(x)\right) \tag{2.5}$$

La sortie de la fonction *sign(.)* est soit 1 (pas de défauts), soit 0 (défaut détecté), fournissant ainsi la classification finale.

Durant l'apprentissage, des techniques de régularisation sont appliquées afin de prévenir le surajustement. Dans le cadre d'AdaBoost, le nombre total d'itérations, *M*, agit comme un paramètre de régularisation essentiel. Un nombre trop élevé d'apprenants faibles peut conduire le modèle à mémoriser le bruit présent dans les données d'entraînement, dégradant ainsi sa capacité de généralisation sur de nouvelles données. Le processus de boosting est donc souvent interrompu après un nombre d'itérations optimal, déterminé par validation croisée, ou dès que les performances sur un ensemble de validation cessent de s'améliorer (arrêt anticipé ou *early stopping*). De plus, la complexité des apprenants faibles eux-mêmes (par exemple, la profondeur des arbres de décision) constitue une autre forme de régularisation.

2.2.3 Méthodes d'équilibrage de données

Dans cette section, nous abordons la manière dont la problématique du déséquilibre entre les classes de données a été traitée dans la littérature relative au secteur manufacturier. En examinant les travaux consacrés au traitement de cette problématique, deux grandes approches se distinguent.

La première regroupe les méthodes classiques de rééchantillonnage des données, tandis que la seconde, plus récente, s'appuie sur l'utilisation de modèles génératifs pour la création de données synthétiques, approche qui connaît un intérêt croissant dans de nombreuses études.

A. Méthodes classiques de rééchantillonnage de données

Les données industrielles de qualité se caractérisent aujourd'hui par un fort déséquilibre entre classes, ce qui constitue un défi majeur pour les approches d'apprentissage automatique. La littérature souligne que ce déséquilibre compromet la performance des modèles, en particulier leur capacité à détecter les classes minoritaires associées aux défauts (de Giorgio et al. 2023). Pour y remédier, plusieurs méthodes d'échantillonnage ont été proposées. Ces techniques peuvent être regroupées en trois classes, le suréchantillonnage des classes rares, le sous-échantillonnage des classes majoritaires, ainsi que des approches hybrides qui combinent ces deux stratégies.

Méthodes de suréchantillonnage

Dans la littérature, une des méthodes classiques utilisée est le suréchantillonnage aléatoire, qui augmente la proportion de la classe minoritaire en dupliquant simplement les exemples existants (Yang et al., 2024). Cependant, cette duplication systématique peut entraîner un surapprentissage, car les modèles risquent de mémoriser des instances redondantes plutôt que d'apprendre des caractéristiques discriminantes. En réponse, des techniques d'interpolation synthétique ont été proposées, notamment SMOTE (Synthetic Minority Over-Sampling Technique). Ces méthodes génèrent de nouveaux exemples minoritaires en interpolant entre points proches (Beckmann et al., 2015). A cet égard, plusieurs études utilisent le synthetic Minority Over-sampling (SMOTE) sur des jeux de données industrielles ont montré que la génération artificielle d'exemples de la classe minoritaire améliore significativement le rappel (recall) des modèles de classification en présence de défauts rares (Han et al., 2024; Khosravi et al., 2024; Matharaarachchi et al., 2024). Toutefois, ce type de données synthétiques peut introduire du bruit et créer des observations peu représentatives de la réalité, ce qui limite la robustesse des modèles. Pour pallier cette difficulté, des variantes comme Borderline-SMOTE ou ADASYN ont été proposées (Han et al., 2024), en ciblant davantage les zones proches des frontières de décision ou en adaptant la génération aux exemples difficiles, permettant ainsi une meilleure séparation entre produits conformes et défectueux. Néanmoins ces stratégies peuvent involontairement amplifier l'impact d'instances bruitées ou mal étiquetées situées à proximité des frontières de décision (Seol et al., 2023). Ces

limites ont orienté les chercheurs vers des approches de rééchantillonnage alternatives, telles que le sous- échantillonnage.

Méthodes de sous-échantillonnage

Les méthodes de sous-échantillonnage constituent une autre approche pour traiter le déséquilibre des classes, en réduisant la taille de la classe majoritaire afin de l'aligner sur celle de la classe minoritaire. La méthode la plus simple est le *Random Undersampling (RUS)*, qui supprime aléatoirement des instances majoritaires. Toutefois, cette approche entraîne fréquemment une perte d'information substantielle comme l'ont montré des études menées sur des systèmes de freinage à haute vitesse (Wang & Liu, 2021). Pour remédier à cette contrainte, des méthodes basées sur le clustering, telles que celles utilisant *K-means* ou *DBSCAN*, visent à conserver des échantillons représentatifs de la classe majoritaire. Park et al. (2024) proposent ainsi une approche intégrant la réduction dimensionnelle par l'Analyse des Composantes Principales (PCA), suivie d'un couplage de *K-means*, *DBSCAN* et *Outliers Based Neighbourhood (OBN)* pour éliminer les instances bruitées ou chevauchantes. Néanmoins, ces techniques nécessitent un réglage précis des paramètres et peuvent éliminer des instances majoritaires pertinentes, réduisant la compréhension des conditions normales de fonctionnement. Enfin, des méthodes centrées sur les frontières, telles que *Tomek Links* (Harandi et al., 2023) et *Edited Nearest Neighbors (ENN)* (Muntasir Nishat et al., 2022), cherchent à nettoyer les frontières de décision, mais leur application reste délicate dans les environnements manufacturiers.

Méthodes hybrides

Les méthodes hybrides ont émergé comme une solution visant à atténuer les limites inhérentes aux approches de suréchantillonnage et de sous-échantillonnage en combinant les deux stratégies. Parmi elles, l'approche SMOTE-ENN s'est révélée particulièrement efficace dans des études liées à la maintenance prédictive. Par exemple, sur un jeu de données présentant 31 % d'échantillons défectueux, elle a atteint un rappel de 0.973 pour la classe minoritaire, démontrant sa capacité à identifier des cas rares de défaillances tout en réduisant le bruit provenant de la classe majoritaire (Muntasir Nishat et al., 2022). Toutefois, son efficacité reste conditionnée par la disponibilité d'un nombre suffisant d'instances défectueuses, car sa capacité à générer des échantillons synthétiques pertinents et à éliminer avec précision les données bruitées diminue lorsque les défauts sont extrêmement rares. Pour renforcer la robustesse face au bruit, Wei et al. (2024) ont proposé le

Noise-Immune Majority Weighted Minority Oversampling Technique (NI-MWMOTE), une approche itérative intégrant un mécanisme de suppression progressive du bruit. Dans le contexte spécifique des données industrielles temporelles, les défis liés aux frontières de décision ont également été abordés par la combinaison de SMOTE avec *Tomek Links*, permettant de réduire efficacement les chevauchements entre classes et d'améliorer la séparation entre produits conformes et défectueux (Harandi et al., 2023). Enfin, d'autres travaux ont montré que ces approches hybrides représentent une avancée significative, bien qu'elles demeurent sensibles aux choix de paramètres et à la distribution intrinsèque des données.

Bien que les méthodes mentionnées ci-dessus permettent de traiter le déséquilibre de classes, leurs limites face au bruit (Matharaarachchi et al., 2024), aux données mixtes (Chen et al., 2024) et aux contextes industriels dynamiques. C. Yang et al. (2024) soulignent la nécessité de solutions plus adaptatives et spécifiques au domaine. Dans les environnements manufacturiers, les données sont souvent hétérogènes, combinant à la fois des mesures continues issues de capteurs, des paramètres catégoriels et des variations temporelles. De plus, la rareté des motifs de défaut complique l'application directe des méthodes génériques d'équilibrage, qui risquent soit de simplifier excessivement le problème, soit d'introduire des données synthétiques ne reflétant pas fidèlement la variabilité réelle des procédés. Ces défis ont motivé l'exploration de techniques plus avancées, parmi lesquelles les *Generative Adversarial Networks (GANs)* qui se distinguent par leur potentiel prometteur pour corriger le déséquilibre des classes dans des contextes industriels complexes (Liu et al. 2021; Lee et al. 2025; Lu et al. 2022).

B. Méthodes génératives de données

L'émergence des GANs a marqué une révolution dans le traitement des jeux de données déséquilibrés (Sauber-Cole & Khoshgoftaar, 2022). Contrairement aux techniques classiques de suréchantillonnage fondées sur l'interpolation, les GANs apprennent directement la distribution des données réelles. La présente section revient sur le mécanisme fondamental des GANs, puis discute des principales applications rapportées dans la littérature industrielle.

Architecture et mécanisme opérationnel des GANs

Les GANs reposent sur une architecture composée de deux réseaux neuronaux, le *générateur* (*Generator*) noté G et le *discriminateur* (*Discriminator*) noté D, engagés dans un processus compétitif (Goodfellow et al., 2014). Le générateur transforme un bruit aléatoire z, échantillonné

selon une distribution $p_z(z)$, en données synthétiques qui cherchent à imiter la classe minoritaire et le discriminateur *D* apprend à distinguer les échantillons réels, issus de la distribution $p_{data}(x)$, des échantillons synthétiques. Lors du processus le discriminateur est entraîné en premier lieu afin de maximiser la probabilité de classer correctement les exemples réels et synthétiques, ce qui se traduit par la fonction de perte L_D suivante :

$$L_D = -(E_{x\sim p_{data}}[logD(x)] + E_{x\sim p_z}[\log(1 - D(G(z))]) \quad (2.6)$$

De son côté, le générateur met à jour ses paramètres de manière à minimiser sa propre fonction de perte L_G, qui mesure sa capacité à "tromper" le discriminateur D :

$$L_G = -(E_{z\sim p_z}[logD(G(z)]) \quad (2.7)$$

Ce processus alterné se poursuit jusqu'à ce qu'un équilibre soit atteint, lorsque les données synthétiques deviennent indiscernables des données réelles.

Application des GANs dans les contextes manufacturiers déséquilibrés

Ces dernières années, la recherche académique et industrielle a largement investi l'utilisation des GANs pour traiter le déséquilibre de données en production, en raison de la difficulté à détecter des défauts rares ou des pannes d'équipement souvent sous-représentés. Plusieurs architectures adaptées ont été proposées. Lu et al. (2022) ont proposé le Mode Seeking GAN (MSGAN), une variante de réseau antagoniste génératif visant à atténuer le phénomène de *mode collapse*, c'est-à-dire la tendance du générateur à produire des sorties répétitives plutôt que diversifiées. Dans cette étude MSGAN introduit une fonction de perte générant des signaux variés. Le modèle a permis de synthétiser des signaux de courant dans des moteurs robotiques, améliorant ainsi la détection de pannes, mais sa performance reste conditionnée par la diversité des données d'apprentissage et il demeure sensible à des sorties répétitives lorsque la distribution est trop complexe. De leur côté, les architectures reposant sur le Wasserstein GAN avec pénalité de gradient (WGAN-GP) se sont révélées efficaces pour stabiliser l'entraînement, comme l'ont montré des travaux sur la détection de défectuosité dans des compresseurs d'air (Shi, 2023), en limitant la sur confiance du discriminateur. Cependant, leur coût computationnel élevé et les difficultés de paramétrage constituent encore un frein à leur adoption en temps réel. Le DB-CGAN, qui reprend WGAN-GP en y ajoutant des mécanismes de correction de biais de distribution (Zhou et al., 2023), a renforcé

la robustesse face aux déséquilibres sévères, mais cette amélioration se fait au prix d'une complexité accrue et d'une instabilité possible lors de l'apprentissage. D'autres recherches ont combiné les GANs avec des modèles séquentiels, tels que des réseaux récurrents comme RNN, LSTM ou des autoencodeurs variationnels (VAE), afin de générer des séries temporelles réalistes. Dans ce contexte, une approche hybride a permis une amélioration de 7,4 % de la reconnaissance des défauts minoritaires (Shen et al., 2023), mais elle reste limitée par la nécessité de grands volumes de données temporelles de qualité et par la difficulté de généralisation à des environnements industriels très hétérogènes. Enfin, pour pallier l'incapacité des GANs classiques à modéliser correctement les données tabulaires mêlant variables numériques et catégorielles, Xu et al. (2019) ont proposé le Conditional Tabular GAN (CTGAN), qui conditionne la génération sur certaines variables et utilise des transformations adaptées aux distributions asymétriques, permettant de produire des données tabulaires synthétiques plus représentatives pour les procédés industriels. Toutefois, malgré ces avancées, CTGAN, requiert un volume important de données normales pour apprendre efficacement ainsi qu'une haute qualité de données, ne présentant pas de bruit ou de corrélations complexes.

Fonctionnement du modèle génératif CTGAN

Comme ce travail utilise CTGAN, notre attention se porte sur ce modèle génératif. Ce modèle conçu par Xu et al. (2019) pour pallier les limites des GANs conventionnels dans le traitement des données tabulaires mixtes. CTGAN introduit un vecteur conditionnel, appliqué à la fois au générateur et au discriminateur. Ce vecteur est construit à partir des variables catégorielles encodées en *one-hot*, garantissant que les échantillons synthétiques respectent les catégories prédéfinies. Durant l'entraînement, une variable catégorielle est sélectionnée de manière aléatoire et l'une de ses modalités est échantillonnée selon une distribution logarithmique des fréquences, ce qui assure une représentation équilibrée des catégories rares et fréquentes. Un vecteur binaire de masquage est ensuite utilisé pour indiquer la modalité sélectionnée, et l'ensemble des masques relatifs aux variables catégorielles est concaténé pour former le vecteur conditionnel. Ce dernier est fourni simultanément au générateur et au discriminateur. Le générateur est pénalisé lorsqu'il ne parvient pas à produire des échantillons correspondant à la modalité imposée, au moyen d'une fonction de perte d'entropie croisée l_{cond} appliquée entre la catégorie générée et le masque conditionnel, ce qui l'oblige à respecter la contrainte et garantit que les échantillons synthétiques produits correspondent à la catégorie souhaitée.

$$l_{cond} = -\sum_{i=0}^{K} y_i \log(\widehat{p_i}) \tag{2.7}$$

où :

K: nombre total de catégories possibles pour la variable considérée ;

y_i: valeur réelle (ou cible) de la i^e catégorie dans le masque conditionnel (1 si la catégorie est sélectionnée, 0 sinon) ;

$\widehat{p_i}$: probabilité prédite par le générateur que la catégorie i soit celle imposée.

Bien que cet apprentissage conditionnel permette de traiter efficacement les variables catégorielles, les variables numériques nécessitent une approche distincte afin de capturer leurs distributions souvent complexes et multimodales. Pour ces colonnes numériques, CTGAN recourt à une normalisation spécifique aux modes : un modèle de mélange gaussien variationnel (*Variational Gaussian Mixture*, VGM) identifie les modes distincts de la distribution. Chaque valeur est ensuite normalisée en fonction du mode auquel elle est associée, à l'aide de la moyenne et de l'écart-type de ce mode. Cette opération permet de borner les entrées et d'éviter les instabilités de gradient. Lors de la génération, le modèle produit à la fois un scalaire normalisé et un vecteur *one-hot* indiquant le mode associé, qui sont ensuite dénormalisés pour reconstruire des valeurs réalistes. Cette représentation duale préserve la structure multimodale des données, tandis que l'entraînement adversarial garantit que les valeurs synthétiques restent alignées avec la distribution d'origine. En modélisant explicitement les modes, CTGAN évite le phénomène de *mode collapse* et capture fidèlement des distributions non gaussiennes, y compris des formes asymétriques ou multi-pics.

2.2.4 Méthodes d'explicabilité des modèles

Les méthodes d'explicabilité des modèles visent à fournir des explications compréhensibles sur le fonctionnement interne des modèles d'apprentissage automatique ou sur les raisons ayant conduit à une décision particulière. La littérature distingue généralement deux formes d'explicabilité, l'explicabilité globale, qui cherche à expliquer le comportement général du modèle en mettant en évidence les variables les plus influentes et leur impact sur les prédictions, et l'explicabilité locale, qui s'attache à comprendre et justifier les décisions individuelles prises par le modèle pour une observation donnée.

A. Méthodes globales d'explicabilité

Les méthodes globales visent à fournir une compréhension d'ensemble du comportement d'un modèle d'apprentissage automatique, en identifiant les variables qui influencent le plus ses décisions et en analysant leur contribution moyenne sur l'ensemble du jeu de données. Parmi les approches les plus répandues, l'importance des variables (*feature importance*) issue des forêts aléatoires ou des modèles de gradient boosting permet de quantifier la contribution relative de chaque attribut à la prédiction globale. Dans le domaine de la maintenance prédictive, Brusa et al. (2023) ont utilisé l'importance des variables pour identifier les capteurs les plus critiques dans la détection de défaillances de machines industrielles, facilitant ainsi l'optimisation des opérations de maintenance. D'autres outils comme les Partial Dependence Plots (PDP) et leurs extensions, telles que les Accumulated Local Effects (ALE), permettent de visualiser les relations entre variables d'entrée et sortie prédite. Ces techniques ont été appliquées dans l'industrie chimique par Molnar (2019) pour comprendre l'influence de paramètres tels que la pression ou la température sur la probabilité de défauts dans les procédés continus. Ces méthodes globales sont particulièrement utiles pour hiérarchiser les facteurs de risque et orienter les experts vers les variables critiques du procédé. Toutefois, dans la détection d'anomalies, où les instances sont rares et souvent très différentes des données normales, une compréhension globale peut ne pas suffire à expliquer pourquoi une instance spécifique est classée comme anormale, nécessitant alors des approches d'explicabilité locale (Oliveira et al., 2022). Ces limitations soulignent la nécessité d'une analyse critique et d'une complémentarité avec d'autres méthodes pour une compréhension exhaustive du comportement du modèle.

B. Méthodes locales d'explicabilité

À l'inverse, les méthodes locales se concentrent sur l'explication d'une prédiction individuelle, en cherchant à comprendre pourquoi un modèle a classé une observation particulière comme normale ou défectueuse. Parmi les approches les plus connues, la méthode *LIME* (*Local Interpretable Model-agnostic Explanations*) approxime localement le modèle complexe par un modèle simple et interprétable, afin de mettre en évidence les variables ayant le plus contribué à une prédiction spécifique. Cette méthode a été appliquée dans le secteur de la cybersécurité industrielle par Ribeiro et al. (2016) pour expliquer les décisions de modèles détectant des intrusions dans des systèmes de contrôle industriels (ICS). De son côté, *SHAP* (*SHapley Additive exPlanations*),

inspiré de la théorie des jeux coopératifs, attribue à chaque variable une contribution additive équitable pour une instance donnée. Dans le cadre de la surveillance de procédés industriels, Lundberg et Lee (2017) ont utilisé SHAP pour expliquer les prédictions de défauts dans des capteurs de turbines à gaz, permettant aux ingénieurs d'identifier les mesures responsables des alertes. Dans une revue systématique récente, Cação et al. (2025) démontrent que SHAP, combiné à des techniques de clustering, permet de localiser les signaux anormaux et de regrouper efficacement les anomalies dans des contextes industriels complexes.

Malgré leur intérêt, ces méthodes demeurent sensibles au bruit et aux corrélations entre variables, ce qui peut compromettre la fiabilité des explications fournies et freiner leur déploiement en temps réel (Cação et al., 2025).

2.3 Synthèse critique

Les travaux présentés ci-dessus mettent en lumière les différentes méthodes visant à améliorer le contrôle de la qualité en temps réel dans l'industrie manufacturière. Cependant, une analyse critique de la littérature révèle plusieurs limites majeures qui freinent leur adoption dans les environnements industriels actuels. Le contrôle de la qualité a été envisagé sous deux perspectives. La première approche du contrôle de la qualité consiste à identifier les dérives au sein des procédés industriels, ce qui correspond aux méthodes de détection d'anomalies. Historiquement, ces approches reposent sur les méthodes statistiques de contrôle de processus (SPC) (Shewart 1931), déclinées selon deux principales catégories : les approches univariées, centrées sur le suivi d'une seule variable, et les approches multivariées, qui tiennent compte simultanément de plusieurs paramètres de procédé corrélés. Ces outils, tels que les cartes de contrôle de Shewhart ou les cartes de Hotelling T^2 (Hotelling, 1947), sont encore largement utilisés dans divers secteurs industriels, notamment l'automobile, le pharmaceutique et l'agroalimentaire. Néanmoins, leur mise en œuvre suppose le respect d'hypothèses statistiques strictes, telles que la normalité des distributions (Bersimis et al., 2006) ou la faible variabilité des données (Ramos et al., 2021), conditions rarement vérifiées dans les environnements de production réels. C'est pourquoi des approches basées sur les données, majoritairement non supervisées, ont été introduites. Celles-ci reposent sur la modélisation du comportement normal du procédé afin de détecter toute déviation significative, et se révèlent particulièrement adaptées à la surveillance en continu des processus complexes, où les conditions d'exploitation évoluent dynamiquement

(Givnan et al., 2022). Parmi ces approches, on retrouve notamment les Autoencodeurs, l'Isolation Forest (iForest) et le One-Class SVM (OCSVM), qui se sont révélés efficaces pour modéliser des relations non linéaires et capturer la structure intrinsèque des données de procédé (Gao et al., 2023; Lu et al., 2022b; Tsai & Jen, 2021; Zope et al., 2019). Toutefois, ces méthodes nécessitent un calibrage précis des seuils de détection et présentent une faible capacité de généralisation face aux anomalies hors distribution (Lu, 2025; Ruff et al., 2021; Schlüter et al., 2022).

