



HAL
open science

Amélioration de la chaîne logistique de pièces de rechange en boucle fermée : application des modèles d'apprentissage

Hamza El Garrab

► To cite this version:

Hamza El Garrab. Amélioration de la chaîne logistique de pièces de rechange en boucle fermée : application des modèles d'apprentissage. Génie des procédés. Université d'Angers, 2020. Français. NNT : 2020ANGE0019 . tel-03210210

HAL Id: tel-03210210

<https://theses.hal.science/tel-03210210>

Submitted on 27 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE L'UNIVERSITE D'ANGERS

COMUE UNIVERSITE BRETAGNE LOIRE

Ecole Doctorale n° 602

Sciences pour l'Ingénieur

Spécialité : Génie Industriel

Laboratoire : LARIS

Amélioration de la chaîne logistique de pièces de rechange en boucle fermée : Application des modèles d'apprentissage

Thèse présentée et soutenue à Angers, le 18/12/2020

Unité de recherche : EA 7315 _____

Thèse N° : _____

Rapporteurs avant soutenance :

Samir LAMOURI Professeur à ENSAM Paris
Sylvie NORRE Professeur à Université d'Auvergne

Composition du Jury :

Président :	Prénom Nom	Fonction et établissement d'exercice (à préciser après la soutenance)
Examineurs :	Mitra FOULADIRAD	Professeur à Université de Technologie de Troyes
	Adnane LAZRAK	Responsable Supply Chain : Solutions Clients à GEHC
Dir. de thèse :	Bruno CASTANIER	Professeur à Université d'Angers
Co-dir. de thèse :	David LEMOINE	Maitre-assistant à IMT Atlantique

Invité(s)

HEIDSIECK Robert	Directeur du Service de Réparation Europe & EAGM à GEHC
CASTANIER Bruno	Professeur à Université d'Angers
LEMOINE David	Maitre-assistant à IMT Atlantique

Titre : Amélioration de la chaîne logistique de pièces de rechange en boucle fermée : Application des modèles d'apprentissage

Mots clés : chaîne logistique, pièces de rechange réparables, prévision, Machine Learning

Résumé : Dans le domaine de service après-vente et particulièrement dans la maintenance, l'intervention rapide et la réparation du bien du client est un élément clé pour sa satisfaction et pour la création de l'image de marque dans le marché.

Le travail présenté dans cette thèse propose une approche de type Machine Learning pour l'amélioration du flux informationnel de la chaîne logistique de pièces de rechange.

Notre contribution se focalise sur la prévision de la charge dans les centres de réparation des pièces de rechange, qui sont les fournisseurs principaux des pièces utilisées pour réparer les systèmes des clients.

La grandeur de la chaîne logistique et sa complexité, le grand nombre des références de pièces ainsi que la multitude des cas spéciaux (pays avec de lois spécifiques, pièces particulières...) fait que les approches classiques n'offrent pas des prévisions fiables pour les services de réparation.

Dans ce projet, nous proposons des algorithmes d'apprentissage permettant la construction de la connaissance à partir de grands volumes de données, au lieu des méthodes de prévision classiques.

Nous allons voir les modèles dans la littérature, présenter notre méthodologie, et ensuite implémenter les modèles et évaluer leur performance en comparaison avec les algorithmes existants.

Title: The improvement of the closed loop spare parts supply chain: Applying Machine Learning models

Keywords: supply chain, repairable spare parts, forecasting, Machine Learning

Abstract: In the field of after-sales service and particularly in maintenance, rapid intervention and repair of the customer's property is a key element for his satisfaction and for the creation of the brand image in the market.

The work presented in this thesis proposes an inspired Machine Learning approach for improving the information flow of the spare parts supply chain.

Our contribution is focused on forecasting the load in spare parts repair centers, which are the primary suppliers of spare parts used to repair customer systems.

The size of the supply chain and its complexity, the large number of part references as well as the multitude of special cases (countries with specific laws, particular parts, etc.) mean that traditional approaches do not offer reliable forecasts for the repair teams.

In this project, we propose learning algorithms allowing the construction of knowledge from large sets of data, instead of classical forecasting methods.