En complément de cette perspective centrée sur la surveillance de procédé, une seconde approche du contrôle de la qualité s'attache à la détection directe des défauts sur les produits finis. Contrairement aux méthodes de détection d'anomalies, ces approches reposent généralement sur des modèles supervisés de classification binaire, visant à distinguer les produits conformes des produits défectueux à partir de données étiquetées issues d'inspections ou de capteurs. Les modèles classiques tels que le Support Vector Machine (SVM), les arbres de décision ou encore les k-plus proches voisins (k-NN) ont été largement mobilisés comme premières solutions de référence dans ce contexte (Herzog et al., 2024; Kausik et al., 2025b). Toutefois, ces approches présentent certaines limites de généralisabilité lorsque les défauts sont rares ou peu représentés dans les jeux de données industriels volumineux (Fan et al., 2022; Mirzaei et al., 2023). Pour pallier ces contraintes, la littérature récente s'est orientée vers des méthodes ensemblistes, telles qu'AdaBoost ou Random Forest, qui offrent une meilleure robustesse face au déséquilibre des classes et une capacité de modélisation de la variabilité des procédés et de grands volumes de données (de Giorgio et al., 2023). Cependant, il a également été constaté que dans des contextes de déséquilibre extrême, tels que ceux rencontrés dans les industries manufacturières modernes visant le zéro défaut, les méthodes ensemblistes peuvent perdre en performance (Imani et al., 2025). En effet, lorsque la proportion de produits défectueux représente moins de 5 % des observations, ces modèles sont moins efficaces. Face à cette limitation, plusieurs auteurs ont tenté d'améliorer la représentativité des classes rares en recourant à des techniques de rééchantillonnage. Les techniques de rééchantillonnage classiques, comme SMOTE et ses variantes, bien qu'elles améliorent la détection des cas rares, tendent à générer des données synthétiques bruitées ou irréalistes (Han et al., 2005 ; He et al., 2008), et restent limitées face aux données hétérogènes typiques de la production, où variables numériques et catégorielles coexistent. Dans la continuité de ces travaux, les modèles génératifs tels que les GAN et leurs extensions récentes comme WGAN-GP, MSGAN et CTGAN ont été proposés. Ces modèles offrent un potentiel prometteur

pour pallier la rareté des défauts, mais demeurent coûteux en calcul, instables à l'entraînement et difficiles à déployer en temps réel (Lu et al., 2019 ; Xu et al., 2019). Une autre limite recensée de la littérature est que même si plusieurs solutions avancées démontrent une efficacité dans des environnements expérimentaux contrôlés, elles présentent des barrières importantes à l'industrialisation. En effet, elles reposent sur des pipelines computationnellement lourds, incompatibles avec les contraintes de rapidité et de fiabilité du milieu manufacturier (He & Garcia, 2009 ; Khosravi et al., 2023). Enfin, même lorsque les performances de détection sont satisfaisantes, le manque d'interprétabilité reste un obstacle majeur à leur adoption : de nombreux modèles se limitent à signaler l'occurrence d'un défaut sans fournir d'explications sur ses causes potentielles, ce qui empêche les praticiens d'agir de manière éclairée (Ribeiro et al., 2016 ; Lundberg & Lee, 2017 ; Cação et al., 2025). Ces limites soulignent la nécessité de concevoir des approches plus robustes, capables de gérer des défauts extrêmement rares (<5%), de modéliser des données hétérogènes, de fonctionner en temps réel et de fournir des explications fiables. Ces conditions sont indispensables pour favoriser une adoption réelle dans l'industrie manufacturière.

CHAPITRE 3 MÉTHODOLOGIE DE RECHERCHE

Ce chapitre a pour objectif de présenter la méthodologie de recherche adoptée dans le cadre de ce mémoire. Il débute par l'énoncé des objectifs de recherche définis pour ce travail, puis expose en détail la démarche méthodologique mise en œuvre pour les atteindre.

3.1 Objectifs de recherche

L'objectif général de cette recherche **est d'améliorer la capacité et l'efficacité du processus de contrôle qualité en temps réel au sein d'une ligne d'assemblage semi-automatique**. Pour atteindre cet objectif, une méthodologie structurée et basée sur les données est adoptée. Cette approche décompose l'objectif principal en deux objectifs spécifiques interdépendants, permettant de relever méthodiquement les défis inhérents à la détection de défauts en milieu industriel. Les objectifs spécifiques sont les suivants :

SO1 : Proposer un modèle d'indentification de défauts en temps réel adapté au contexte de données fortement débalancées : Cet objectif consiste à construire un classificateur robuste pour l'identification automatique des défauts de produits finis. Le processus inclut la collecte et le prétraitement des données de production, suivis d'une analyse comparative de plusieurs algorithmes de classification. La performance des modèles sera évaluée à l'aide de métriques spécifiques au contexte de données débalancées afin de sélectionner l'architecture la plus performante.

SO2 : Identifier les facteurs d'influence de qualité : Cet objectif vise à comprendre les facteurs associés à l'apparition des défauts en s'appuyant sur l'interprétation des décisions du modèle.

3.2 Méthodologie générale

La méthodologie adoptée dans ce travail s'inspire du cadre *Cross Industry Standard Process for Data Mining (CRISP-DM*), initialement proposé dans les années 1990 sous l'impulsion d'IBM et largement reconnu depuis comme la norme de facto pour la conduite de projets en sciences de données. CRISP-DM définit un cycle de vie structuré et itératif composé de six phases interdépendantes (Wirth & Hipp, 2000). Le processus débute par une phase de compréhension du métier (*business understanding*), visant à définir les objectifs et les contraintes propres au contexte industriel étudié. Elle est suivie par une phase de compréhension des données (*data understanding*),

qui consiste à collecter, explorer et analyser les données disponibles. La troisième étape correspond à la préparation des données (*data preparation*), où celles-ci sont nettoyées, transformées et organisées afin de les rendre exploitables pour l'étape suivante. La phase de modélisation (*modeling*) mobilise ensuite des algorithmes appropriés pour construire des modèles prédictifs adaptés à la problématique. Vient ensuite la phase d'évaluation (*evaluation*), qui permet de juger de la robustesse des modèles obtenus au regard des objectifs métier. Enfin, la dernière étape est celle du déploiement (*deployment*), où les modèles validés sont intégrés dans l'environnement opérationnel de l'organisation. Dans le cadre de ce travail, l'ensemble des phases définies par la méthodologie CRISP-DM ont été retenues, à l'exception de la phase finale de déploiement qui demeure en cours d'exécution. Toutefois, afin d'assurer une meilleure adéquation avec la problématique étudiée et les spécificités des données disponibles, chacune des étapes sera détaillée et contextualisée. Cette démarche permettra de mettre en évidence les choix méthodologiques effectués et de souligner les contraintes propres au jeu de données ainsi que les adaptations mises en œuvre pour répondre aux objectifs de la recherche.

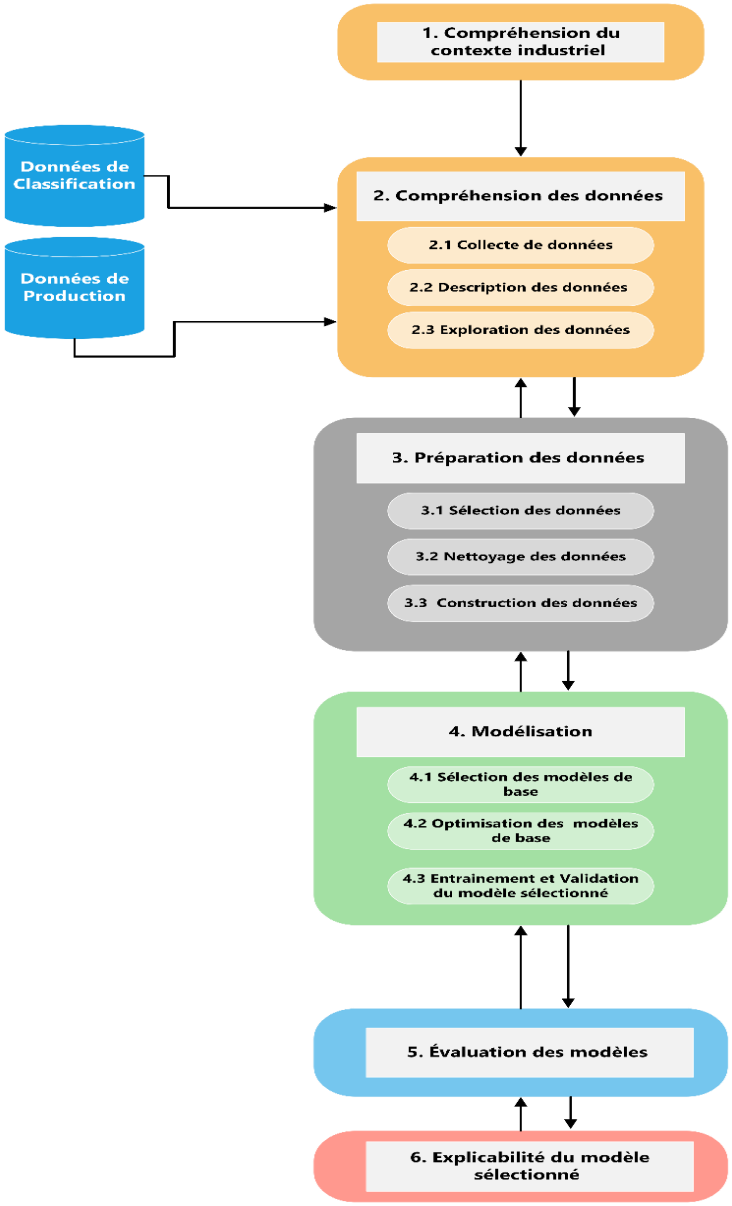


Figure 3.1 Méthodologie générale de recherche

Ce travail débute par une compréhension du contexte industriel, suivie de la collecte des données, incluant à la fois celles relatives au processus de production et celles associées à qualité en sortie. L'étape d'après consiste en une compréhension approfondie des données. Celle-ci vise à confronter la réalité du contexte d'application aux informations extraites, à travers une description détaillée et une analyse exploratoire permettant d'identifier les caractéristiques principales et les éventuelles incohérences. La deuxième étape est celle de la préparation des données, qui englobe la sélection des variables pertinentes pour la détection des défauts, le nettoyage des données, ainsi que leur

enrichissement pour améliorer la représentativité et la performance des modèles. Enfin, la phase de modélisation consiste à développer et affiner un modèle prédictif capable non seulement d'identifier les défauts mais également de fournir des éléments d'explication quant aux facteurs qui contribuent à leur apparition. Les prochaines sous-sections présentent en détail chacune des étapes énoncées ci-dessus.

3.2.1 Compréhension du contexte industriel

La première étape méthodologique consiste à comprendre le contexte industriel dans lequel s'inscrit ce projet. Cette phase préliminaire vise à acquérir une connaissance du procédé de fabrication, de la nature des produits fabriqués ainsi que de l'organisation des opérations le long de la chaîne de production. Pour ce faire, il s'agir d'effectuer des visites industrielles sur le site de production, permettant d'observer les différentes étapes d'assemblage, les équipements utilisés et les conditions de fonctionnement des machines. Ces observations sont complétées par des discussions approfondies avec les acteurs clés du processus, notamment les opérateurs de ligne, les ingénieurs de procédés, ainsi que les experts en assurance qualité. Ces échanges permettent de clarifier les rôles et interactions entre les différentes unités de production, de mieux comprendre la logique de séquencement des opérations, ainsi que les sources potentielles susceptibles d'affecter la qualité finale du produit. Ils ont également permis d'identifier les paramètres critiques du procédé (*Critical Process Parameters, CPP*) et les indicateurs de performance qualité (*Critical Quality Attributes, CQA*). Ces paramètres constituent la base des variables étudiées dans les phases ultérieures de l'analyse de données.

3.2.2 Compréhension des données

Cette étape consiste à analyser les données disponibles en les replaçant dans le contexte opérationnel du processus industriel étudié. Il s'agit d'articuler les informations issues des bases de données avec la réalité du terrain afin de situer clairement la problématique et d'expliciter les attentes du projet. Elle permet également de faire un état des lieux de la situation initiale du partenaire industriel, en identifiant ses priorités, ses contraintes ainsi que les difficultés anticipées. Une telle démarche permet de définir des objectifs cohérents et adaptés aux enjeux spécifiques de la détection de défauts dans un environnement manufacturier semi-automatisé.

A. Collecte des données

Cette étape a pour objectif d'identifier les sources de données ainsi que les modalités de leur collecte. Dans un contexte industriel, les données peuvent provenir de différentes origines à savoir les mesures issues de capteurs, enregistrements automatiques des machines ou encore saisies manuelles. Une fois les sources identifiées, il est nécessaire de préciser le niveau de granularité de la collecte, qui peut s'effectuer au niveau du lot, de la pièce ou du cycle machine. La définition de ce cadre permet ensuite d'évaluer la quantité d'informations disponibles, la nature des variables collectées, la taille des tables de données et le type d'indications qu'elles véhiculent. Ces éléments sont essentiels pour déterminer l'horizon temporel couvert par la collecte et juger de la représentativité des données pour l'analyse. Dans le cadre de cette étude, centrée sur l'identification des défauts au sein du processus d'assemblage, deux tables principales sont mobilisées.

- La première correspond aux données de production, qui retracent l'ensemble des produits assemblés et renseignent sur les étapes successives du processus étudié.
- La seconde regroupe les données de classification des défauts, permettant de caractériser la présence éventuelle d'anomalies, d'en préciser la nature ainsi que la localisation.

B. Description des données

Une fois les données collectées, une étape essentielle consiste à en réaliser une description détaillée pour une meilleure compréhension. Cette phase implique d'examiner les relations entre les différentes tables disponibles et de caractériser les attributs qu'elles contiennent. L'analyse descriptive permet ainsi de préciser la taille des tables, la nature des variables collectées, ainsi que les différentes valeurs possibles pour chaque attribut. Elle met en évidence les informations utiles à l'étude, et les éventuelles lacunes, telles que des valeurs manquantes ou incomplètes. Chaque table est analysée de manière approfondie afin d'évaluer la capacité des variables à représenter adéquatement le processus de production et les événements de défauts associés. À ce stade, il devient possible d'identifier les éléments qui demeurent ambigus pour une compréhension complète du cas d'étude, ainsi que de vérifier si toutes les connaissances métiers nécessaires sont effectivement traduites en données exploitables. Cela permet d'affiner la compréhension du contexte avant d'entamer les phases ultérieures de préparation et de modélisation.

C. Exploration des données

L'exploration de données constitue la dernière étape de la compréhension de données et joue un rôle de guide pour orienter efficacement le processus de nettoyage. Cette phase s'articule autour de trois étapes principales : le profilage statistique, la visualisation des données et l'analyse des corrélations. Le profilage statistique repose sur l'utilisation de méthodes de statistique descriptive classique telles que la moyenne, la médiane, le nombre d'occurrences et l'écart-type. Ces indicateurs permettent de résumer les principales caractéristiques des données, comme les tendances centrales, la dispersion des variables numériques ainsi que les distributions de fréquences pour les variables catégorielles. La visualisation des données, à l'aide d'outils tels que les histogrammes, boîtes à moustaches (*box plots*), diagrammes circulaires (*pie charts*), graphiques en violon (*violin plots*) ou encore l'estimation de densité par noyaux (*Kernel Density Estimation, KDE*), facilite l'exploration des distributions et la détection d'éventuelles valeurs aberrantes. Les nuages de points (*scatter plots*) sont également utiles pour observer visuellement la présence d'anomalies. Enfin, l'analyse des corrélations, souvent réalisée à l'aide de cartes thermiques (*heatmaps*), permet d'identifier la multicolinéarité entre variables, de guider l'élimination d'attributs redondants et d'éclairer le processus de sélection des variables pertinentes. D'autres étapes peuvent s'ajouter en fonction de la nature et de la richesse des données disponibles.

3.2.3 Préparation des données

La phase de préparation des données constitue une étape cruciale du processus, souvent reconnue comme la plus chronophage. Elle a pour objectif de transformer les données brutes en un jeu exploitable et cohérent, adapté aux exigences du modèle à développer. Une préparation rigoureuse devrait à la fois optimiser les performances des algorithmes d'apprentissage, et garantir une meilleure capacité de généralisation. La méthodologie de préparation de données de ce travail est illustrée dans la Figure 3.2.

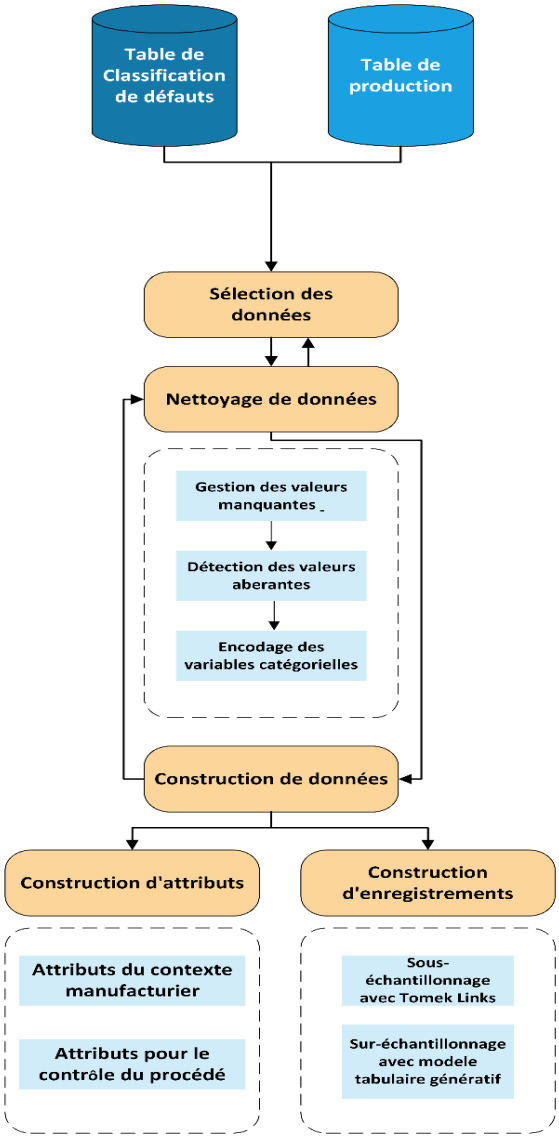


Figure 3.2 Méthodologie de préparation de données

A. Sélection des données

La sélection des données concerne la conservation des informations les plus pertinentes pour la détection des défauts d'assemblage. Cette étape repose sur une compréhension du cas d'étude et des données disponibles, menée en collaboration étroite avec les experts métier et les opérateurs de terrain. Elle vise à constituer un jeu de données cohérent, représentatif, et exploitable pour l'analyse et la modélisation.

Concernant les données de production, celles-ci doivent permettre de retracer l'assemblage de chaque produit de manière unique, en intégrant le type de produit, la machine et l'opérateur

responsables, ainsi que les instants de début et de fin d'assemblage. Des informations supplémentaires, telles que les consignes de production, les mesures réellement effectuées sur les composants et les valeurs de pression enregistrées, sont également prises en compte afin de vérifier la conformité des opérations. Ces variables fournissent une base essentielle pour comparer l'écart entre les valeurs prescrites et les valeurs observées, et peuvent être enrichies par d'autres attributs dans la perspective d'analyses complémentaires ou de la construction de nouvelles variables dérivées.

En ce qui concerne les données de classification, elles permettent d'identifier les produits défectueux et de préciser la nature des défauts détectés.

D'autres informations, telles que les ajustements réalisés par les opérateurs sur les paramètres du processus, comme les mesures ou les pressions, sont également intégrées afin de capturer les interactions homme–machine pouvant influencer l'apparition des défauts.

Enfin, l'intégration de ces sources de données est indispensable, puisqu'elle permet d'associer directement chaque défaut identifié aux produits correspondants. Cette opération d'appariement peut être réalisée dès la phase de collecte, lorsque les systèmes d'information le permettent, ou bien ultérieurement lors du processus de préparation des données.

B. Nettoyage des données

Le nettoyage des données constitue une étape essentielle pour garantir leur qualité. Il vise à corriger les erreurs de collecte ou de saisie de données et aussi à adapter leur format afin de limiter les biais susceptibles d'affecter les performances des modèles. Dans le cadre de cette méthodologie, trois opérations principales ont été menées : la gestion des valeurs manquantes, la détection des anomalies dans les données et l'encodage des variables catégorielles.

Gestion des valeurs manquantes

Les données manquantes constituent une problématique couramment rencontrée dans la majorité des jeux de données, en particulier dans les environnements de fabrication intelligente où des défaillances de capteurs, des retards de communication ou encore des erreurs de saisie manuelle peuvent survenir. Différentes stratégies peuvent être mobilisées pour y remédier. Le choix d'une méthode d'imputation appropriée doit être guidé par la distribution des données sous-jacentes, la nature de la variable concernée et dans certains cas, le contexte spécifique de l'attribut. Les

imputations statistiques univariées, telles que la moyenne, la médiane ou la valeur la plus fréquente, sont largement utilisées en raison de leur simplicité et de leur efficacité computationnelle. Toutefois, lorsque les données manquantes présentent des structures de dépendance complexes, des techniques d'imputation multivariées, telles que l'imputation par les *k plus proches voisins* (k-NN) ou l'imputation par régression, s'avèrent plus adaptées. L'imputation par k-NN exploite la similarité entre observations dans l'espace des variables et se montre particulièrement efficace lorsque les données présentent des corrélations locales ou des structures de regroupement. De leur côté, les méthodes d'imputation par régression utilisent les relations observées entre variables pour estimer les valeurs manquantes et conviennent lorsque des associations linéaires ou non linéaires existent.

Détection des valeurs aberrantes

Les valeurs aberrantes sont fréquemment présentes dans les données issues de capteurs, en raison de dysfonctionnements ponctuels ou de pannes. Ces observations atypiques peuvent avoir un impact disproportionné sur les analyses statistiques et les modèles d'apprentissage automatique, ce qui rend leur détection et leur traitement indispensables. Les méthodes statistiques classiques, telles que l'analyse du *z-score* (Yaro et al., 2023) ou les approches fondées sur des modèles statistiques (Hussain & Zhang, 2025), permettent d'identifier les valeurs aberrantes en quantifiant les écarts par rapport à la distribution attendue des données. Un critère couramment utilisé consiste à signaler comme valeurs aberrantes les points situés à plus de 1,5 fois l'intervalle interquartile (IQR) au-dessus du troisième quartile (Q3) ou en dessous du premier quartile (Q1). Dans un cadre multivarié, des méthodes basées sur des modèles peuvent être appliquées puisqu'elles prennent en compte les interdépendances entre variables, permettant ainsi de détecter des observations atypiques résultant de configurations particulières de valeurs conjointes.

Encodage des variables catégorielles

Les données industrielles comportent une proportion importante de variables catégorielles, telles que le type de produit, la machine utilisée ou encore l'opérateur associé au processus d'assemblage. Toutefois, ces variables ne peuvent être exploitées directement par la plupart des algorithmes d'apprentissage automatique. Donc la transformation des variables catégorielles en représentations numériques constitue une étape essentielle afin d'assurer leur compatibilité avec la majorité des algorithmes d'apprentissage automatique. Selon Poslavskaya et Korolev (2023), les techniques

d'encodage traditionnelles, telles que l'*one-hot encoding* ou le *label encoding*, sont utilisées notamment lorsque les variables catégorielles comportent un nombre limité de modalités distinctes et qu'aucune relation d'ordre n'existe entre elles. Toutefois, dans les situations où il est important de préserver la distribution des catégories, l'encodage basé sur les fréquences peut s'avérer plus adapté. Cette approche, sensible à la répartition des modalités, permet aux modèles de capturer les schémas dominants associés aux catégories les plus fréquentes tout en conservant les signaux informatifs provenant des classes sous-représentées.

C. Construction des données

La construction de données peut être envisagée sous deux formes : la construction d'attributs et la création d'enregistrements. La construction d'attributs (*Feature Enginneeing*) consiste à enrichir le jeu de données en dérivant de nouvelles variables à partir d'informations déjà présentes de manière implicite dans les tables existantes. Quant à la construction d'enregistrements, elle répond à la nécessité d'augmenter le volume de données disponibles, soit pour compenser une taille d'échantillon insuffisante lors de l'apprentissage du modèle, soit, comme c'est le cas dans notre méthodologie, pour pallier la rareté des défauts.

Construction d'attributs

À ce stade, la méthodologie proposée met l'accent sur le *feature engineering*. Ce processus peut inclure l'application de diverses transformations mathématiques, telles que les transformations logarithmiques ou de puissance, particulièrement utiles dans le cas de distributions fortement asymétriques, afin de stabiliser la variance et de limiter l'influence des valeurs extrêmes. Dans certains cas, il est également nécessaire d'intégrer des informations contextuelles liées au processus de production, à l'environnement opérationnel ou encore aux caractéristiques intrinsèques des produits. Ces variables peuvent inclure, par exemple, des agrégations temporelles permettant de capturer des régularités dans les opérations, ou des indicateurs dérivés des procédés établissant un lien entre les événements de production et les signaux issus des machines ou des capteurs. Sur la base de ces transformations, des attributs spécifiques au domaine manufacturier peuvent ainsi être construits. Deux approches complémentaires, présentées dans les sous-sections suivantes, sont proposées dans ce travail pour la détection de défauts.

Construction d'attributs liés au contexte manufacturier

Cette étape consiste à intégrer la connaissance métier relative à l'environnement de production dans le jeu de données, en créant des variables reflétant les événements opérationnels clés ainsi que leur dynamique temporelle. Dans les contextes manufacturiers, le comportement d'une ligne de production n'est pas uniquement déterminé par les mesures issues des capteurs en temps réel, mais également par des changements structurels tels que le remplacement d'équipements, la reconfiguration des lignes ou les interventions de maintenance. Afin de capturer ces effets, plusieurs types de variables peuvent être construites. Par exemple, des indicateurs binaires signalant si un remplacement ou une modification est survenu dans une fenêtre temporelle donnée précédant un lot de production, des variables de comptage enregistrant le nombre d'interventions ou de changements sur des horizons temporels glissants ou encore des variables mesurant le temps écoulé depuis la dernière intervention sur un équipement. Une telle ingénierie de variables contextualisée permet de mettre en évidence d'éventuelles associations entre les modifications opérationnelles et les déviations de qualité ou les variations de performance qui en résultent.

Construction d'attributs liés au contrôle statistique du procédé

Comme discuté dans la revue de littérature, l'instabilité du processus peut constituer un indicateur précurseur de l'apparition de défauts. L'ingénierie de variables fondée sur le SPC permet ainsi de traduire la stabilité du processus en indicateurs exploitables par les modèles prédictifs. Dans cette méthodologie, nous proposons d'intégrer les cartes de contrôle statistique dans l'ingénierie de variables en introduisant des indicateurs dérivés de ces cartes afin d'encoder les anomalies statistiques ou les motifs révélant une dérive du processus. Par exemple, des variables binaires peuvent être construites pour signaler si une mesure dépasse les limites de contrôle prédéfinies, indiquant ainsi une situation potentiellement hors contrôle. De même, la survenue de points consécutifs au-dessus ou au-dessous de la ligne de tolérance peut être traduite en variables reflétant des tendances subtiles et quantifiant leur écart par rapport au seuil de tolérance. L'intégration de ces signaux SPC dans l'ensemble de caractéristiques offre ainsi au modèle des représentations interprétables de la stabilité du processus. Bien que l'ingénierie de variables fondée sur le SPC ait déjà été utilisée pour le suivi de la stabilité et la détection de déviations dans les procédés industriels (Xia et al., 2021), son intégration dans une démarche de modélisation prédictive pour la détection de défauts demeure encore peu explorée.

Construction d'enregistrements

En détection de défauts, le déséquilibre des données constitue un problème très abordé, car les procédés industriels sont spécifiquement conçus pour réduire au maximum les occurrences de non-conformités. Cette rareté impose parfois la génération de nouveaux enregistrements afin de soutenir l'apprentissage des modèles prédictifs. Les techniques classiques de suréchantillonnage, telles que le *Random Oversampling (ROS)* ou les méthodes basées sur l'interpolation, sont couramment mobilisées pour atténuer ce problème. Toutefois, dans les scénarios réels, les observations rares sont fréquemment entremêlées aux données normales. Dans un cas pareil, les approches conventionnelles montrent leurs limites. Le ROS se contente de répliquer des exemples minoritaires sans résoudre le problème du chevauchement dans l'espace des caractéristiques, tandis que SMOTE repose sur l'hypothèse d'une géométrie convexe des classes, rarement vérifiée dans les données industrielles. Afin de dépasser ces contraintes, la méthodologie proposée introduit une approche hybride combinant l'échantillonnage par Tomek Links et la génération de données synthétiques par Conditional Tabular GAN (CTGAN).

La méthode des *Tomek Links* introduite par Tomek en 1976, vise à réduire l'ambiguïté aux frontières de décision en supprimant les instances bruitées de la classe majoritaire situées à proximité immédiate de la classe minoritaire. Plus précisément, pour chaque observation $X_i \in C_{min}$ (classe minoritaire), l'algorithme cherche son plus proche voisin $X_j \in C_{max}$ (classe majoritaire) à l'aide d'une mesure de distance $d(X_i, X_j)$. Lorsqu'un tel couple est identifié, l'instance majoritaire X_j est retirée du jeu de données. Ce processus permet de nettoyer l'espace des caractéristiques et d'améliorer la séparation entre classes.

En complément, **le CTGAN** est mobilisé comme technique de suréchantillonnage afin de traiter le déséquilibre résiduel. Particulièrement adapté aux ensembles de données manufacturières, il est capable de générer des données synthétiques réalistes à variables mixtes, en préservant la distribution des attributs catégoriels et en reproduisant les structures complexes, souvent non gaussiennes, des variables numériques de procédé. Comme expliqué dans le chapitre de revue de littérature, l'approche consiste à équilibrer la représentation des catégories dominantes et rares en conditionnant la génération sur les attributs catégoriels, tandis que les variables numériques sont modélisées de manière à capturer les comportements multimodaux des processus. Les échantillons

synthétiques ainsi produits sont ensuite intégrés au jeu de données initial afin de renforcer la représentation de la classe minoritaire sans altérer la structure sous-jacente des données.

3.2.4 Modèle de détection de défauts

Cette section expose la stratégie de modélisation adoptée pour la détection de défauts dans un contexte marqué par un fort déséquilibre de classes. Dans cette méthodologie, la détection de défauts est formalisée comme un problème de classification binaire supervisée. Les sous-sections suivantes détaillent les différentes étapes de la méthodologie de modélisation, telles qu'illustrées dans la Figure 3.3.

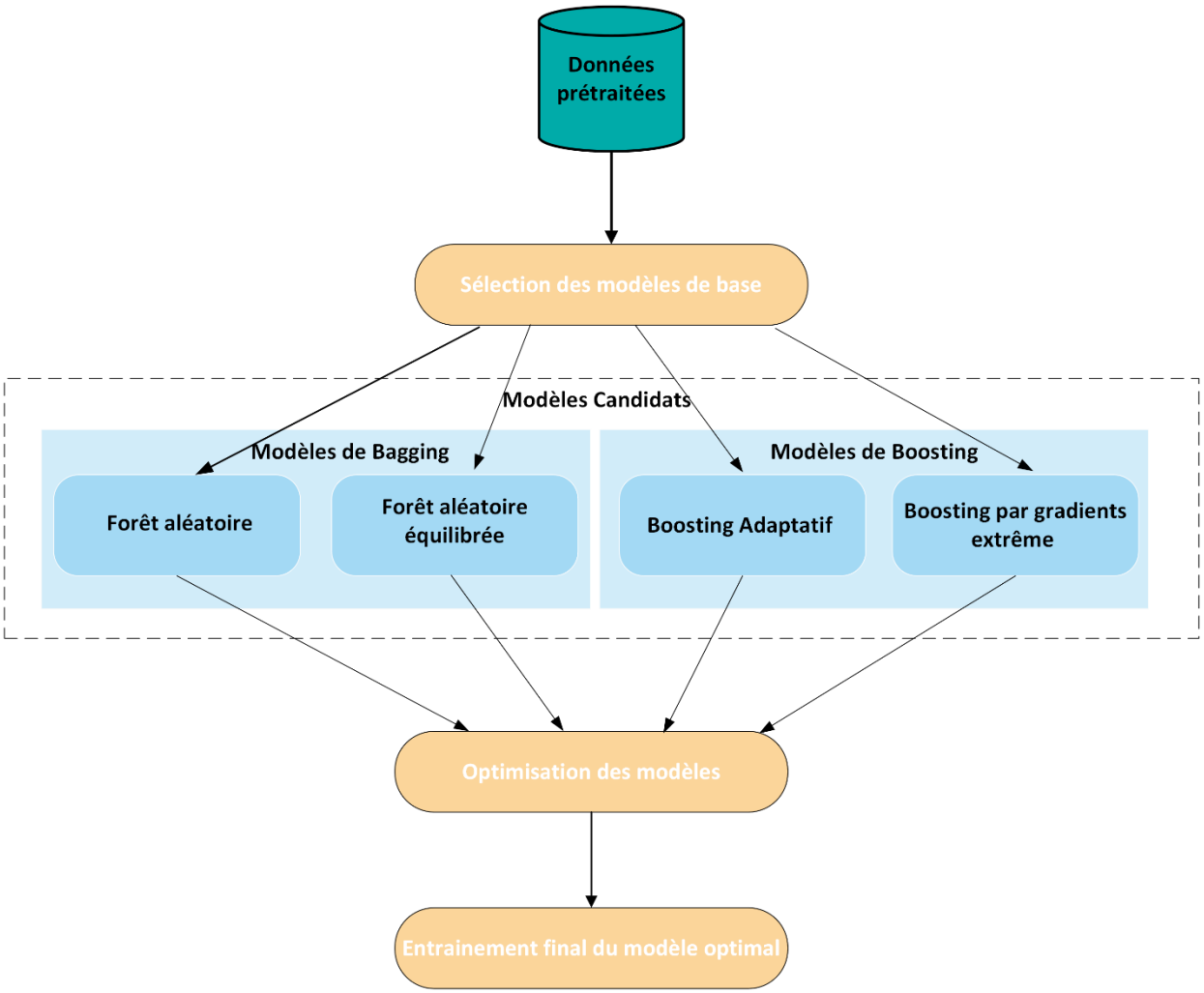


Figure 3.3 Méthodologie de détection

A. Sélection des modèles de base

Ce travail de recherche s'inscrit dans un contexte de fort déséquilibre des classes. Pour la phase de modélisation, il est donc question d'étudier les modèles présentant une robustesse face à ce problème. Dans une revue de la littérature sur les méthodes de traitement du déséquilibre de données en contexte industriel, conduite par de Giorgio et al. (2023), il a été établi que les approches ensemblistes sont reconnues pour leur robustesse et sont, de ce fait, fréquemment utilisées pour les tâches de classification. Ces modèles se distinguent principalement en deux grandes catégories. D'une part, *les approches* de *bagging* (Bootstrap Aggregating), qui visent à réduire la variance en combinant plusieurs modèles entraînés sur des sous-échantillons aléatoires des données, et d'autre part, *les approches* de *boosting*, qui cherchent à améliorer la performance prédictive en construisant de manière séquentielle une série de modèles, chacun corrigeant les erreurs commises par le précédent.

Modèle de Bagging

Foret aléatoire (Random Forest) : C'est un modèle ensembliste qui repose sur la construction de plusieurs arbres, communément appelé « forêt ». Chaque arbre de la forêt est entraîné sur un échantillon aléatoire des données obtenu par rééchantillonnage avec remise (*bootstrap sampling*) et ne considère qu'un sous-ensemble aléatoire de variables lors des divisions (*feature randomization*). Lors de la phase de classification, chaque arbre produit une prédiction, et la décision finale est obtenue par vote majoritaire. La robustesse de ce modèle selon Imani et al. (2025) découle de son principe de « sagesse collective » (*wisdom of crowds*) : si chaque arbre pris isolément est susceptible de produire des erreurs, l'agrégation des décisions permet d'atténuer ces biais individuels par un effet de moyenne, aboutissant ainsi à des prédictions globalement plus stables.

Foret aléatoire équilibrée (Balanced Random Forest) : Lors de la construction des arbres de décision, ce modèle remplace le bootstrapping classique des données par un rééchantillonnage équilibré entre classes. Concrètement, à chaque itération de construction d'un arbre, un échantillon aléatoire de la classe minoritaire est sélectionné, puis un sous-échantillon de taille identique est tiré de la classe majoritaire. Ce mécanisme permet de réduire le biais en faveur de la classe majoritaire lors d'un problème de déséquilibre de classes comme la détection de défauts (He et al., 2023).

Modèle de Boosting

Boosting Adaptatif (Adaptive Boosting) : est un modèle d'apprentissage supervisé fondé sur le principe du boosting. L'idée c'est d'entrainer successivement une série de classificateurs faibles, le plus souvent des arbres de décision peu profonds (appelés *stumps*). Chaque nouvel arbre cherche à corriger les erreurs commises à l'itération précédente. Plus précisément, les observations mal classées reçoivent un poids plus élevé lors des itérations suivantes, ce qui incite le modèle à se concentrer sur les exemples difficiles à prédire (Haixiang, 2016). L'agrégation finale des classificateurs faibles se fait à travers une combinaison pondérée. Au terme du processus, le but est de construire un méta-modèle dont la performance est censée excéder celle de ses composants individuels.

Boosting par gradients extrême (eXtreme Gradient Boosting) : Ce modèle est une version améliorée du modèle Adaboost. Sa spécificité réside dans plusieurs optimisations algorithmiques et numériques, telles que la régularisation explicite pour limiter le surapprentissage, l'approximation efficace des meilleurs points de coupure, la gestion des valeurs manquantes et l'exploitation parallèle du calcul. Ces améliorations permettent d'obtenir un modèle à la fois robuste, performant et particulièrement adapté aux grands volumes de données comme l'a indiqué Taha (2025).

B. Optimisation des modèles

L'optimisation des modèles est une étape aussi importante que les étapes précédentes. La performance et la capacité de généralisation d'un modèle d'apprentissage automatique ne dépendent pas uniquement de son architecture, mais également du réglage approprié de ses hyperparamètres définis en amont. Les arbres de décision possèdent plusieurs hyperparamètres influençant à la fois leur complexité et leur capacité de généralisation. Dans cette méthodologie, l'objectif principal est de favoriser la généralisation des modèles tout en améliorant la détection des défauts, c'est-à-dire la classe minoritaire. À cette fin, un réglage ciblé des hyperparamètres est réalisé.

Pour les modèles de type bagging, les hyperparamètres *max_depth*, *min_samples_split* et *n_estimators* devraient être ajustés. La profondeur maximale (*max_depth*) doit être limitée afin d'éviter que les arbres ne deviennent excessivement complexes, ce qui réduirait leur capacité de généralisation en les exposant au surapprentissage. Le paramètre *min_samples_split* est fixé de manière à contrôler le nombre minimal d'observations nécessaires pour diviser un nœud, ce qui

permet de régulariser la croissance de l'arbre et de réduire la fragmentation excessive des données. Enfin, le nombre d'arbres (*n_estimators*) doit être calibré afin de trouver un compromis entre performance et temps de calcul.

En complément, pour les modèles de boosting, le paramètre *learning_rate* doit également être ajusté. Celui-ci contrôle la contribution de chaque nouvel arbre dans le processus séquentiel de correction des erreurs. L'apprentissage*e* trop élevée accélère l'apprentissage mais risque de conduire au surapprentissage, tandis qu'une valeur plus faible permet un apprentissage plus progressif et robuste.

Par ailleurs, un apprentissage sensible aux coûts est appliqué pour renforcer l'attention accordée à la classe minoritaire. Cela est mis en œuvre via le paramètre *class_weight* pour la Forêt Aléatoire, et *scale_pos_weight* pour AdaBoost et XGBoost. Ce mécanisme permet d'augmenter la pénalisation associée aux erreurs de classification de la classe minoritaire, améliorant ainsi sa détection. Seul le Balanced Random Forest ne nécessite pas cette adaptation, dans la mesure où son mécanisme de construction intègre déjà un équilibrage explicite entre classes.

Une fois les hyperparamètres identifiés, il est nécessaire de mettre en œuvre une procédure permettant de sélectionner la combinaison la plus performante. Dans ce travail, *RandomizedSearchCV* est proposé. Afin de mieux en comprendre le fonctionnement, nous proposons d'illustrer son déroulement à travers trois étapes principales : la construction de l'espace des hyperparamètres et l'échantillonnage aléatoire, l'application d'un processus de validation croisée et l'évaluation des performances obtenues et sélection du modèle optimal.

Construction de l'espace des hyperparamètres et l'échantillonnage aléatoire

Tout d'abord, chaque hyperparamètre peut être représenté comme une dimension dans un espace de recherche. Par exemple, si l'on considère deux hyperparamètres tels que *max_depth* et *n_estimators*, chaque combinaison possible correspond à un point dans un plan bidimensionnel. Ainsi, l'ensemble des combinaisons peut être visualisé comme une grille de points dans un espace 2D, comme illustré dans la Figure 3.4.

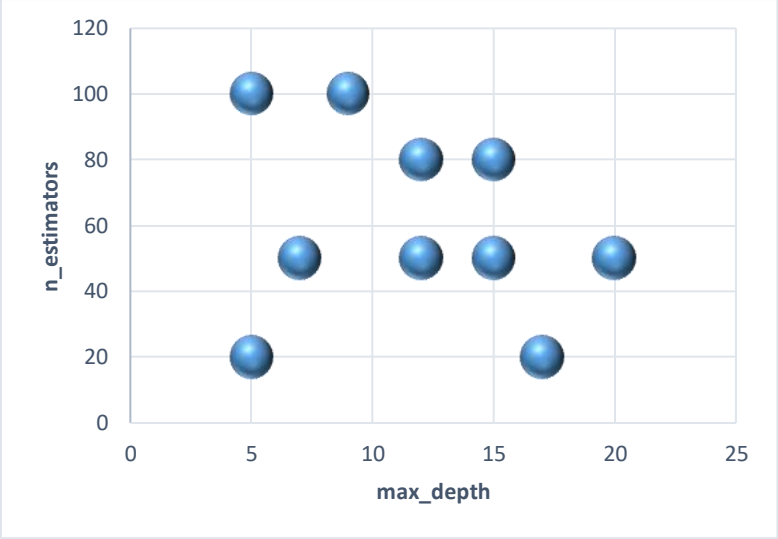


Figure 3.4 Exemple d'espace des hyperparamètres

L'objectif est alors d'identifier le point, c'est-à-dire la combinaison d'hyperparamètres, qui conduit à la meilleure performance du modèle. L'approche *RandomizedSearchCV* procède en sélectionnant aléatoirement un nombre prédéfini de points au sein de cet espace d'hyperparamètres. Les points retenus, illustrés en orange sur la Figure 3.5, sont choisis de manière non uniforme, de façon à couvrir différentes régions de l'espace de recherche.

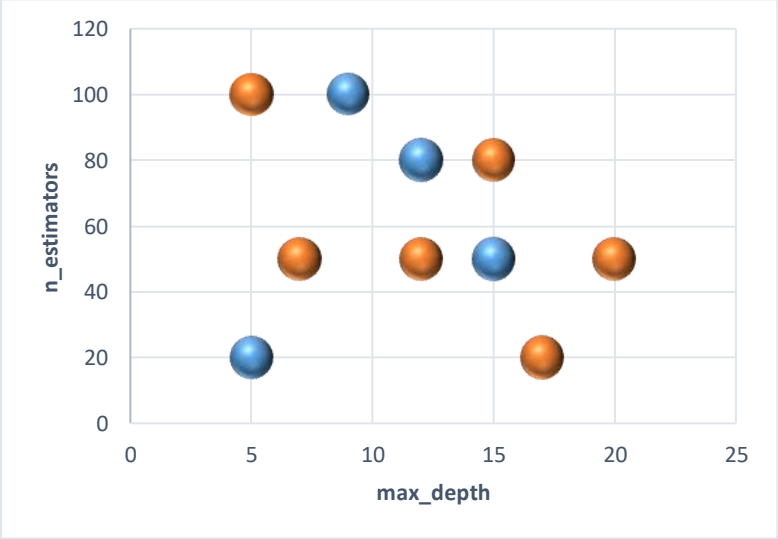


Figure 3.5 Sélection des combinaisons d'hyperparamètres

Application d'un processus de validation croisée (K-Fold)

Afin d'évaluer chaque combinaison d'hyperparamètres échantillonnée, RandomizedSearchCV utilise la validation croisée. C'est une technique qui permet d'estimer la performance d'un modèle sur des données non vues afin de réduire le risque de surapprentissage. Ce processus commence par diviser le jeu de données d'entrainement en *k* sous-ensembles (appelés plis). Pour chaque itération de la validation croisée, un pli est désigné comme l'ensemble de validation, et les *k-1* plis restants sont combinés pour former l'ensemble d'entraînement. Le modèle est entraîné sur l'ensemble d'entraînement et évalué sur l'ensemble de validation. Ce processus est répété K fois, de sorte que chaque pli serve une fois d'ensemble de validation. Les scores de performance obtenus à chaque itération sont ensuite moyennés pour donner une estimation.

La Figure 3.6 illustre le processus de validation croisée à *k* plis (ici, *k*=5). Chaque ligne représente une itération, montrant comment les données sont divisées en ensembles d'entraînement et de validation.