We will study models in the literature, present our methodology, and then implement the models and evaluate their performance in comparison with existing models.

Table des matières

Chapitre I : Contexte de la problématique industrielle : La chaîne logistique de pièces de rechange réparables en boucle fermée	27
I Introduction.....	27
II La chaîne logistique	28
1 Définition	28
2 La structure de la chaîne logistique.....	29
3 Les flux de la chaîne logistique.....	30
3.1 Le flux Physique	30
3.2 Le flux informationnel	31
3.3 Le flux financier.....	32
III La chaîne logistique en boucle fermée.....	33
1 Définition	33
2 Structure de la chaîne logistique en boucle fermée.....	34
3 Objectifs de la logistique en boucle fermée	34
3.1 La réutilisation des biens	34
3.2 Le traitement et recyclage.....	35
IV La chaîne logistique de pièces de rechange	36
1 Définition	36
2 Les particularités de la chaîne logistique de pièces de rechange	37
2.1 La nature de la demande :.....	37
2.2 La structure de la chaîne logistique de pièces de rechange :	38
2.3 Le nombre de références de pièces de rechange :.....	38
V La chaîne logistique en boucle fermée de pièces de rechange réparables	38
1 Les pièces de rechange réparables et consommables.....	39
1.1 Les pièces de rechange consommables.....	39
1.2 Les pièces de rechange réparables	40
VI Conclusion	41
Chapitre II : La chaîne logistique de pièces de rechange de General Electric Healthcare.	42
I Le conglomérat General Electric	42
II La branche General Electric Healthcare	43
III Le service de maintenance de GE Healthcare.....	43

1	La chaine logistique de pièces de rechange réparables de GEHC	44
2	Le service de réparation des pièces de rechange de GEHC :	45
3	Particularités de la chaine logistique de pièces de rechange de GEHC : complexité, volume et variabilité	46
3.1	Diversité du portfolio des pièces de rechange de GEHC :.....	46
3.2	La variation de la demande en fonction des pièces rechange et des régions	47
3.3	La compatibilité entre les pièces de rechange.....	47
IV	Description du processus de gestion du point de vue des centres de réparation et du point de vue du client.....	49
1	Le retour des pièces de rechange.....	50
2	La réparation des pièces de rechange.....	50
3	Le processus d’approvisionnement	52
4	Présentation de la problématique industrielle	53
V	Conclusion	55
Chapitre III : Positionnement scientifique de la problématique et contributions		57
I	Positionnement scientifique	57
1	Littérature de la problématique de prévision dans la chaine logistique de pièces de rechange	57
II	Méthodologies de prévision dans la littérature et l’utilisation des modèles d’apprentissage dans le contexte de la chaine logistique.....	59
1	La tendance croissante des contributions en application d’intelligence artificielle en chaine logistique	61
2	Analyses des approches présentes dans la littérature :	61
3	Analyses des différents cas d’utilisation des modèles :	63
4	Contributions de la recherche.....	64
4.1	Modélisation et mise en place d’un modèle Machine Learning de prévision de la charge dans les centres de réparation dans une CL-PR-BF	65
4.2	Implémentation d’un système de prévision de la charge, combinant des modèles Machine Learning Hybrides et de la simulation	67
4.3	Mise en place d’un système d’aide à la décision et de prévision pour le service de réparation de GEHC.....	67
III	Conclusion	68
Chapitre IV : Modélisation UML de la chaine logistique de pièces de rechange en boucle fermée		70
I	Introduction.....	70
II	Le langage de modélisation UML.....	70
1	Définition	70
2	Les différents diagrammes UML	71
2.1	Les diagrammes de structure	72