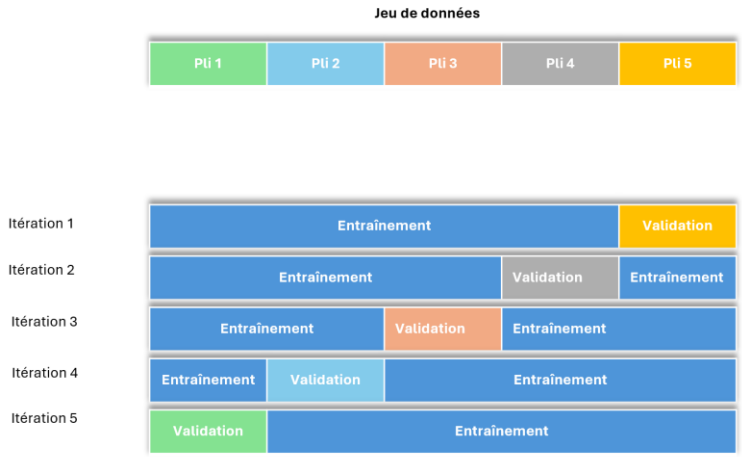


Figure 3.6 Processus de validation croisée

Évaluation des performances et sélection du modèle optimal

Une fois l'évaluation de chaque combinaison d'hyperparamètres échantillonnée au moyen de la validation croisée est effectuée, *RandomizedSearchCV* compare les scores de performance obtenus. L'objectif est d'identifier la combinaison qui maximise le score moyen de validation croisée. Une fois cette combinaison déterminée, le modèle est réentraîné sur l'ensemble complet des données

d'apprentissage (non partitionné en plis), en utilisant les hyperparamètres optimaux. Le modèle ainsi finalisé constitue celui qui sera retenu pour réaliser des prédictions sur de nouvelles données.

C. Entraînement et validation du modèle sélectionné

Après l'entraînement des modèles proposés à la section A, à l'aide d'hyperparamètres optimisés, l'étape suivante consiste à sélectionner le modèle le plus performant, en l'occurrence AdaBoost, puis à le réentraîner sur l'ensemble complet des données. AdaBoost construit un ensemble d'arbres de décision, chacun étant conçu pour corriger les erreurs commises par les arbres précédents. Ce processus repose sur les principes du boosting et vise, dans le cadre de cette étude, à améliorer la détection des défauts issus du processus d'assemblage.

Entraînement du modèle sélectionné

Afin d'optimiser les performances du modèle, il a été jugé pertinent d'adopter une approche systématique de recherche des hyperparamètres, incluant *n_estimators*, *learning_rate*, *class_weight* et *max_depth*, plutôt qu'une approche aléatoire. Dans ce contexte, la méthode *GridSearchCV* a été privilégiée. Cette technique explore exhaustivement toutes les combinaisons possibles d'hyperparamètres définies dans la grille de recherche. Pour chaque hyperparamètre, un ensemble de valeurs candidates est spécifié, et toutes les combinaisons issues de ce maillage sont évaluées. L'évaluation des modèles obtenus, utilisant des méthodes spécifiées à la section 3.2.5, est réalisée au moyen de la validation croisée, garantissant une estimation robuste et impartiale des performances.

Analyse de la généralisabilité du modèle du modèle

Une validation des performances s'avère indispensable dans la perspective d'un déploiement industriel. Afin de rendre le modèle opérationnel, il est nécessaire de procéder à une évaluation reposant sur plusieurs approches. Dans cette méthodologie, deux stratégies complémentaires de validation ont été mises en œuvre. La première consiste à tester le modèle sur des données n'ayant subi aucune méthode d'échantillonnage, mais uniquement un processus de nettoyage, de manière à éviter tout biais et à garantir une évaluation sur un jeu de données représentatif de la réalité industrielle. La seconde repose sur un test effectué sur une autre période temporelle, éloignée de plus d'une année, permettant ainsi d'examiner la robustesse et la capacité de généralisation du modèle dans des conditions évolutives.

3.2.5 Évaluation des modèles

L'évaluation d'un modèle d'apprentissage, dans le contexte étudié, devrait mettre en évidence sa fiabilité ainsi que sa validité opérationnelle au sein d'environnements industriels. Les métriques classiques de classification sont principalement dérivées de la matrice de confusion, présentée ci-dessous :

		Classes prédites
Classes réelles	TP	FN
	FP	TN

Vrais positifs (TP – *True Positives*) : nombre d'observations appartenant réellement à la classe positive et correctement prédites comme telles par le modèle. Dans le cadre de cette étude, il s'agit des produits effectivement sans défaut et correctement identifiés comme tels.

Vrais négatifs (TN – *True Negatives*) : nombre d'observations appartenant réellement à la classe négative et correctement prédites comme telles. Ici, cela correspond aux produits défectueux qui ont été effectivement détectés par le modèle.

Faux positifs (FP – *False Positives*) : nombre d'observations appartenant réellement à la classe négative mais incorrectement prédites comme appartenant à la classe positive (erreur de type I). Dans ce cas, il s'agit de produits défectueux que le modèle n'a pas réussi à identifier comme tels.

Faux négatifs (FN – *False Negatives*) : nombre d'observations appartenant réellement à la classe positive mais prédites à tort comme négatives (erreur de type II). Cela correspond aux produits sans défaut qui ont été erronément signalés comme défectueux (fausses alarmes).

À partir de la matrice de confusion, plusieurs métriques de performance peuvent être dérivées afin d'évaluer la qualité de classification du modèle. Parmi les plus couramment utilisées figurent l'accuracy, le rappel et le F1-score.

L'exactitude (Accuracy) : Cette métrique correspond à la proportion de prédictions correctes réalisées par le modèle par rapport au nombre total de prédictions. Elle se calcule comme le rapport suivant :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

Le rappel (Recall) : correspond à la proportion d'éléments positifs correctement identifiés par le modèle parmi l'ensemble des éléments réellement positifs.

$$Recall = \frac{TP}{TP + FN} \tag{3.2}$$

La Précision (Precision) : mesure la proportion d'observations prédites comme positives qui sont réellement positives. Elle renseigne donc sur la fiabilité des prédictions positives du modèle.

$$Precision = \frac{TP}{TP + FP} \tag{3.3}$$

Le F1-score : C'est la moyenne harmonique entre la précision et le rappel. Il constitue une métrique synthétique qui équilibre la capacité du modèle à éviter les faux positifs (précision) et les faux négatifs (rappel).

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3.4}$$

Bien que plusieurs travaux reposent souvent principalement sur l'exactitude, cette mesure présente une limite fondamentale dans les contextes d'apprentissage déséquilibré, car elle ne prend pas en compte la prévalence des classes. En conséquence, un modèle peut obtenir des scores artificiellement élevés en privilégiant la classe majoritaire. Ce phénomène, désigné sous le terme de « paradoxe de l'accuracy », peut ainsi donner l'illusion d'une performance élevée tout en échouant complètement à détecter des défauts critiques.

Pour pallier ces limites, un ensemble spécifique de métriques d'évaluation est mobilisé. L'exactitude équilibrée (*Balanced Accuracy)* et la Moyenne Géométrique (*G-mean*) se révèlent particulièrement pertinentes, car elles reflètent la capacité du modèle à maintenir une performance équilibrée entre toutes les classes, respectivement en moyennant ou en multipliant les valeurs de rappel par classe comme illustré ci-dessous.

$$Exactitude\ équilibrée = \frac{Recall(Classe\ 0) + Recall(Classe\ 1)}{2} \tag{3.5}$$

$$Moyenne\ Géométrique = \sqrt{Recall(Classe\ 0) \times Recall(Classe\ 1)} \tag{3.6}$$

L'utilisation de ces métriques garantit que l'évaluation du modèle demeure alignée sur les exigences opérationnelles de la détection de défauts, où la non-détection d'anomalies rares peut entraîner des conséquences majeures.

3.2.6 Explicabilité du modèle

La dernière étape de la méthodologie proposée porte sur l'explicabilité des modèles à travers les techniques d'*Explainable Artificial Intelligence* (XAI). Le caractère de « boîte noire » des modèles d'intelligence artificielle et d'apprentissage automatique constitue en effet une limite majeure, en particulier dans les environnements industriels critiques, où les sorties du modèle influencent directement des décisions relatives à la qualité des produits, à la sécurité des procédés et à l'efficacité économique. Deux besoins fondamentaux en découlent : la latence et l'interprétabilité. Les systèmes de production en temps réel requièrent des modèles capables de fournir des prédictions dans des délais très restreints tout en justifiant leurs décisions de manière intelligible. L'intégration de techniques XAI vise à combler l'écart entre performance prédictive et exploitabilité opérationnelle, en rendant le comportement du modèle compréhensible pour les experts métiers. Ces explications se déclinent à deux niveaux complémentaires global et local.

A. Explicabilité globale du modèle

Les méthodes d'explicabilité globale permettent d'identifier les tendances générales et l'importance des variables explicatives sur l'ensemble du jeu de données. Elles visent à estimer la contribution moyenne de chaque variable aux prédictions du modèle, offrant ainsi une vision d'ensemble des facteurs déterminants. Parmi les approches courantes figurent l'importance par permutation (Elshawi et al., 2019), qui mesure la baisse de performance du modèle lorsqu'une variable est aléatoirement réarrangée, et les graphiques de dépendance partielle (*PDPs*) (Friedman, 2001), qui visualisent l'effet marginal d'une variable sur la sortie prédite. Des méthodes plus récentes ont été mobilisées dans cette méthodologie, telles que les SHapley Additive exPlanations (SHAP) (Lundberg et Lee 2017), qui étendent ces principes en attribuant à chaque variable une valeur d'importance cohérente, issue de la théorie des jeux coopératifs, reflétant sa contribution moyenne aux prédictions dans toutes les combinaisons possibles de variables.

Lorsque SHAP est appliqué aux modèles de bagging et boosting il ne traite pas l'ensemble comme une boîte noire unique. Au lieu de cela, il décompose la contribution de chaque caractéristique en considérant son impact à travers tous les apprenants faibles du modèle ensembliste. L'idée est d'attribuer une valeur SHAP à chaque caractéristique pour chaque apprenant faible, puis d'agréger ces valeurs de manière pondérée.

Pour une instance x est une caractéristique j, la valeur SHAP $\varphi_j\ (x)$ pour l'ensemble du modèle entrainé peut être approximée comme la somme pondérée des valeurs SHAP de cette caractéristique j pour chaque apprenant faible h_m

$$\varphi_j\ (x) \approx \sum_{m=1}^{M} \alpha_m . \varphi(x, h_m) \tag{3.7}$$

Où : *$\varphi(x, h_m)$est la valeur SHAP de la caractéristique j pour l'instance x calculée pour le m-ième apprenant faible h_m. * $\alpha_m$est le poids de l'apprenant faible h_m dans l'ensemble du modèle entrainé.

Cette agrégation pondérée est cruciale car elle reflète l'importance relative de chaque apprenant faible dans la prédiction finale du modèle ensembliste. La visualisation dans la Figure 3.7 illustre ce processus d'agrégation :

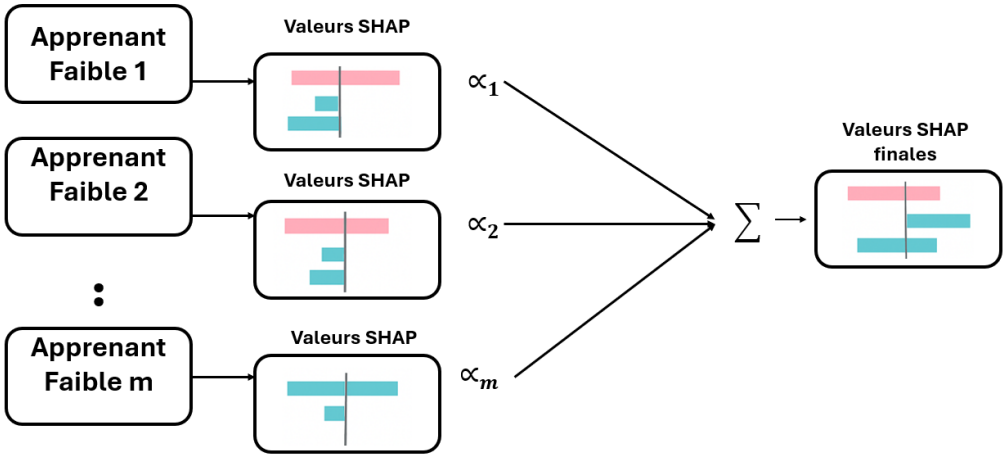


Figure 3.7 Agrégation des valeurs SHAP pour les modèles ensemblistes

En moyennant les valeurs SHAP absolues sur de nombreuses prédictions, on peut obtenir un classement des caractéristiques les plus importantes pour le modèle. Ce classement des variables contribue non seulement à affiner les modèles, mais fournit également aux experts industriels des indications exploitables pour comprendre les déterminants critiques des prédictions.

B. Explicabilité locale du modèle

En complément des analyses globales, les méthodes d'interprétabilité locale sont essentielles pour expliquer des prédictions individuelles, notamment dans le cadre de décisions critiques en temps réel, telles que l'arrêt d'une ligne de production ou l'inspection ciblée d'une unité. Parmi les approches représentatives, on distingue LIME, SHAP appliqué au niveau instance (Lundberg et Lee 2017) et les explications contrefactuelles (Fatemi et al., 2025). LIME construit un modèle local de substitution autour d'une prédiction donnée, SHAP attribue des scores de contribution cohérents à chaque variable pour une instance particulière, et les méthodes contrefactuelles indiquent les modifications minimales nécessaires pour changer la décision du modèle. La section suivante présente plus en détail la méthode LIME appliquée dans ce travail, en mettant l'accent sur la manière dont elle a été adaptée au modèle AdaBoost.

Concrètement, LIME fonctionne en construisant un modèle local interprétable qui approxime le comportement d'AdaBoost dans le voisinage immédiat d'une instance particulière. Pour ce faire, des perturbations sont générées autour de l'observation à expliquer, puis évaluées par AdaBoost afin d'obtenir les prédictions correspondantes. Chaque perturbation est ensuite pondérée en

fonction de sa proximité avec l'instance d'origine, ce qui permet d'accorder plus d'importance aux observations les plus similaires. Sur la base de ces données pondérées, un modèle simple, généralement une régression linéaire, est ajusté afin de reproduire localement la décision complexe issue de la combinaison d'arbres d'AdaBoost. Les coefficients de ce modèle de substitution fournissent alors une mesure directe de l'influence relative des variables explicatives sur la prédiction considérée

L'algorithme ci-dessous décrit les principales étapes de ce processus.

Entrée :	Modèle AdaBoost $f(x)$, instance à expliquer x_0, nombre de perturbations N, noyau de proximité π_{x0}, fonction de perte ℓ, pénalisation Ω.
Sortie :	Modèle local $g(x) = \beta_0 + \beta^T x$, contributions locales β, fidélité locale $\mathcal{L}(f,g,\pi_{x0})$.
Étape 1 :	Générer N perturbations x_i autour de l'instance x_0.
Étape 2 :	Pour chaque perturbation x_i, calculer la prédiction d'AdaBoost $\hat{y}_i = f(x_i)$ et son poids de proximité $w_i = \pi_{x0}(x_i)$.
Étape 3 :	Ajuster un modèle linéaire local $g(x)$ en minimisant la perte pondérée : $\Sigma\, w_i \cdot \ell(\hat{y}_i, g(x_i)) + \Omega(g)$.
Étape 4 :	Retourner le modèle local g, les coefficients β comme explication locale et la fidélité $\mathcal{L}(f,g,\pi_{x0})$.

Cette technique apporte ainsi une granularité d'explication indispensable pour renforcer la confiance dans le modèle et faciliter son intégration dans des environnements de production industriels.

Conclusion

Ce chapitre a permis de définir les objectifs de notre recherche et de proposer une méthodologie fondée sur la science des données. Inspirée du modèle CRISP-DM, cette approche vise à améliorer le processus de contrôle qualité au sein d'une ligne d'assemblage. La mise en œuvre de cette méthodologie dans un cas industriel concret sera présentée au chapitre suivant.

CHAPITRE 4 CAS D'ÉTUDE

Dans ce chapitre, une brève introduction du contexte de ce travail est présentée à travers la description du partenaire industriel ainsi que du contexte industriel spécifique au cas d'étude. Par la suite, la méthodologie établie précédemment est appliquée au cas d'étude concret en collaboration avec ce partenaire.

4.1 Mise en œuvre de la méthodologie pour la détection de défauts dans un processus d'assemblage

4.1.1 Contexte industriel

A. Présentation du partenaire industriel

Le partenaire industriel de ce projet est une usine de la division nord-américaine d'une multinationale manufacturière, située à Québec. Ce site de production est spécialisé dans la fabrication de produits destinés à une large gamme de véhicules, incluant les voitures particulières, les camions légers, les poids lourds, les autobus, ainsi que les véhicules agricoles et les motocyclettes. En matière d'innovation, cette usine a constamment fait évoluer ses équipements et a déployé, au fil des années, de multiples dispositifs de collecte de données. Ces données sont destinées à l'amélioration continue de ses processus de production, notamment par l'intégration de technologies innovantes.

B. Compréhension du cas industriel

Cette étude se focalise sur une étape précise de la ligne de production, à savoir l'assemblage. En effet, l'usine a opéré une transition technologique en remplaçant la plupart de ses anciennes machines par 15 nouvelles unités semi-automatiques, ne conservant que 3 postes manuels. Ces machines sont multiproduits, dépendamment de la demande, des transitions peuvent se faire. L'optimisation de la productivité n'a cependant pas éliminé l'occurrence de certains défauts critiques. Le processus actuel ne permet de les identifier qu'à l'étape finale d'inspection manuelle. Ceci signifie que des pièces défectueuses sont traitées tout au long de la chaîne de production avant d'être finalement écartées. C'est dans le contexte de développer une solution de détection de défauts que s'inscrit ce travail de recherche. Le mandat de ce projet est donc de déterminer s'il est possible d'exploiter les données de production pour identifier les défauts d'assemblage au moment même ou même avant qu'ils se produisent.

Dans un premier temps, il est donc nécessaire de comprendre en détail le processus d'assemblage ainsi que les types de défauts susceptibles d'apparaître. Le processus d'assemblage illustré à travers la Figure 4.1 est réalisé à l'aide de la machine VMI MAXX, un équipement industriel de pointe largement utilisé pour la fabrication de pneus. Cette machine repose sur une architecture à deux tambours, l'un dédié à la construction de la carcasse (*Carcass Drum*) et l'autre à l'assemblage de la ceinture et de la bande de roulement (Belt *& tread Drum*). Les composants internes de la carcasse sont d'abord appliqués sur le tambour carcasse, puis transférés via un anneau de transfert fixe vers le tambour ceinture où s'ajoutent la ceinture et la bande de roulement.

Figure 4.1 Processus d'assemblage avec la machine VMI MAXX

Au cours de ce processus, divers types de défauts peuvent apparaître, avec des degrés de gravité variables. Certains peuvent être corrigés par des reprises mineures ou majeures, tandis que d'autres conduisent au rebut du pneu. Parmi ces défauts, notre attention se porte spécifiquement sur deux anomalies critiques et irréparables, liées à une mauvaise adhérence entre les plis de la carcasse, susceptibles de survenir dès la première étape de l'assemblage.

D'un point de vue production, les temps de cycle sur les lignes d'assemblage semi-automatiques sont courts généralement de 50 à 60 secondes. Cela crée un besoin pressant de détection précoce des anomalies afin que les opérateurs puissent réagir en temps réel, soit en réparant les produits

défectueux lorsqu'une intervention est possible et que le défaut est détecté avant qu'il se produise, soit en stoppant la progression de pneus défectueux sur la ligne. À titre de contextualisation, la Figure 4.1 mets en évidence l'étape où les données étudiées ont été collectées et où apparaissent typiquement les défauts liés à l'adhérence. Les sous-sections suivantes détaillent la mise en œuvre de la méthodologie proposée pour détecter ces défauts critiques.

4.1.2 Compréhension des données

A. Collecte des données

Les données utilisées dans cette étude ont été collectées automatiquement par le partenaire industriel et stockées dans une base de données *SQL Server*. Ces données brutes proviennent de trois sources distinctes : les paramètres de configuration saisis sur les machines d'assemblage, les mesures issues des capteurs intégrés, et enfin, les informations enregistrées manuellement par les opérateurs. Pour les besoins de l'analyse, les données pertinentes ont été extraites de la base de données et exportées au format *CSV*, afin de permettre leur exploitation dans un environnement *Jupyter Notebook*. Les données collectées pour cette étude couvrent les premiers 7 mois de l'année 2023. Cette période a été retenue car elle correspondait aux données les plus récentes et complètes disponibles au moment de l'initiation de ce projet. Les principales tables *SQL* exploitées pour cette analyse, qui structurent les informations relatives au processus d'assemblage, sont les suivantes :

Table de production : Cette table constitue la source d'information centrale, regroupant les données générées lors de l'assemblage de chaque unité. Elle regroupe des informations de traçabilité telles qu'un identifiant unique du produit, l'identifiant de la machine, l'horodatage de l'opération ainsi que le type de produit. D'autres éléments, tels que le numéro de série du rouleau de caoutchouc utilisé, peuvent également y être renseignés. Cette table consigne également les paramètres clés du processus comme illustrées sous un petit exemple du jeu de données dans le tableau 4.1, présentés sous trois formes : les paramètres introduits automatiquement par la machine (paramètres consignés), les mesures correspondantes enregistrées par les capteurs intégrés à la machine, et finalement les traces des interventions manuelles, telles que les ajustements réalisés par les opérateurs (ajustements des paramètres). Elle inclut notamment les pressions appliquées ainsi que les mesures dimensionnelles du produit, en particulier celles des couches de carcasse et des joints d'adhérence.

Tableau 4.1 Exemple de paramètres du processus

Paramètres consignes	**Mesures réelles**	**Ajustements des paramètres**
Opérateurs (*UserName*)	-	-
Type de produit (*RecipeName*)	-	-
Longueurs des plis de carcasse (*BP_x_Length_Setpoint*)	Longueur réelle après la découpe (*BP_x_Length_Measured*)	Ajustements de la longueur du plus (+/-) (*BP_x_Length_Offset*)
Pression_x_consignes (*Pressure_x*)	-	Pression ajustée (*Pressure_x_Change*)

Table de classification : cette table contient les résultats de l'inspection finale. Elle fait le lien avec la première table via l'identifiant unique du produit et lui associe une étiquette de classification indiquant si le produit est conforme ou défectueux, en spécifiant, le cas échéant, la nature du défaut constaté.

B. Description des données

La table de classification analysée dans cette étude regroupe 77 929 observations, correspondant à l'ensemble des produits défectueux enregistrés au cours de 7 mois de l'année 2023. Ces produits couvrent un large éventail de défaillances potentielles liées au processus de fabrication.

Chaque enregistrement est associé à un identifiant unique assurant la traçabilité du produit, ainsi qu'à une variable catégorielle décrivant le code de défaut observé lors de l'inspection qualité en fin de chaîne de production.

Dans le cadre de ce travail, un seul type de défauts a été retenu, considéré comme critique par le partenaire industriel en raison de leur impact significatif sur la qualité du produit final.

Cette catégorie correspond à :

- **Une mauvaise adhérence de la couche de carcasse.**

Après filtrage, le jeu de données final comprend 757 observations correspondant à ces deux types de défauts.

Ce sous-ensemble a été sélectionné pour les étapes ultérieures d'analyse et de modélisation, dans le but de mieux comprendre les facteurs de défaillance et d'améliorer la détection des défauts d'adhésion au sein du procédé industriel.

C. Exploration des données

Profilage statistique et Visualisation

Découverte du processus

Notre analyse exploratoire a débuté par l'étude de la répartition des types de produits, ou "recettes", sur les différentes machines d'assemblage. Cette démarche visait principalement à évaluer la fréquence des changements de recette. L'enjeu était de confirmer si ces transitions étaient suffisamment fréquentes pour justifier une analyse de leur impact, sachant que chaque changement implique un recalibrage de la machine et donc un risque pour la stabilité du processus.

La Figure 4.2 met en évidence que les recettes sont, en pratique, majoritairement assignées à des machines spécifiques. Cette répartition favorise la stabilité des réglages et limite les interruptions liées aux changements fréquents de configuration. Toutefois, certaines recettes, telles que *Recette_A* et *Recette_E*, sont produites sur plusieurs machines différentes, traduisant une flexibilité intentionnelle du système de production.

Selon le partenaire industriel, cette stratégie organisationnelle résulte d'un compromis entre deux impératifs industriels. D'une part, la réduction des temps d'arrêt et de la variabilité associée aux recalibrages répétés et d'autre part, la nécessité d'assurer une répartition équilibrée de la charge de production pour répondre aux fluctuations de la demande.

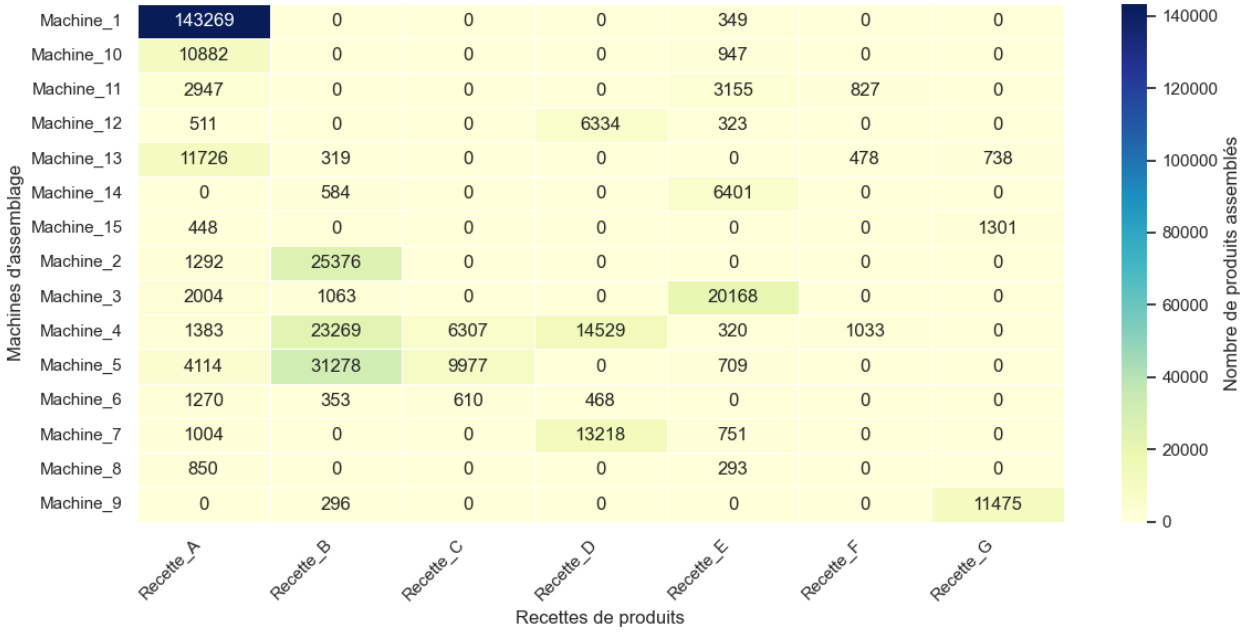


Figure 4.2 Profil de production des différentes machines d'assemblage

D'un point de vue qualité, cette situation peut engendrer que la spécialisation d'une machine sur une recette unique tend à renforcer la régularité et la reproductibilité du procédé, tandis que la coexistence de plusieurs versions sur une même machine accroît le risque de dérives transitoires, notamment lors des phases de recalibrage.

Analyse de la non-qualité

Il est maintenant question de comprendre comment les défauts sont répartis dans le jeu de données ainsi qu'examiner les relations entre les défauts et les différentes variables de production. Dans l'ensemble du jeu de données, les défauts d'adhérence des plis de carcasse représentent environ 2 % de la production totale (Figure 4.3). Toutefois, lorsqu'ils sont regroupés avec d'autres défauts issus des étapes d'assemblage, ne conduisant pas nécessairement à un rebut du produit, les défauts d'adhésion constituent près de 17 % de l'ensemble des observations défectueuses. En raison de leur caractère irréparable et de leur impact direct sur la qualité finale du produit, ces défauts ont été retenus comme objet principal des analyses menées dans cette étude.

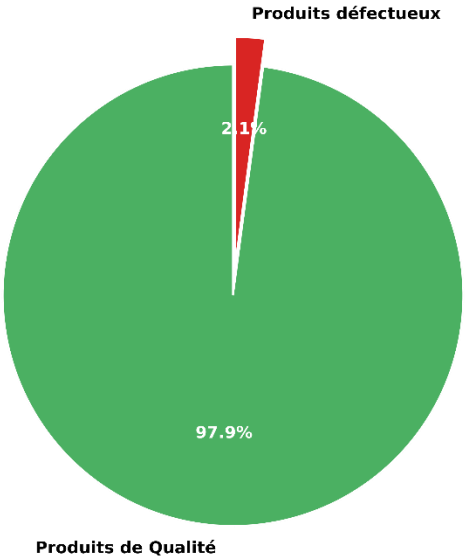
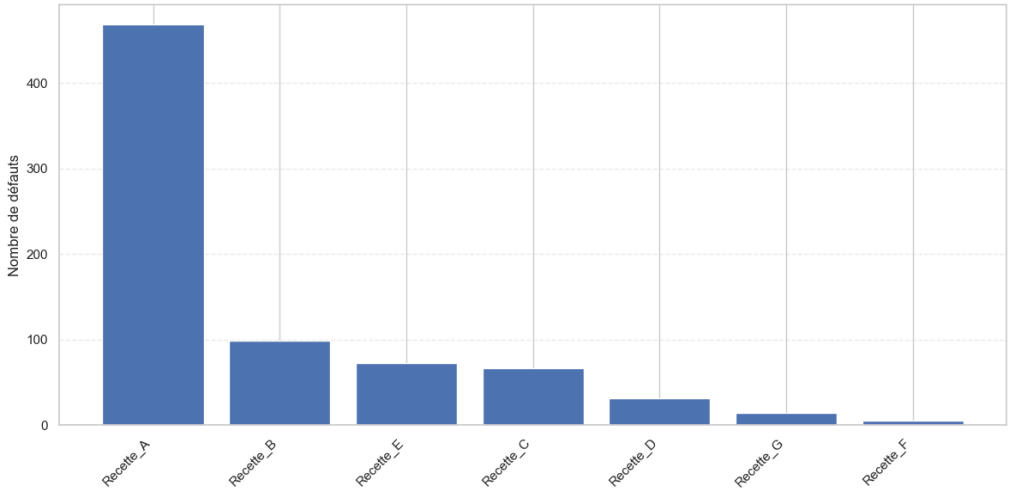
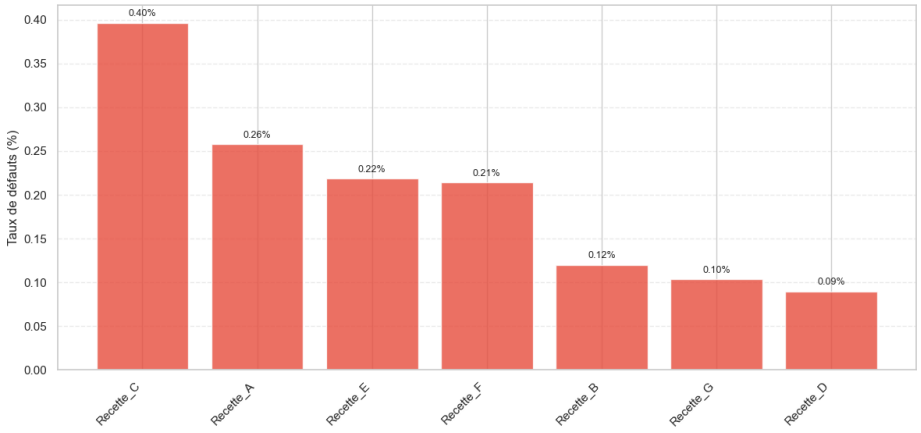


Figure 4.3 Fréquence des défauts d'adhérence des plis de carcasse dans la production totale

Ayant identifié les défauts d'adhérence des plis de carcasse comme l'objet central de l'analyse, des hypothèses d'investigation ont été formulées afin d'en explorer les causes potentielles. Dans un premier temps, l'analyse s'intéresse à la manière dont les défauts se répartissent selon les différentes recettes de production, dans le but de déterminer si certaines recettes présentent une sensibilité accrue aux anomalies. L'analyse présentée dans la Figure 4.4 (a) met en évidence que la Recette_A concentre un nombre de défauts plus important. Toutefois, une analyse normalisée (Figure 4.4 (b)) montre que cette observation ne traduit pas une moindre qualité intrinsèque, mais résulte directement du volume de production beaucoup plus élevé de cette recette.



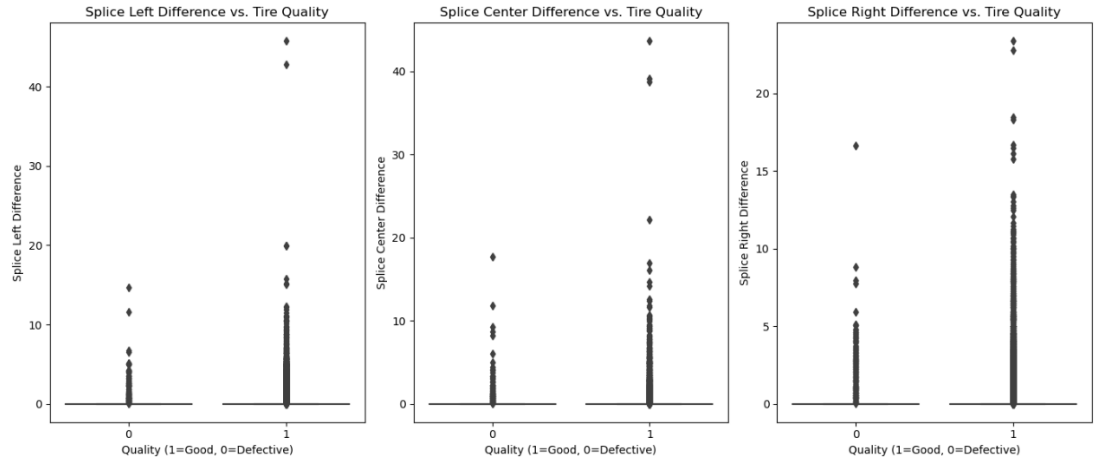
(a)



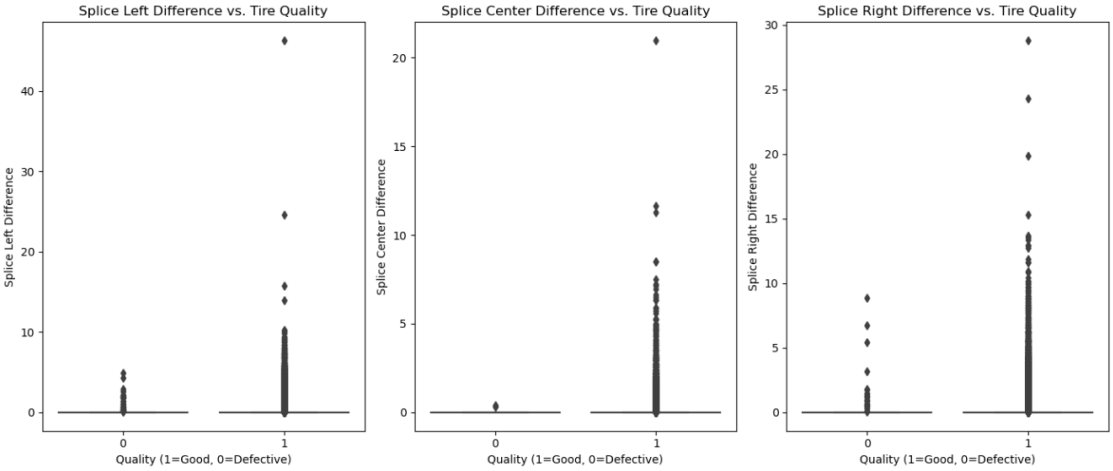
(b)

Figure 4.4 (a) Répartition des défauts par recette et (b) Taux de défauts par recette

La deuxième piste d'investigation consiste à examiner les variables de procédé directement liées à l'adhérence des plis de carcasse. Comme l'illustre la Figure 4.5, des diagrammes en boîte (boxplots) présentent la distribution des longueurs de joints observées par rapport à leurs limites de tolérance inférieure et supérieure. Ces visualisations mettent en évidence non seulement des écarts et la présence de valeurs aberrantes, mais également un chevauchement important entre les assemblages défectueux et non défectueux. Les deux classes comportent en effet des valeurs atypiques, ce qui suggère que le dépassement des limites de tolérance ne constitue pas, à lui seul, un indicateur suffisant pour prédire l'apparition de défauts d'adhérence. Cette observation souligne la complexité des mécanismes en jeu et justifie le recours à des méthodes d'ingénierie de variables enrichies ainsi qu'à des approches de modélisation prédictive allant au-delà des simples règles basées sur les tolérances.



(a) Écarts par rapport à la tolérance inférieure



(b) Écarts par rapport à la tolérance supérieure

Figure 4.5 Diagrammes en boites des écarts entre les longueurs réelles et les limites de tolérance

Dans une démarche similaire, la Figure 4.6 compare les valeurs de la consigne assignée à la machine d'assemblage aux valeurs réellement mesurées pour les longueurs des plis de carcasse. Cette comparaison met en évidence un décalage visible et systématique entre ce qui est attendu et le résultat retenu, suggérant des déviations récurrentes par rapport aux spécifications. Toutefois, il est crucial de noter que ces écarts, bien que présents, n'expliquent pas non plus de manière systématique la totalité des cas de défauts.

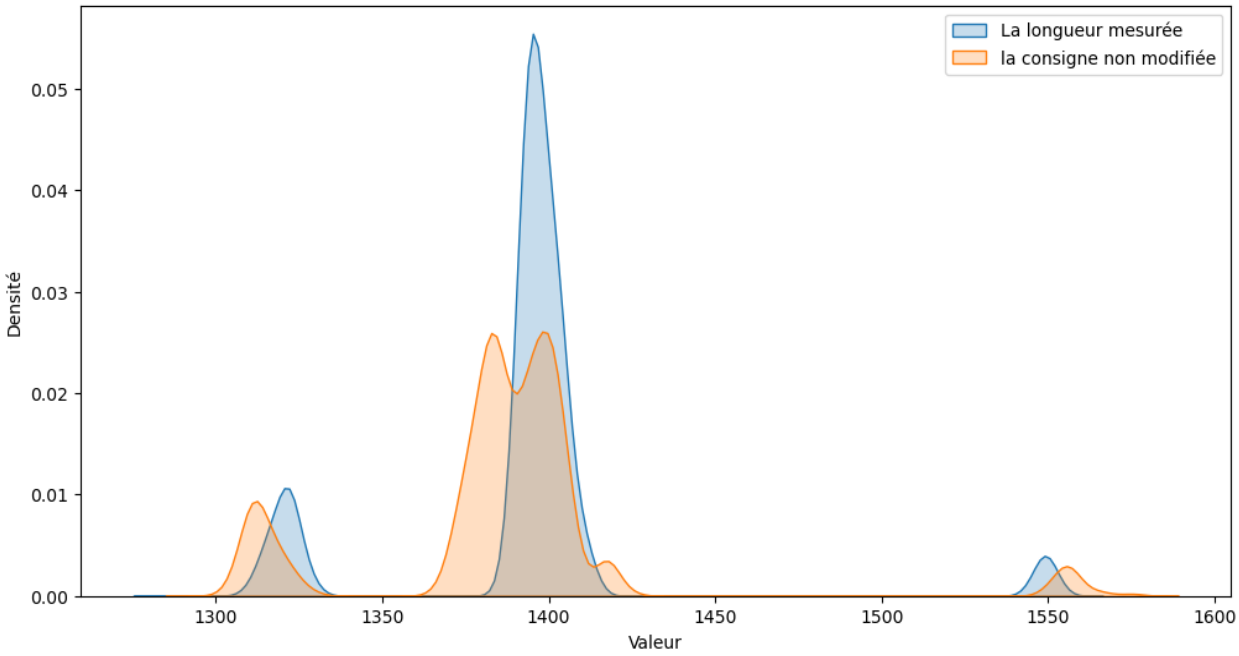


Figure 4.6 Comparaisons entre les distributions des longueurs consignes et les longueurs réelles de pli de carcasse

Les incohérences observées dans les mesures de processus nous ont conduits à explorer une piste alternative qui est l'influence des facteurs environnementaux. Une troisième hypothèse a donc été formulée, postulant un lien entre les variations de température saisonnières et la fréquence des défauts. Cette piste est particulièrement pertinente étant donné la localisation géographique de l'usine à Québec, une région soumise à des conditions climatiques extrêmes (hivers très froids, étés potentiellement chauds et humides).

En effet, le matériau central de notre étude, le caoutchouc, est connu pour ses propriétés de dilatation et de contraction thermique. Bien que l'usine s'efforce de maintenir une température stable à l'aide de systèmes de ventilation, une observation sur site a révélé que seules les zones de stockage des matières premières sont efficacement tempérées, laissant les lignes de production plus exposées aux variations ambiantes.

Pour tester cette hypothèse, une première analyse simple a consisté à examiner la répartition temporelle des défauts sur l'ensemble de 7 mois de l'année (Figure 4.7), à la recherche de pics saisonniers. Les résultats mettent en évidence une augmentation notable du nombre de défauts durant les mois de mai et juin, suggérant une possible influence des conditions environnementales

ou des paramètres de production propres à cette période. La baisse observée en juillet s'explique probablement par une interruption partielle des activités de production, ce qui réduit mécaniquement le volume d'observations. Ainsi, bien qu'aucune relation causale ne puisse être établie à ce stade, cette tendance indique une occurrence plus fréquente des défauts durant la période estivale, méritant une investigation plus approfondie sur les facteurs opérationnels ou environnementaux susceptibles de l'expliquer.

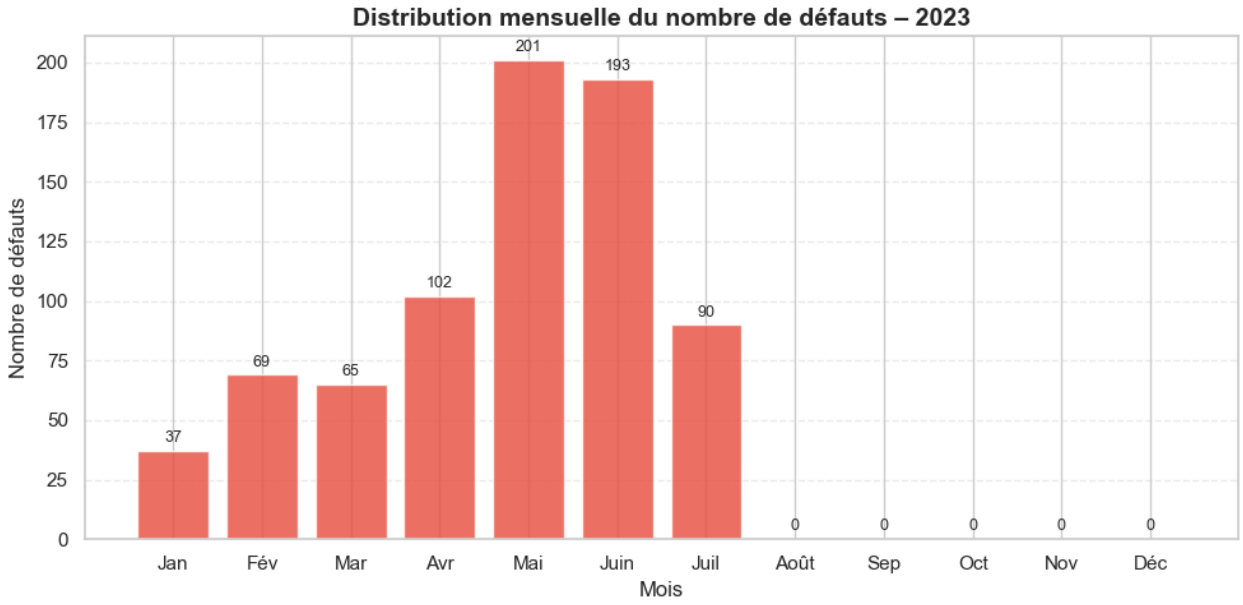


Figure 4.7 Distribution du nombre de défauts sur l'année

En synthèse, bien que les résultats de l'analyse exploratoire et des visualisations de données ouvrent des pistes potentielles d'investigation sur les défauts d'adhésion mais elles indiquent également que les défauts ne peuvent être distingués de manière fiable par de simples analyses univariées.

Cette conclusion souligne la nécessité d'une exploration plus approfondie des interactions entre les caractéristiques et des relations multivariées. C'est uniquement en examinant ces dépendances complexes que l'on peut espérer découvrir des prédicteurs plus informatifs et, in fine, améliorer la détection des défauts. L'échec des approches simples justifie ainsi pleinement le recours à la modélisation prédictive pour la suite de cette étude.

Analyse de corrélation

Dans cette étude, les corrélations entre les 52 variables explicatives initiales et la variable cible ont été calculées, puis représentées sous forme de carte thermique afin de faciliter leur interprétation. Aucune association linéaire forte avec l'étiquette de défaut n'a été mise en évidence, la corrélation absolue maximale observée étant de 0.37. Une carte thermique des corrélations inter-variables a toutefois révélé l'existence de regroupements marqués de colinéarité. Cette observation a conduit à l'élimination de trois variables fortement corrélées, à savoir les longueurs des deux plis de carcasse et la longueur du préassemblé de la première couche de carcasse, toutes présentant une corrélation absolue supérieure à 0.80, tout en conservant un représentant par cluster afin de limiter les effets de multi colinéarité. Les associations entre les identifiants catégoriels et la présence de défauts ont ensuite été examinées à l'aide d'un encodage *one-vs-rest* et de cartes thermiques. Le type de produit n'a montré aucune relation significative avec la variable cible, tandis que deux machines présentaient des corrélations positives modérées avec l'occurrence des défauts, avec des coefficients de 0.57 et 0.51 respectivement. Ces résultats demeurent exploratoires et feront l'objet de validations supplémentaires au moyen d'analyses stratifiées ou de modèles ajustés, afin d'écarter d'éventuels effets de confusion.

4.1.3 Préparation des données

A. Sélection des données

La sélection des données a d'abord été réalisée en collaboration avec le partenaire industriel, ce qui a permis d'identifier et de cibler les variables pertinentes liées aux étapes d'adhésion dans le processus d'assemblage. Dans un second temps, une analyse de corrélation a été conduite afin d'évaluer les relations entre les variables explicatives et la variable cible, ainsi qu'entre les variables elles-mêmes. La matrice de corrélation met en évidence que le coefficient de corrélation entre chaque variable et la variable cible ne dépasse pas 0,3, suggérant l'absence de dépendance linéaire forte. Néanmoins, l'ensemble des variables a été conservé, la collaboration avec l'expert ayant confirmé leur pertinence pour l'explication des défauts d'adhésion. Par ailleurs, l'analyse inter-variables a révélé de fortes corrélations entre certaines mesures relatives aux plis de carcasse, au pré-assemblage et aux longueurs de joints. Ces variables n'ont pas été éliminées, car l'étude porte sur deux types distincts de défauts d'adhésion, chacun étant associé à une couche spécifique de pli.

B. Nettoyage des données

Gestion des valeurs manquantes

La première étape du pipeline de prétraitement des données est consacrée au traitement des valeurs manquantes, qui peuvent résulter de pannes machine ou de dysfonctionnements de capteurs. Pour ce faire, une stratégie d'imputation en trois temps a été mise en œuvre. Premièrement, une méthode de remplissage par propagation avant (forward fill) a été appliquée, en partant de l'hypothèse que des produits consécutifs aux configurations de fabrication identiques sont susceptibles de présenter des valeurs de processus similaires. Cette méthode comble les lacunes en propageant les dernières valeurs valides connues. Deuxièmement, lorsque les produits adjacents différaient et que la propagation n'était pas applicable, une imputation par régression linéaire a été utilisée. Cette approche estime les valeurs manquantes en exploitant les relations entre les variables, apprises sur les observations complètes du jeu de données. Enfin, après l'application de ces deux stratégies, une faible proportion des données, soit environ 4 % des observations, contenait encore des valeurs manquantes. Ces lignes résiduelles, ne contenant aucun enregistrement associé à un défaut, ont été supprimées afin de garantir l'intégrité des données pour les étapes de modélisation subséquentes.

Détection des valeurs aberrantes

L'étape suivante du prétraitement se concentre sur la détection des valeurs aberrantes. Compte tenu de la complexité du processus d'assemblage semi-automatique et de la possibilité d'interventions humaines ou d'anomalies de capteurs, certaines valeurs enregistrées présentaient des déviations extrêmes. Pour détecter ces anomalies, une approche hybride combinant des méthodes statistiques et l'expertise métier des ingénieurs de procédé a été adoptée. Spécifiquement, la méthode des diagrammes en boîtes (boxplots) a été utilisée pour identifier les valeurs aberrantes sur l'ensemble des caractéristiques de mesure, y compris les valeurs de pression et les métriques d'alignement des couches. Cette procédure a révélé plusieurs valeurs extrêmes, jugées peu susceptibles de refléter des conditions opératoires normales. Ces observations ont ensuite été examinées au regard des seuils opérationnels définis par les experts et, si nécessaire, retirées du jeu de données pour éviter de biaiser le processus d'entraînement du modèle.

Encodage des variables catégorielles

Finalement, le prétraitement aborde le cas des variables catégorielles. Dans notre jeu de données, certaines de ces variables présentent un déséquilibre significatif, avec des catégories fortement surreprésentées par rapport à d'autres. Pour que le modèle puisse tenir compte de cette distribution, un encodage par fréquence (frequency encoding) a été appliqué.

Dans cette approche, chaque catégorie est remplacée par sa fréquence relative d'occurrence dans l'ensemble de données. Cette technique permet au modèle de conserver l'information relative à la prédominance ou à la rareté de chaque catégorie, ce qui peut constituer un signal prédictif utile.

C. Construction de données

Ingénierie des caractéristiques

À la suite des étapes de prétraitement, une phase d'ingénierie de caractéristiques (*Feature Engineering*) a été entreprise. L'objectif est de transformer les données brutes en un espace de caractéristiques qui capture la physique des mécanismes opérationnels, afin de permettre au modèle d'apprendre les relations subtiles menant aux défauts. Guidée par les connaissances des experts du domaine et les observations de l'analyse exploratoire, cette phase vise à extraire des motifs pertinents en incorporant une connaissance approfondie du processus de production directement dans le jeu de données.

Ingénierie des caractéristiques contextuelles

Dans cette étude, les caractéristiques dites "contextuelles" sont des variables qui encodent des événements opérationnels discrets et leur chronologie au sein du processus d'assemblage, traduisant ainsi des interventions ou des changements de configuration en prédicteurs mesurables. Parmi ces événements, on retrouve les réinitialisations d'équipement (equipment resets) et les interventions périodiques pour la reconfiguration du système de contrôle, qui peuvent introduire des instabilités transitoires dans le processus et potentiellement dégrader la qualité. Pour quantifier cet effet, deux caractéristiques temporelles ont été définies :

Compteur de Production Post-Réinitialisation (N_a) : cette variable mesure le nombre d'unités produites depuis la dernière réinitialisation de l'équipement. Elle est calculée comme suit :

$$N_a = n - n_0 \tag{4.1}$$

où n est le compteur de production cumulé et n_0 correspond au compteur au moment de la réinitialisation la plus récente. Cette variable reflète durée opérationnelle de la configuration machine actuelle.

Continuité Normalisée Post-Réinitialisation (Nc) : Pour contextualiser la valeur de N_a, nous introduisons une normalisation basée sur un volume de production de référence historique (N_{ref})

$$N_c = \frac{\mathrm{n} - \mathrm{n0}}{N_{ref}} \tag{4.2}$$

Ingénierie de caractéristiques basées sur le contrôle statistique des procédés

Comme introduit dans la méthodologie, une stratégie clé d'ingénierie de caractéristiques consiste à extraire des indicateurs de stabilité du processus à l'aide de cartes de contrôle statistiques. Cette partie vise à évaluer la pertinence de ces indicateurs, spécifiquement les signaux "hors de contrôle" issus des cartes X-barre et R (X-bar and R charts), comme prédicteurs de l'occurrence de défauts.

Pour surveiller la stabilité du processus, des cartes de contrôle ont été construites pour plusieurs variables continues critiques. Puisque ces variables suivent approximativement une distribution normale, nous avons appliqué une stratégie de sous-groupes rationnels, une pratique courante dans le cadre de la méthodologie Six Sigma. Cette approche a permis un suivi précis des variations à travers différents types de produits et de configurations de machines. Quatre caractéristiques critiques ont été sélectionnées en collaboration avec les experts du domaine, et des cartes de contrôle individuelles ont été générées pour chacune. La Figure 4.8 fournit un exemple visuel des cartes X-barre et R pour une caractéristique donnée, correspondant à une configuration produit-machine spécifique. Les points hors de contrôle, marqués en rouge, représentent des anomalies potentielles dans le processus d'assemblage.

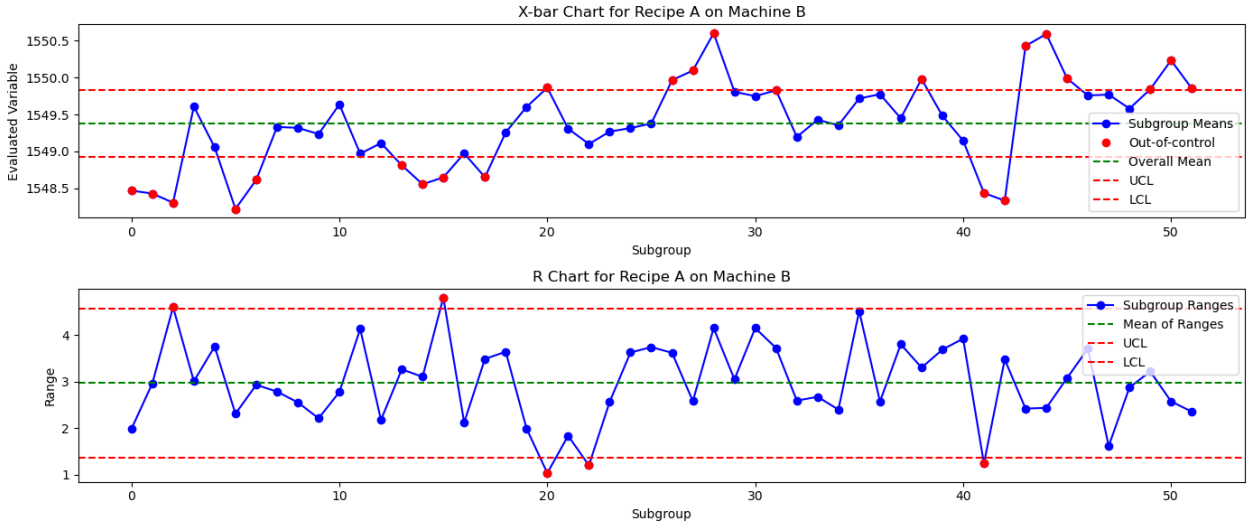


Figure 4.8 Cartes de contrôle X-barre et R pour les caractéristiques du processus d'assemblage

À partir de ces points, nous avons dérivé deux caractéristiques principales :

- Indicateur Hors de Contrôle (OOCi) : Une variable binaire, fixée à 1 lorsque la statistique du groupe *i* dépasse la limite de contrôle supérieure (*UCL*) ou tombe en dessous de la limite de contrôle inférieure (*LCL*), et à 0 sinon. Elle encode l'occurrence d'une violation de contrôle.

- Amplitude de l'Écart Hors-Contrôle : Une mesure de sévérité qui capture l'ampleur avec laquelle une observation hors de contrôle dépasse la limite la plus proche. Pour une statistique de carte *Zi*, la magnitude est calculée par

$$A_i = \max(0, Z_i - UCL, LCL - Z_i) \tag{4.3}$$

où :

- A_i: amplitude de l'écart hors-contrôle associée à l'observation i;
- Z_i: valeur normalisée (ou statistique de la carte de contrôle) pour l'observation i;
- UCL: Upper Control Limit, limite supérieure de contrôle ;
- LCL: Lower Control Limit, limite inférieure de contrôle ;

Pour donner suite à la construction de ces caractéristiques, une analyse de l'ensemble des points hors de contrôle a révélé qu'environ 460 produits défectueux figuraient parmi les échantillons hors de contrôle, représentant approximativement 61 % de tous les défauts critiques. Cette découverte

confirme la pertinence des violations de cartes de contrôle comme indicateurs de problèmes de qualité potentiels.

Cependant, il a également été observé qu'une proportion non négligeable d'environ 2000 de points hors de contrôle ne correspondait pas à des pièces défectueuses, suggérant que toutes les instabilités de processus ne conduisent pas à des défaillances critiques. Par conséquent, bien que les cartes de contrôle soient précieuses pour mettre en évidence les anomalies de processus, elles doivent être interprétées dans leur contexte et complétées par une modélisation additionnelle pour une prédiction efficace des défauts.

Construction d'enregistrements

Pour traiter le déséquilibre de classes extrême présent dans notre jeu de données, où seulement 757 échantillons défectueux étaient disponibles parmi plus de 300 000 instances une méthode hybride d'augmentation de données en deux étapes a été mise en œuvre, combinant le nettoyage de frontière et la modélisation générative.

La première étape a consisté à appliquer l'algorithme des Liens de Tomek (*Tomek Links*) afin d'éliminer les échantillons ambigus de la classe majoritaire situés près de la frontière de décision. Cette opération a légèrement réduit la taille du jeu de données d'environ 5 %, résultant en un espace de caractéristiques plus "propre", avec moins de chevauchement entre les classes. Bien que la séparation des classes ne soit pas visuellement marquée en raison du nombre écrasant d'instances appartenant à la classe majoritaire, l'effet des Liens de Tomek demeure perceptible. Cela se reflète notamment dans les performances du modèle, qui seront présentées à la section 4.1.4..

Dans un second temps, nous avons utilisé un GAN Tabulaire Conditionnel (*CTGAN - Conditional Tabular GAN*) pour générer des échantillons synthétiques réalistes et ainsi corriger le déséquilibre de classes résiduel. Le modèle a été entraîné sur 1 000 époques, avec une gestion des métadonnées adaptée à la nature mixte de nos caractéristiques. Pour ce faire, les variables catégorielles ont été traitées par encodage one-hot afin d'assurer une représentation équilibrée des catégories rares. Parallèlement, les variables numériques ont été normalisées à l'aide d'un modèle de mélange gaussien variationnel (*Variational Gaussian Mixture Model*), une approche capable de capturer des distributions multimodales et asymétriques. L'objectif était de créer un échantillon plus représentatif à la fois des produits conformes et des produits défectueux.

Cette phase d'augmentation synthétique a permis de générer 3 350 nouveaux échantillons de la classe "défectueux", augmentant ainsi la représentation de la classe minoritaire de 343 % tout en préservant les propriétés statistiques des données originales.

La qualité des données synthétiques a été évaluée à l'aide de la boîte à outils d'évaluation de *Synthetic Data Vault*. Le score de tendance des paires de colonnes ("*Column Pair Trends*") a atteint 88,91 %, et le score de qualité global s'est élevé à 86,62 %, indiquant un haut niveau de fidélité entre les jeux de données synthétique et original. La distribution post-CTGAN, visualisée dans l'espace des caractéristiques, révèle une classe minoritaire plus dense et mieux délimitée, suggérant une amélioration significative de sa représentation.

4.1.4 Résultats de détection de défauts sur la ligne d'assemblage

Cette section a pour objectif de présenter les résultats obtenus avec le modèle AdaBoost appliqué à la détection de défauts, puis d'en proposer une évaluation comparative avec différents modèles de référence afin de confirmer la pertinence de ce choix. Elle vise également à valider la méthode de rééchantillonnage retenue en la confrontant à d'autres approches issues de la littérature sur les données manufacturières.

A. Performance du modèle AdaBoost

Cette section présente une évaluation approfondie de l'approche proposée au moyen d'une série d'expérimentations comparatives visant à mesurer son efficacité dans le contexte de la détection de défauts sur données manufacturières déséquilibrées. Le Tableau 4.2 synthétise les différentes configurations de l'approche, évaluées à l'aide de quatre métriques de performance. Ces dernières ont été choisies afin de garantir une interprétation nuancée des performances respectives sur les classes majoritaire et minoritaire.

Le modèle de référence repose sur un classificateur AdaBoost, dont les hyperparamètres ont été optimisés par une recherche systématique sur grille, avec validation croisée stratifiée en cinq plis, selon l'espace de paramètres détaillé au Tableau 4.2.

Tableau 4.2 Espace de recherche des hyperparamètres

Hyperparamètre	**Plage de recherche**
max_depth	Entier : [5, 10, 15, 20]
min_split	Entier : [7,10,15,20]
learning_rate	Réel : [0.01, 0.015, 0.02] sélectionné à l'aide d'une approximation de type régression
n_estimators	Entier : [50, 100, 150, 200, 250, 300]

Par ailleurs, une optimisation supplémentaire a été effectuée en recourant à une méthode d'apprentissage sensible aux coûts (cost-sensitive learning). Cette approche ne modifie pas la distribution des données, mais ajuste l'algorithme lui-même pour qu'il accorde plus d'importance aux erreurs commises sur la classe minoritaire. Concrètement, un hyperparamètre scale_pos_weight (souvent appelé class_weight dans d'autres librairies) a été intégré au modèle. Cet hyperparamètre est calculé pour pénaliser de manière inversement proportionnelle à leur fréquence les erreurs sur chaque classe. Il est défini par le ratio suivant :

$$scale_{pos_{weight}} = \frac{Nombre\ total\ d'échantillons\ de\ la\ classe\ majoritaire}{Nombre\ total\ d'échantillons\ de\ la\ classe\ minoritaire} \quad (4.4)$$

Afin de garantir une représentativité adéquate des classes lors de l'évaluation, une division stratifiée du jeu de données a été effectuée selon un ratio de 70:30 entre apprentissage et test. Il convient de souligner que cette opération a été réalisée en amont du processus de rééchantillonnage. Ce choix méthodologique vise à préserver l'intégrité des données de test en les maintenant strictement réelles, sans inclusion d'instances synthétiques. De cette manière, l'évaluation des performances reflète de manière fidèle la capacité généralisable du modèle, évitant tout biais potentiel introduit par des données artificiellement générées.

Les résultats du modèle de référence sur l'ensemble de test ont affiché une précision globale proche de 100 % sur les données initiales non rééquilibrées. Toutefois, cette métrique s'est révélée

trompeuse. Une analyse complémentaire à l'aide de métriques plus adaptées, telles que le rappel de la classe « défaut », l'accuracy équilibrée, le G-mean et le F1-score, a mis en évidence des performances nettement plus faibles. Cette divergence illustre un problème classique dans les contextes de classification déséquilibrée, où le modèle tend à prédire systématiquement la classe majoritaire tout en négligeant la classe minoritaire. Bien que ce comportement conduise à une précision globale élevée, il empêche la détection des défauts critiques, rendant le modèle inopérant pour une utilisation pratique en détection de défauts.

Le tableau 4.3 présente les différentes configurations méthodologiques proposées afin de comparer les contributions respectives apportées au modèle de détection des défauts d'adhésion.

Tableau 4.3 Performance comparative des configurations de classification pour la détection de défauts

Méthode	**Exactitude équilibrée**	**Rappel (Défauts)**	**Moyenne F1-score**	**Moyenne Géométrique**
AdaBoost	0.63	0.27	0.54	0.52
Adaboost + CAF	0.66	0.34	0.53	0.58
Adaboost + CAF + SPCF	0.71	0.47	0.5	0.66
Méthode proposée: Adaboost + CAF + SPCF + Res	**0.845***	**0.69***	**0.84***	**0.83***

L'approche proposée intègre une ingénierie de variables fondée sur l'expertise métier, combinant des caractéristiques contextuelles (*Context-Aware Features*, CAF) et des variables basées sur le contrôle statistique du procédé (Statistical *Process Control Features*, SPCF), permettant de

caractériser l'instabilité des procédés en identifiant les comportements hors contrôle. L'introduction de ces variables construites a conduit à une amélioration notable du rappel de la classe « défaut », qui a atteint 47 %, traduisant une sensibilité accrue aux instances défectueuses. Toutefois, une légère baisse du F1-score a été observée, ce qui suggère que le modèle, bien qu'il identifie davantage de produits défectueux, a également généré plus de classifications erronées, illustrant ainsi le compromis classique entre précision et rappel dans les contextes déséquilibrés.

L'amélioration la plus marquée a néanmoins été obtenue grâce à l'application de la méthode de rééchantillonnage proposée (Res), qui a permis d'atténuer efficacement le déséquilibre de classes en renforçant la représentation des instances minoritaires. Cet ajustement a significativement accru la capacité du modèle à détecter les cas défectueux, comme en témoignent les gains notables enregistrés pour l'ensemble des métriques d'évaluation. Comme l'indique le Tableau 4.3, la méthode proposée surpasse systématiquement le modèle de référence et les approches comparatives en termes d'exactitude équilibrée, de rappel, de F1-score et de la moyenne géométrique, confirmant sa robustesse et son adéquation aux scénarios de détection de défauts déséquilibrés.

B. Analyse comparative des différentes méthodes conventionnelles de rééchantillonnage

Afin de valider l'efficacité de notre approche, des expérimentations complémentaires ont été menées pour comparer notre méthode de rééchantillonnage hybride à des techniques établies dans la littérature. Le Tableau 4.4 synthétise les performances globales de chaque méthode selon les principales métriques d'évaluation.

Tableau 4.4 Comparaison des performances entre la stratégie de rééchantillonnage proposée et les méthodes conventionnelles

Méthode	**Exactitude équilibrée**	**Rappel (Défauts)**	**Moyenne F1-score**	**Moyenne Géométrique**
SMOTE	*0.53*	*0.06*	*0.54*	*0.24*
ADASYN	*0.53*	*0.06*	*0.54*	*0.24*
SMOTEENN	*0.545*	*0.09*	*0.55*	*0.3*
NearMiss	*0.57*	***0.73****	*0.42*	*0.55*
Méthode proposée	***0.845****	*0.69*	***0.84****	***0.83****

Les techniques de sur-échantillonnage (oversampling) telles que SMOTE et ADASYN ont montré une efficacité limitée, avec des taux de détection de défauts n'atteignant que 6 %. Ce résultat indique leur incapacité à préserver et à exploiter les motifs informatifs de la classe minoritaire dans notre contexte. À l'inverse, l'approche de sous-échantillonnage NearMiss présente un rappel supérieur de 4 % par rapport au modèle proposé, soulignant son potentiel pour la détection d'événements rares. Cependant, cette amélioration s'est faite au détriment d'une chute substantielle de 42 % du F1-score. Ce déclin suggère que, bien que davantage d'instances défectueuses soient détectées, une proportion élevée des prédictions positives est incorrecte, le modèle se trompant dans plus de 50 % des cas lorsqu'il prédit la classe minoritaire. Une technique hybride combinant sur et sous-échantillonnage (SMOTE-ENN) a également été testée, mais elle a affiché de faibles performances en matière de détection de défauts, avec une tendance marquée à favoriser la classe majoritaire.

Ces résultats comparatifs démontrent les limites des approches de rééchantillonnage standards sur notre jeu de données et justifient, par contraste, la pertinence de notre méthode hybride personnalisée (Tomek Links + CTGAN) qui a permis d'obtenir un meilleur compromis entre la détection des défauts et la fiabilité globale des prédictions.

C. Analyse comparative des performances de différents modèles de classification de défauts

Après avoir validé notre stratégie de rééchantillonnage, l'étape suivante a consisté à évaluer et comparer les performances de plusieurs modèles de classification ensemblistes pour la tâche de détection de défauts. Cette analyse comparative, dont les résultats sont synthétisés dans le Tableau 4.5, vise à identifier le modèle offrant le meilleur compromis entre la sensibilité à la classe minoritaire et la fiabilité globale des prédictions.

Tableau 4.5 Évaluation des modèles de classification de défauts avec le modèle proposé

Modèle d'apprentissage automatique	**Exactitude équilibrée**	**Rappel (Défauts)**	**Moyenne F1-score**	**Moyenne Géométrique**
Random Forest	0.685	0.40	0.52	0.62
LightGBM	0.69	0.42	0.51	0.63
Balanced Random Forest	0.735	0.52	0.51	0.70
XGBoost	0.705	0.44	0.52	0.65
Méthode proposée	**0.845***	**0.69***	**0.84***	**0.83***

Parmi les modèles de référence, Random Forest et LightGBM ont obtenu les valeurs de rappel (Recall) les plus faibles (40-42 %), ne parvenant à détecter que moins de la moitié des cas défectueux. XGBoost, de son côté, a atteint un rappel comparativement meilleur, mais a affiché une faible moyenne géométrique (G-mean), ce qui indique une sensibilité limitée à la classe positive.

Le modèle Balanced Random Forest a démontré une performance globale solide, avec un rappel plus élevé pour la classe minoritaire, une meilleure exactitude équilibrée (Balanced Accuracy) et une G-mean supérieure. Cependant, son F1-score moyen de 51 % suggère un taux élevé de faux positifs. Cette faiblesse est particulièrement problématique dans un contexte manufacturier, où de

fausses alarmes peuvent entraîner des interruptions de processus inutiles et des coûts supplémentaires.

En comparaison, la méthode que nous proposons surpasse tous les autres modèles de référence, atteignant un F1-score moyen de 84 %. Cette performance reflète un équilibre efficace entre un taux de détection de défauts élevé et un faible taux de faux positifs. Comme l'illustre la matrice de confusion à la Figure 4.10, le modèle n'a classifié à tort que 74 produits non défectueux comme étant défectueux. Ce résultat met en évidence la robustesse et la pertinence opérationnelle de l'approche proposée, qui parvient à maximiser la capacité de détection tout en minimisant les perturbations coûteuses pour la production

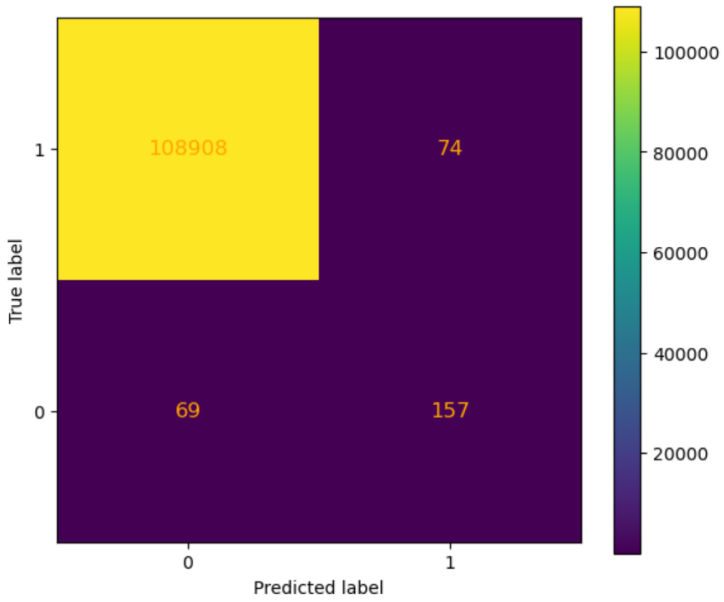


Figure 4.9 Matrice de confusion du modèle AdaBoost pour la classification de défauts

D. Analyse comparative des performances de différents modèles de détection de défauts combinés NearMiss

Les premières expérimentations ayant montré que les méthodes de rééchantillonnage influencent fortement la performance de détection des défauts, l'étape suivante consiste à évaluer la

combinaison des classificateurs de référence avec la technique la plus prometteuse, à savoir NearMiss. Les résultats, synthétisés dans le Tableau 4.6, sont comparés à ceux de la méthode proposée afin de mettre en évidence les différences en termes de capacité de détection et de robustesse.

Tableau 4.6 Comparaison des différents modèles de classification de défauts combinés à NearMiss

Modèle d'apprentissage automatique	**Exactitude équilibrée**	**Rappel (Défauts)**	**Moyenne F1-score**	**Moyenne Géométrique**
Random Forest + NearMiss	0.59	**0.81***	0.27	0.54
LightGBM+ NearMiss	0.585	0.79	0.28	0.54
Balanced Random Forest+NearMiss	0.61	0.80	0.30	0.57
XGBoost+NearMiss	0.575	0.73	0.30	0.55
AdaBoost+Nearmiss	0.51	0.68	0.30	0.53
Méthode proposée	**0.845***	0.69	**0.84***	**0.83***

L'examen des modèles de référence combinés à la méthode NearMiss révèle que, bien que les valeurs de rappel pour la détection des défauts soient relativement élevées (allant de 0.68 à 0.81), les F1-scores associés demeurent très faibles (0.27-0.30). Ce constat indique que ces approches ont tendance à sur-prédire la classe minoritaire, obtenant ainsi une meilleure sensibilité mais au prix d'un nombre excessif de fausses alarmes. Les exactitudes équilibrées, comprises entre 0.51 et 0.61, confirment par ailleurs la faible efficacité globale de ces modèles.

E. Évaluation comparative des stratégies prédictive temps réel en contexte manufacturier

Afin d'évaluer les implications pratiques de notre modèle prédictif dans le processus d'assemblage de pneus, deux stratégies distinctes ont été comparées en fonction de l'inclusion ou non des mesures liées à l'adhérence inter-plis à travers la Figure 4.11. Le Modèle 1 (*Alerte de Rebut*) intègre ces variables dans son espace de caractéristiques. Il permet ainsi la détection précoce des défauts d'adhérence avant la phase de cuisson, rendant possible la mise au rebut des pneus assemblés défectueux avant l'engagement de coûts additionnels liés aux étapes subséquentes de production. Dans cette configuration, le modèle agit comme un mécanisme d'alerte pré-cuisson, contribuant à la réduction des déchets et des coûts.

Le Modèle 2 (*Déclencheur de réparation*), en revanche, exclut les variables relatives à l'adhérence et repose uniquement sur les autres capteurs en amont ainsi que sur les paramètres de procédé. Cette approche s'inscrit dans une logique de réparation : la prédiction sert alors de déclencheur pour interrompre le processus, permettant une inspection manuelle ou une correction rapide, et évitant ainsi que des pneus défectueux ne progressent davantage sur la ligne.

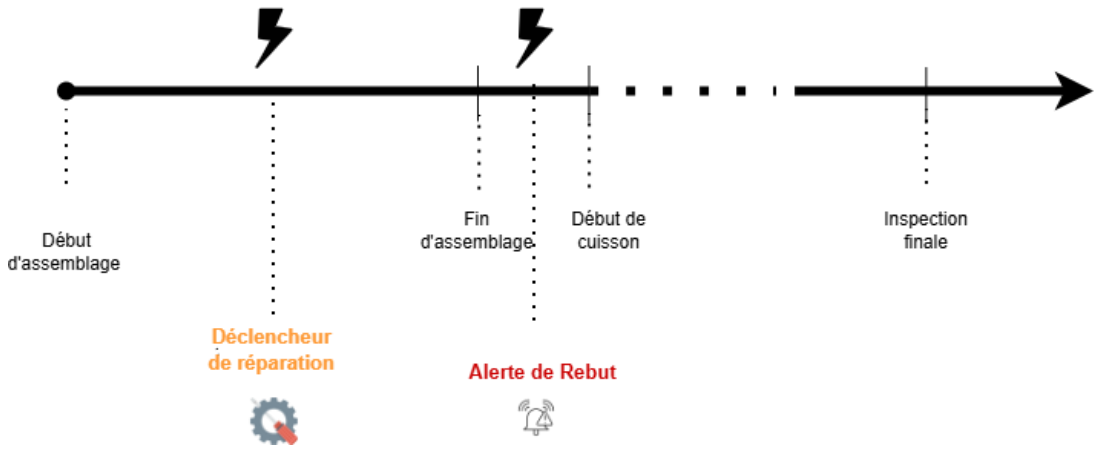


Figure 4.10 Comparaison de deux stratégies de modélisation selon l'inclusion des mesures d'adhérence inter-plis

Le Tableau 4.7 synthétise l'évaluation de ces deux stratégies prédictives. Les résultats mettent en évidence des compromis distincts entre performances prédictives et efficacité computationnelle. Le Modèle 1 atteint de meilleures performances globales, avec un F1-score de 0.84. Son temps moyen d'inférence, de 8.003 millisecondes, demeure compatible avec une intégration en temps réel

dans un système d'exécution de la production, sans retards significatifs ni perturbation du flux opérationnel, selon les experts du domaine. Le Modèle 2, bien que présentant des performances inférieures (F1-score de 0.52), conserve des résultats acceptables. Son principal avantage réside dans son efficacité computationnelle, avec un temps d'inférence moyen de 3.027 millisecondes par instance, très inférieur aux contraintes temporelles du processus d'assemblage (50 à 60 secondes par unité). De ce fait, il se révèle adapté aux stratégies orientées réparation, où l'objectif n'est pas la mise au rebut précoce mais plutôt la détection rapide de problèmes potentiels permettant une correction manuelle ou un ajustement du procédé.

Tableau 4.7 Analyse des performances des stratégies prédictives (*Alerte de Rebut* et *Déclencheur de réparation*)

Stratégie prédictive	**Paramètres de joint inclus**	**Exactitude équilibrée**	**Rappel (Défaut)**	**Moyenne f1-score**	**Moyenne Géométrique**	**Temps d'inférence (ms)**
Modèle 1 (Alerte de Rebut)	Oui	0.845	0.69	0.84	0.83	8.003
Modèle 2 (Déclencheur de réparation)	Non	0.805	0.65	0.52	0.63	3.027

Enfin, il convient de noter que la disponibilité effective des données et la capacité des modèles à soutenir une prise de décision en temps réel sont rarement abordées dans la littérature, rendant difficile l'évaluation de leur applicabilité dans des environnements industriels opérationnels.

4.1.5 Explicabilité du modèle

A. Analyse des facteurs de défauts globale

Afin d'améliorer l'interprétabilité du modèle prédictif, une approche globale a d'abord été appliquée. Les valeurs SHAP (*SHapley Additive exPlanations*) ont été calculées et analysées afin d'identifier les principaux facteurs contribuant à l'apparition de défauts. Il est important de préciser que, dans le cadre de notre configuration, l'étiquette *Target = 1* correspond aux produits défectueux, tandis que *Target = 0* désigne les produits conformes. En conséquence, des valeurs SHAP négatives contribuent à la prédiction d'un défaut, tandis que des valeurs positives indiquent une probabilité accrue de conformité.

Cette technique a été appliquée aux deux modèles prédictifs. Pour le modèle *Alerte de Rebut*, l'analyse SHAP offre aux opérateurs une visibilité sur les facteurs les plus fortement associés à l'apparition de défauts, leur permettant de surveiller proactivement ces aspects lors des cycles de production suivants. Pour le modèle *Déclencheur de réparation*, les résultats SHAP jouent un rôle de diagnostic, en mettant en évidence les variables de procédé les plus susceptibles d'être liées au défaut observé, guidant ainsi les opérateurs vers des actions correctives ciblées et efficaces.

Le graphique d'exemple SHAP présenté à la Figure 4.12 met en évidence les variables les plus influentes dans la décision du modèle. Les couleurs bleu et rose indiquent respectivement les valeurs faibles et élevées des variables, permettant ainsi de visualiser leur contribution positive ou négative à la prédiction. L'un des prédicteurs dominants est la distance par rapport à la limite de contrôle dans les cartes de contrôle, des écarts plus importants étant généralement associés à une probabilité accrue de classer un assemblage comme défectueux. Ceci peut ainsi être interprétée comme un indicateur précoce de dérive du procédé ou un signal annonciateur d'un besoin de recalibrage des équipements. Les caractéristiques liées aux opérateurs apparaissent également comme significatives : la présence récurrente de certains opérateurs semble corrélée à une probabilité plus élevée de défauts, suggérant de possibles erreurs non détectées ou des actions correctives incomplètes. Enfin, les mesures de pression, en particulier celles relevées aux extrémités du pli de carcasse (telles que *Pressure 2* et *Pressure 8*), montrent une forte association avec l'occurrence de défauts, des valeurs plus faibles à ces positions étant généralement liées à des assemblages défectueux.

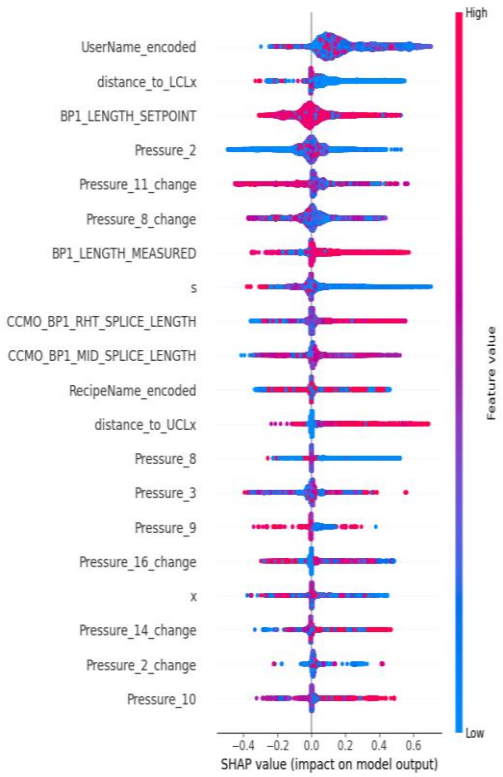


Figure 4.11 Exemple d'analyse globale de l'explicabilité du modèle à l'aide des valeurs SHAP

Une analyse plus fine, obtenue à partir des graphiques SHAP spécifiques aux variables, met en évidence des schémas plus nuancés est présenté par la Figure 4.13. Ainsi, pour les variables liées aux cartes de contrôle, les instances présentant des écarts importants par rapport à la limite de tolérance inférieure sont généralement associées à des valeurs SHAP négatives, suggérant une probabilité accrue d'apparition de défauts dans ces cas. De manière similaire, la Figure 4.13 illustre que les pressions brutes comme les pressions ajustées par les opérateurs suivent une tendance cohérente : des pressions comprises entre 0.1 et 0.3 bars sont le plus souvent corrélées à des valeurs SHAP négatives, indiquant qu'une augmentation des pressions contribue à la probabilité de survenue d'un défaut. Enfin, les variables associées aux changements d'équipement présentent également des valeurs SHAP négatives, ce qui suggère que ces transitions peuvent être liées à des instabilités transitoires du procédé, augmentant ainsi le risque de défaut.

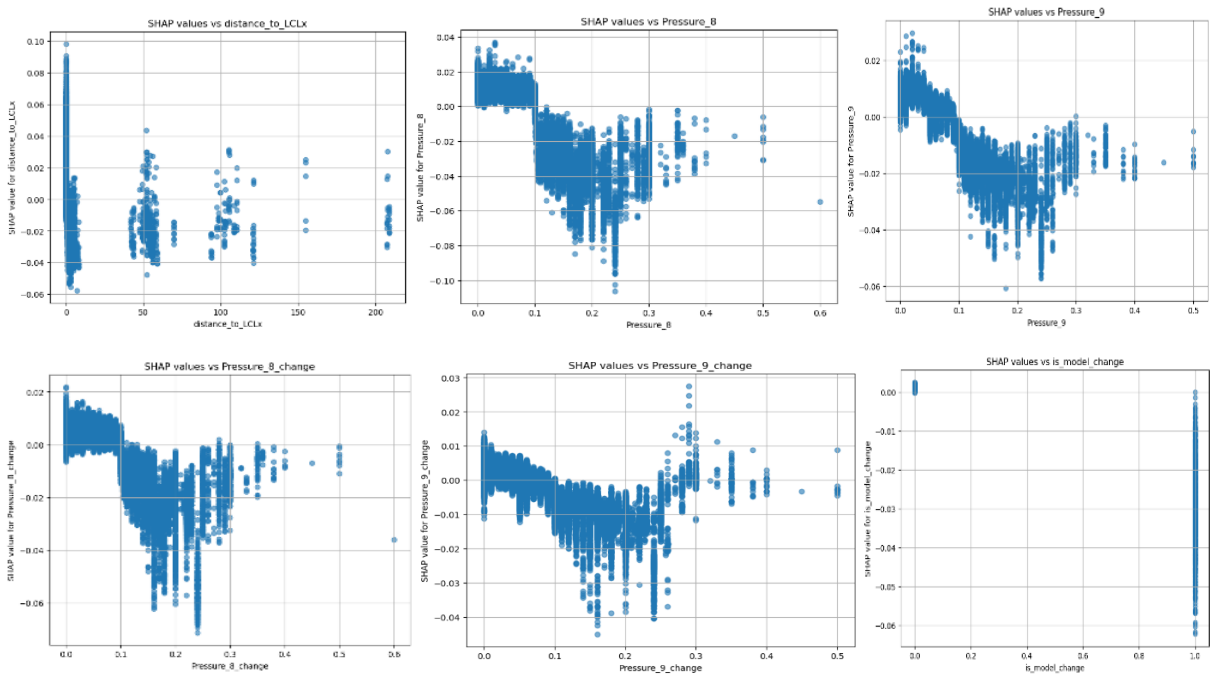


Figure 4.12 Analyse de dépendance par variable à l'aide des valeurs SHAP

B. Analyse des facteurs de défauts locale

En complément de l'interprétation globale du modèle à l'aide de SHAP, nous avons mobilisé LIME (*Local Interpretable Model-Agnostic Explanations*) afin de générer des explications au niveau des instances. Cette approche d'explicabilité locale s'avère particulièrement pertinente dans le cadre du modèle *Déclencheur de réparation*, où elle peut informer les opérateurs sur les paramètres les plus susceptibles d'avoir contribué à une prédiction spécifique de défaut. Comme l'illustre la Figure 4.14, LIME fournit non seulement la probabilité estimée qu'une instance donnée corresponde à un défaut, mais met également en évidence la contribution individuelle de chaque variable, en indiquant si elle accroît ou réduit la probabilité de survenue d'un défaut. Ce retour d'information en temps réel peut jouer un rôle essentiel dans la prise de décision des opérateurs, en leur permettant d'ajuster proactivement certains paramètres afin de limiter le risque de défauts.

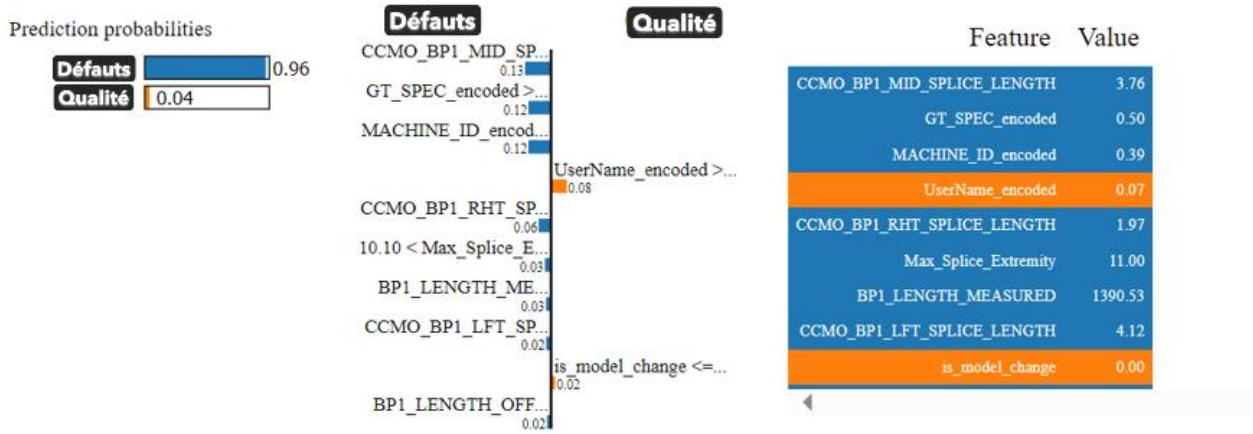


Figure 4.13 Explicabilité locale d'une observation

Conclusion

Ce chapitre a présenté la mise en œuvre pratique de la méthodologie développée dans le cadre de ce travail, appliquée à un cas d'étude industriel issu du secteur pneumatique. L'approche proposée a permis de démontrer la faisabilité et l'efficacité de la détection de défauts au sein d'une ligne d'assemblage semi-automatisée, en intégrant les différentes étapes de préparation, de modélisation et d'interprétation des résultats.

CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS

Ce chapitre final conclut ce projet de maîtrise, dont l'objectif était d'optimiser le contrôle qualité en temps réel sur une ligne d'assemblage semi-automatique. Il propose une synthèse des travaux réalisés et des principaux résultats obtenus, suivie d'une discussion sur les limites de la solution développée. Enfin, des pistes pour de futurs travaux et des améliorations potentielles seront présentées.

5.1 Synthèse des travaux

Ce travail a proposé et validé une méthodologie, inspirée du cadre CRISP-DM, pour la détection de défauts rares en contexte d'assemblage industriel. L'objectif général de ce mémoire était d'améliorer la capacité et l'efficacité du processus de contrôle qualité en temps réel au sein d'une ligne d'assemblage semi-automatique tout en répondant a deux sous objectifs spécifiques : celui de proposer un modèle d'identification de défauts en temps réel adapté au contexte de données fortement débalancées et celui d'identifier les facteurs d'influence de qualité.

Pour répondre à ces objectifs, une étude de cas a été menée en collaboration avec un partenaire industriel du secteur pneumatique. Le premier sous-objectif a été atteint à travers la conception d'un modèle prédictif performant intégrant les principes du Contrôle Statistique des Procédés (SPC) et des techniques d'apprentissage automatique adaptées à des données extrêmement déséquilibrées. Le modèle final a permis d'atteindre un taux de détection de 69 % des défauts au sein de la ligne d'assemblage, tout en maintenant un F1-score de 84 %, traduisant un équilibre optimal entre la détection efficace et la maîtrise des fausses alarmes.
Cette performance résulte d'un prétraitement rigoureux des données, d'une ingénierie de caractéristiques reflétant la dynamique du procédé, ainsi que d'une stratégie hybride de rééchantillonnage (Tomek Links + CTGAN), essentielle pour accroître la sensibilité du modèle à la classe minoritaire.

Au-delà de la performance prédictive, cette étude a fourni des renseignements opérationnels concrets répondant ainsi au deuxième sous objectif. L'analyse des facteurs les plus influents a permis d'identifier plusieurs causes potentielles de non-conformité, notamment :

- La nécessité d'un calibrage plus fréquent des machines pour contrer la dérive des processus.

- L'importance de vérifier la bonne application des corrections par les opérateurs lors des arrêts machine.
- Le besoin d'un ajustement plus fin des pressions pour garantir une adhésion optimale des matériaux.
- Une corrélation notable entre les changements de série et l'apparition de défauts, suggérant une vulnérabilité lors des phases de transition.

Ces découvertes offrent au partenaire industriel des leviers d'action directs pour améliorer la stabilité et la qualité de son processus de production.

5.2 Limites de cette étude

Cette étude présente néanmoins certaines limites qu'il convient de les analyser avec un regard critique. Tout d'abord, le périmètre de l'étude a été volontairement restreint à un seul type de défaut qui est la mauvaise adhésion des matériaux. Bien que ce choix ait été justifié par la criticité de cette non-qualité, ceci restreint la généralisabilité des conclusions à d'autres catégories de défauts.

Par ailleurs, bien que la méthodologie proposée permette de trouver un compromis entre la maximisation de la détection des défauts et la réduction des fausses alarmes, elle demeure entachée par un taux non négligeable de faux positifs. Sur l'ensemble des défauts détectés, 47 % des alertes étaient en réalité des faux positifs. Dans un contexte opérationnel, un tel ratio pourrait encore engendrer des coûts d'inspection et des interruptions de production non nécessaires.

En outre, si la méthode de suréchantillonnage employée prend en considération les variables catégorielles, la stratégie de sous-échantillonnage utilisée ne les intègre pas. Ceci pourrait entraîner la suppression d'échantillons de la classe majoritaire porteurs d'une information catégorielle importante, et donc une perte potentielle d'information.

Enfin, cette étude se positionne sur un plan prédictif et explicatif, mais non prescriptif. L'intégration d'outils d'explicabilité tels que SHAP et LIME permet d'identifier les facteurs associés à l'apparition des défauts, mais elle ne fournit pas de recommandations opérationnelles directes quant aux actions correctives à entreprendre comme « de combien ajuster la pression ? ».

5.3 Recommandations futures

Parmi les perspectives les plus immédiates, le déploiement de ce travail en environnement industriel réel s'impose comme une étape cruciale.

Cette validation en conditions opérationnelles permettra de confirmer la robustesse du modèle sur le long terme et de quantifier précisément son impact sur l'efficacité du contrôle qualité et les gains économiques associés.

Au-delà de cette application spécifique, l'élargissement de l'étude à d'autres industries ou à d'autres typologies de défauts renforcerait la généralisabilité de la méthodologie.

En réponse aux limites identifiées, des pistes d'amélioration méthodologiques sont également envisagées : l'exploration de méthodes de rééchantillonnage hybrides intégrant explicitement les variables catégorielles pourrait réduire la perte d'information et optimiser la représentativité des classes minoritaires.

Par ailleurs, le développement d'approches prescriptives, allant au-delà de la simple identification des facteurs associés aux défauts pour proposer des recommandations correctives concrètes, constituerait une avancée pour l'opérationnalisation du modèle.

Enfin, l'intégration de mécanismes d'apprentissage continu (Online Learning), capables de s'adapter dynamiquement aux évolutions du procédé et aux changements contextuels, renforcerait la pertinence et la durabilité des modèles en environnement industriel.

RÉFÉRENCES

Beckmann, M., Ebecken, N. F. F., & Pires De Lima, B. S. L. (2015). A KNN Undersampling Approach for Data Balancing. *Journal of Intelligent Learning Systems and Applications*, *07*(04), 104-116. https://doi.org/10.4236/jilsa.2015.74010

Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9584-9592. https://doi.org/10.1109/CVPR.2019.00982

Bersimis, S., Psarakis, S., & Panaretos, J. (2006). *Multivariate statistical process control charts : An overview*.

Brusa, E., Cibrario, L., Delprete, C., & Di Maggio, L. G. (2023). Explainable AI for Machine Fault Diagnosis : Understanding Features' Contribution in Machine Learning Models for Industrial Condition Monitoring. *Applied Sciences*, *13*(4), 2038. https://doi.org/10.3390/app13042038

Cação, J., Santos, J., & Antunes, M. (2025). Explainable AI for industrial fault diagnosis : A systematic review. *Journal of Industrial Information Integration*, *47*, 100905. https://doi.org/10.1016/j.jii.2025.100905

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection : A survey. *ACM Computing Surveys*, *41*(3), 1-58. https://doi.org/10.1145/1541880.1541882

Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. P. (2024). A survey on imbalanced learning : Latest research, applications and future directions. *Artificial Intelligence Review*, *57*(6), 137. https://doi.org/10.1007/s10462-024-10759-6

Deepika, Khamar Taj, & Parashuram Bedar. (2024). Automation in Production Systems : Enhancing Efficiency and Reducing Costs in Mechanical Engineering. *Nanotechnology Perceptions*. https://doi.org/10.62441/nano-ntp.vi.3895

de Giorgio, A., Cola, G., & Wang, L. (2023a). Systematic review of class imbalance problems in manufacturing. *Journal of Manufacturing Systems*.

Elshawi, R., Al-Mallah, M. H., & Sakr, S. (2019). On the interpretability of machine learning-based model for predicting hypertension. *BMC Medical Informatics and Decision Making*, *19*(1), 146. https://doi.org/10.1186/s12911-019-0874-0

Fan, S.-K. S., Cheng, C.-W., & Tsai, D.-M. (2022). Fault Diagnosis of Wafer Acceptance Test and Chip Probing Between Front-End-of-Line and Back-End-of-Line Processes. *IEEE Transactions on Automation Science and Engineering*, *19*(4), 3068-3082. https://doi.org/10.1109/TASE.2021.3106011

Fatemi, P., Sharifian, E., & Yassaee, M. H. (2025). *A New Approach to Backtracking Counterfactual Explanations : A Unified Causal Framework for Efficient Model Interpretability* (No. arXiv:2505.02435). arXiv. https://doi.org/10.48550/arXiv.2505.02435

Friedman, J. H. (2001). Greedy function approximation : A gradient boosting machine. *The Annals of Statistics*, *29*(5). https://doi.org/10.1214/aos/1013203451

Gao, Y., Yin, X., He, Z., & Wang, X. (2023). A deep learning process anomaly detection approach with representative latent features for low discriminative and insufficient abnormal data. *Computers & Industrial Engineering*, *176*, 108936. https://doi.org/10.1016/j.cie.2022.108936

Givnan, S., Chalmers, C., Fergus, P., Ortega-Martorell, S., & Whalley, T. (2022). Anomaly Detection Using Autoencoder Reconstruction upon Industrial Motors. *Sensors*, *22*(9), 3166. https://doi.org/10.3390/s22093166

Godina, R., Matias, J. C. O., University of Beira Interior, Covilhã, Portugal, Azevedo, S. G., & University of Beira Interior, Covilhã, Portugal. (2016). Quality Improvement With Statistical Process Control in the Automotive Industry. *International Journal of Industrial Engineering and Management*, *7*(1), 1-8. https://doi.org/10.24867/IJIEM-2016-1-101

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks* (No. arXiv:1406.2661). arXiv. https://doi.org/10.48550/arXiv.1406.2661

Gupta, H., & Madan, A. K. (2023). Reconfigurable Inspection in Manufacturing : State of the Art and Taxonomy. *2023 7th International Conference on Automation, Control and Robots (ICACR)*, 202-209. https://doi.org/10.1109/ICACR59381.2023.10314609

Haixiang, G. (2016). *Learning from class-imbalanced data : Review of methods and applications*.

Han, Y., Wei, Z., & Huang, G. (2024). An imbalance data quality monitoring based on SMOTE-XGBOOST supported by edge computing. *Scientific Reports*, *14*(1), 10151. https://doi.org/10.1038/s41598-024-60600-x

Harandi, M. A. Z., Li, C., Schou, C., Villumsen, S. L., Bøgh, S., & Madsen, O. (2023). STAD-FEBTE, a shallow and supervised framework for time series anomaly detection by automatic feature engineering, balancing, and tree-based ensembles : An industrial case study. *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 840-846. https://doi.org/10.1109/AIM46323.2023.10196288

He, M., Petering, M., LaCasse, P., Otieno, W., & Maturana, F. (2023). Learning with supervised data for anomaly detection in smart manufacturing. *International Journal of Computer Integrated Manufacturing*, *36*(9), 1331-1344. https://doi.org/10.1080/0951192X.2023.2177747

Herzog, T., Brandt, M., Trinchi, A., Sola, A., & Molotnikov, A. (2024). Process monitoring and machine learning for defect detection in laser-based metal additive manufacturing. *Journal of Intelligent Manufacturing*, *35*(4), 1407-1437. https://doi.org/10.1007/s10845-023-02119-y

Hotelling, H. (1947). Techniques of Statistical Analysis. In *Multivariate Quality Control* (p. 111-184).

Huang, H., Wang, P., Pei, J., Wang, J., Alexanian, S., & Niyato, D. (2025). *Deep Learning Advancements in Anomaly Detection : A Comprehensive Survey* https://doi.org/10.48550/arXiv.2503.13195

Hussain, M., & Zhang, T. (2025). Machine learning-based outlier detection for pipeline in-line inspection data. *Reliability Engineering & System Safety*, *254*, 110553. https://doi.org/10.1016/j.ress.2024.110553

Imani, M., Beikmohammadi, A., & Arabnia, H. R. (2025). Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels. *Technologies*, *13*(3), 88. https://doi.org/10.3390/technologies13030088

Kalgonda, A. A., & Kulkarni, S. R. (2004). Multivariate Quality Control Chart for Autocorrelated Processes. *Journal of Applied Statistics*, *31*(3), 317-327. https://doi.org/10.1080/0266476042000184000

Kausik, A. K., Rashid, A. B., Baki, R. F., & Jannat Maktum, M. M. (2025b). Machine learning algorithms for manufacturing quality assurance : A systematic review of performance metrics and applications. *Array*, *26*, 100393. https://doi.org/10.1016/j.array.2025.100393

Khan, S. S., & Madden, M. G. (2014). One-Class Classification : Taxonomy of Study and Review of Techniques. *The Knowledge Engineering Review*, *29*(3), 345-374. https://doi.org/10.1017/S026988891300043X

Khosravi, H., Farhadpour, S., Grandhi, M., Raihan, A. S., Das, S., & Ahmed, I. (2024). *Strategic Data Augmentation with CTGAN for Smart Manufacturing : Enhancing Machine Learning Predictions of Paper Breaks in Pulp-and- Paper Production*.

Klarák, J., Andok, R., Malík, P., Kuric, I., Ritomský, M., Klačková, I., & Tsai, H.-Y. (2024). From Anomaly Detection to Defect Classification. *Sensors*, *24*(2), 429. https://doi.org/10.3390/s24020429

Kosztyán, Z. T., & Katona, A. I. (2016). Risk-based multivariate control chart. *Expert Systems with Applications*, *62*, 250-262. https://doi.org/10.1016/j.eswa.2016.06.019

Kujawińska, A., & Vogt, K. (2015). Human Factors in Visual Quality Control. *Management and Production Engineering Review*, *6*(2), 25-31. https://doi.org/10.1515/mper-2015-0013

Lal Bhaskar, H. (2020). Lean Six Sigma in Manufacturing : A Comprehensive Review. In F. Pedro García Márquez, I. Segovia Ramirez, T. Bányai, & P. Tamás (Éds.), *Lean Manufacturing and Six Sigma—Behind the Mask*. IntechOpen. https://doi.org/10.5772/intechopen.89859

Li, Z., Yan, Y., Wang, X., Ge, Y., & Meng, L. (2025). A survey of deep learning for industrial visual anomaly detection. *Artificial Intelligence Review*, *58*(9), 279. https://doi.org/10.1007/s10462-025-11287-7

Lowry, C. A., & Montgomery, D. (1995). A review of multivariate control charts. *IIE Transactions (Institute of Industrial Engineers)*, 800-810.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A Multivariate Exponentially Weighted Moving Average Control Chart. *Taylor & Francis, Ltd.*, 46-53.

Lu, H. (2025). Evaluating the Performance of SVM, Isolation Forest, and DBSCAN for Anomaly Detection. *ITM Web of Conferences*, *70*, 04012. https://doi.org/10.1051/itmconf/20257004012

Lu, H., Du, M., Qian, K., He, X., & Wang, K. (2022). *GAN-Based Data Augmentation Strategy for Sensor Anomaly Detection in Industrial Robots*.

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* https://doi.org/10.48550/arXiv.1705.07874

Matharaarachchi, S., Domaratzki, M., & Muthukumarana, S. (2024). Enhancing SMOTE for imbalanced data with abnormal minority instances. *Machine Learning with Applications*, *18*, 100597. https://doi.org/10.1016/j.mlwa.2024.100597

Mattera, G., Mattera, R., Vespoli, S., & Salatiello, E. (2025). Anomaly detection in manufacturing systems with temporal networks and unsupervised machine learning. *Computers & Industrial Engineering*, *203*, 111023. https://doi.org/10.1016/j.cie.2025.111023

Melchore, J. A. (2011). Sound Practices for Consistent Human Visual Inspection. *AAPS PharmSciTech*, *12*(1), 215-221. https://doi.org/10.1208/s12249-010-9577-7

Mirzaei, M., Sadat, M. H., & Naderkhani, F. (2023). Application of Machine Learning for Anomaly Detection in Printed Circuit Boards Imbalance Date Set. *2023 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 128-133. https://doi.org/10.1109/ICPHM57936.2023.10193957

Molnar, C. (2019). *Interpretable Machine Learning*.

Montgomery, D. C. (2009). *Introduction to Statistical Quality Control* (7th éd.). https://books.google.ca/books/about/Introduction_to_Statistical_Quality_Cont.html?id=oh7zDwAAQBAJ&redir_esc=y

Muntasir Nishat, M., Faisal, F., Jahan Ratul, I., Al-Monsur, A., Ar-Rafi, A. M., Nasrullah, S. M., Reza, M. T., & Khan, M. R. H. (2022). A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset. *Scientific Programming*, *2022*, 1-17. https://doi.org/10.1155/2022/3649406

Oliveira, D. F. N., Vismari, L. F., Nascimento, A. M., De Almeida, J. R., Cugnasca, P. S., Camargo, J. B., Almeida, L., Gripp, R., & Neves, M. (2022). A New Interpretable Unsupervised Anomaly Detection Method Based on Residual Explanation. *IEEE Access*, *10*, 1401-1409. https://doi.org/10.1109/ACCESS.2021.3137633

Pang, G., Shen, C., Cao, L., & Hengel, A. van den. (2022). Deep Learning for Anomaly Detection : A Review. *ACM Computing Surveys*, *54*(2), 1-38. https://doi.org/10.1145/3439950

Park, & Wang, M. (2020). A Study on the X and S Control Charts with Unequal Sample Sizes. *Mathematics*, *8*(5), 698. https://doi.org/10.3390/math8050698

Park, Y.-J., Brito, P., & Ma, Y.-C. (2024). Anomaly detection-based undersampling for imbalanced classification problems. *Engineering Optimization*, *56*(12), 2565-2578. https://doi.org/10.1080/0305215X.2024.2315501

Pignatiello Jr, J. J., & Runger, G. C. (1990). Comparisons of Multivariate CUSUM Charts. *Journal of Quality Technology*, 173-186.

Pittino, F., Puggl, M., Moldaschl, T., & Hirschl, C. (2020). Automatic Anomaly Detection on In-Production Manufacturing Machines Using Statistical Learning Methods. *Sensors*, *20*(8), 2344. https://doi.org/10.3390/s20082344

Poslavskaya, E., & Korolev, A. (2023). *Encoding categorical data : Is there yet anything « hotter » than one-hot encoding?* https://doi.org/10.48550/arXiv.2312.16930

Psarommatis, F., May, G., Dreyfus, P.-A., & Kiritsis, D. (2020). Zero defect manufacturing : State-of-the-art review, shortcomings and future directions in research. *International Journal of Production Research*, *58*(1), 1-17. https://doi.org/10.1080/00207543.2019.1605228

Ramos, M., Ascencio, J., Hinojosa, M. V., Vera, F., Ruiz, O., Jimenez-Feijoó, M. I., & Galindo, P. (2021). Multivariate statistical process control methods for batch production : A review focused on applications. *Production & Manufacturing Research*, *9*(1), 33-55. https://doi.org/10.1080/21693277.2020.1871441

Riaz, M., & Muhammad, F. (2012). *An Application of Control Charts in Manufacturing Industry*.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). « Why Should I Trust You? » : Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144. https://doi.org/10.1145/2939672.2939778

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., & Muller, K.-R. (2021). A Unifying Review of Deep and Shallow Anomaly Detection. *Proceedings of the IEEE*, *109*(5), 756-795. https://doi.org/10.1109/JPROC.2021.3052449

Sauber-Cole, R., & Khoshgoftaar, T. M. (2022). The use of generative adversarial networks to alleviate class imbalance in tabular data : A survey. *Journal of Big Data*, *9*(1), 98. https://doi.org/10.1186/s40537-022-00648-6

Schapire, R. E. (2003). The Boosting Approach to Machine Learning : An Overview. In D. D. Denison, M. H. Hansen, C. C. Holmes, B. Mallick, & B. Yu (Éds.), *Nonlinear Estimation and Classification* (Vol. 171, p. 149-171). Springer New York. https://doi.org/10.1007/978-0-387-21579-2_9

Schlüter, H. M., Tan, J., Hou, B., & Kainz, B. (2022). *Natural Synthetic Anomalies for Self-Supervised Anomaly Detection and Localization* (No. arXiv:2109.15222). arXiv. https://doi.org/10.48550/arXiv.2109.15222

See, J. E. (2012). *Visual Inspection : A Review of the Literature*.

Seol, D. H., Choi, J. E., Kim, C. Y., & Hong, S. J. (2023). Alleviating Class-Imbalance Data of Semiconductor Equipment Anomaly Detection Study. *Electronics*, *12*(3), 585. https://doi.org/10.3390/electronics12030585

Shen, B., Yao, L., Jiang, X., Yang, Z., & Zeng, J. (2023). *Time series data augmentation classifier for industrial process imbalanced fault diagnosis*.

Shewart, W. A. (1931a). *Economic control of quality of manufactured product*.

Shewart, W. A. (1931b). *Economic Control of Quality of Mnaufactured Product*.

Shi, Y. (2023). An Imbalanced Data Augmentation and Assessment Method for Industrial Process Fault Classification With Application in Air Compressors. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, *72*.

Shrestha, N. (2021). Application of Statistical Process Control Chart in Food Manufacturing Industry. *International Journal of Engineering, Business and Management*, *4*(5), 82-87. https://doi.org/10.22161/ijebm.4.5.2

Taha, K. (2025). Observational and experimental insights into machine learning-based defect classification in wafers. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-024-02521-0

Talib, F., & Rahman, Z. (2012). Total quality management practices in manufacturing and service industries : A comparative study. *International Journal of Advanced Operations Management*, *4*(3), 155. https://doi.org/10.1504/IJAOM.2012.047634

Teixeira, E. L. S., Dias, T. S., Andrade, M. de O., & Oliveira, A. B. S. (2017). *COBEM-2017-2905 STATISTICAL PROCESS CONTROL APPLICATION IN AUTOMOTIVE INDUSTRY*.

Tsai, D.-M., & Jen, P.-H. (2021). Autoencoder-based anomaly detection for surface defect inspection. *Advanced Engineering Informatics*, *48*, 101272. https://doi.org/10.1016/j.aei.2021.101272

Wang, C., & Liu, J. (2021). An Efficient Anomaly Detection for High-Speed Train Braking System Using Broad Learning System. *IEEE Access*, *9*, 63825-63832. https://doi.org/10.1109/ACCESS.2021.3074929

Wei, J., Wang, J., Huang, H., Jiao, W., Yuan, Y., Chen, H., Wu, R., & Yi, J. (2024). Novel extended NI-MWMOTE-based fault diagnosis method for data-limited and noise-imbalanced scenarios. *Expert Systems with Applications*, *238*, 121799. https://doi.org/10.1016/j.eswa.2023.121799

Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a Standard Process Model for Data Mining*.

Woodall, W. H., & Ncube, M. M. (1985). Multivariate CUSUM Quality-Control Procedures. *Technometrics*, *27*(3), 285-292. https://doi.org/10.1080/00401706.1985.10488053

Xia, Z., Ye, F., Dai, M., & Zhang, Z. (2021). Real-time fault detection and process control based on multi-channel sensor data fusion. *The International Journal of Advanced Manufacturing Technology*, *115*(3), 795-806. https://doi.org/10.1007/s00170-020-06168-y

Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). *Modeling Tabular data using Conditional GAN* (No. arXiv:1907.00503). arXiv. https://doi.org/10.48550/arXiv.1907.00503

Yang, C., Fridgeirsson, E. A., Kors, J. A., Reps, J. M., & Rijnbeek, P. R. (2024). Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data. *Journal of Big Data*, *11*(1), 7. https://doi.org/10.1186/s40537-023-00857-7

Yang, Khorshidi, H. A., & Aickelin, U. (2024). A review on over-sampling techniques in classification of multi-class imbalanced datasets : Insights for medical problems. *Frontiers in Digital Health*, *6*, 1430245. https://doi.org/10.3389/fdgth.2024.1430245

Yaro, A. S., Maly, F., & Prazak, P. (2023). Outlier Detection in Time-Series Receive Signal Strength Observation Using Z-Score Method with Sn Scale Estimator for Indoor Localization. *Applied Sciences*, *13*(6), 3900. https://doi.org/10.3390/app13063900

Zhang, P., Jia, Y., & Shang, Y. (2022). Research and application of XGBoost in imbalanced data. *International Journal of Distributed Sensor Networks*, *18*(6), 155013292211069. https://doi.org/10.1177/15501329221106935

Zhang, Yang, D., & Wang, H. (2019). Data-Driven Methods for Predictive Maintenance of Industrial Equipment : A Survey. *IEEE Systems Journal*, *13*(3), 2213-2227. https://doi.org/10.1109/JSYST.2019.2905565

Zhou, X., Hu, Y., Wu, J., Liang, W., Ma, J., & Jin, Q. (2023). Distribution Bias Aware Collaborative Generative Adversarial Network for Imbalanced Deep Learning in Industrial IoT. *IEEE Transactions on Industrial Informatics*, *19*(1), 570-580. https://doi.org/10.1109/TII.2022.3170149

Zope, K., Singh, K., Nistala, S. H., Basak, A., Rathore, P., & Runkana, V. (2019). *Anomaly Detection and Diagnosis in Manufacturing Systems : A Comparative Study of Statistical, Machine Learning and Deep Learning Techniques*.

https://doi.org/10.36001/phmconf.2019.v11i1.815