2.2	Les diagrammes dynamiques.....	73
3	Le langage UML en pratique	74
III	Approche méthodologique proposée : La méthodologie ASCI.....	74
1	L'approche ASCI dans la littérature.....	75
2	La démarche méthodologique	75
2.1	Analyse	76
2.2	Spécification	77
2.3	Conception	79
2.4	Implémentation	79
IV	Modélisation UML de la CL-PR-BF	80
1	Le sous-système physique.....	80
1.1	Les systèmes médicaux.....	82
1.2	La pièce de rechange.....	83
1.3	Les centres de distribution	84
1.4	Le transport de pièces de rechange	85
1.5	Les centres de réparation	86
2	Le sous-système logique	87
3	Le sous-système décisionnel	88
3.1	Les règles de gestion DRP	89
3.2	Les règles de gestion logistique	91
3.3	Les règles de gestion de réparation et d'intervention	92
V	Le modèle de simulation de la CL-PR-BF – Cas d'application : GEHC.....	92
VI	Le modèle de résultats de la simulation de la CL-PR-BF.....	103
1	Implémentation du modèle de simulation.....	103
2	Résultats du modèle de simulation sur les pièces de rechange de GEHC.....	106
2.1	Catégorie I : Les pièces avec l'achat seul (NB) comme solution optimale.....	106
2.2	Catégorie II : Les pièces avec la réparation seule (RP) comme solution optimale.	107
2.3	Catégorie III : Les pièces nécessitant une combinaison entre la réparation (RP) et l'achat de pièces neuves (NB).....	109
3	Performance du modèle pour la prévision de la charge de réparation	111
3.1	Processus de test de la performance du modèle.....	111
3.2	Métriques d'évaluation de la performance des modèles.....	112
3.3	Résultats.....	113
VII	Conclusion.....	120
Chapitre V : Modélisation et implémentation d'un model Machine Learning pour la prévision de la charge dans les centres de réparation dans une CL-PR-BF.....		123
I	Introduction.....	123

II	Notions de base en Machine Learning.....	124
1	Définition	124
2	L'apprentissage supervisée et non supervise	125
2.1	L'apprentissage supervisé.....	125
2.2	L'apprentissage non-supervisé.....	126
2.3	Autres méthodes.....	127
3	La classification et la régression	127
4	La classification.....	127
4.1	Définition	127
4.2	Evaluation de performance des modèles de classification.....	128
5	La régression	129
5.1	Définition	129
6	Le sur apprentissage et le sous apprentissage	130
III	Implémentation d'un modèle Machine Learning pour la prévision de la charge des centres de réparation dans une CL-PR-BF.....	130
1	Méthodologie d'implémentation de modèles d'apprentissage dans la littérature ...	130
2	Implémentation du modèle de résolution	131
2.1	La collecte de données	132
2.2	Le nettoyage de données	134
2.3	Apprentissage du modèle	139
2.4	Evaluation de la performance des modèles de régression	141
IV	Conclusion	151
Chapitre VI : Implémentation d'un modèle Machine Learning Hybride pour la prévision de la charge dans les centres de réparation dans une CL-PR-BF		
I	Introduction.....	155
II	Les modèles Machine Learning Hybrides dans la littérature.....	155
III	Implémentation du modèle Machine Learning Hybride pour la prévision de la charge de réparation dans une CL-PR-BF	157
1	La structure du modèle Hybride.....	157
2	Modèles additionnels pour la classification	158
2.1	KNN.....	158
2.2	SVM (Support Vector Machines)	159
2.3	Les arbres de décision.....	160
2.4	Les forets aléatoires	161
3	Processus d'implémentation.....	162
IV	Evaluation de la performance du modèle hybride	164
V	Conclusion	167
Chapitre VII : Conclusion, application industrielle et perspectives.....		
		169

1	Conclusion.....	169
2	Application industrielle des travaux chez GEHC	170
3	Perspectives.....	173

Table des figures

Figure I.1 : Les types de chaines logistiques (Mentzer <i>et al.</i> (2001))	29
Figure I.2 : Les flux de gestion de la chaine logistique (Tanyas <i>et al.</i> 2014).....	30
Figure I.3 Les éléments du fonds de roulement (schéma inspiré du modèle Hofmann <i>et al.</i> (2011)).....	33
Figure I.4: La forme générale d'une CL-BF (Khor <i>et al.</i> (2012)).....	34
Figure I.5: Le lien entre la CL-PR et le service de maintenance (Huiskonen <i>et al.</i> (2001)) ..	36
Figure I.6: La chaine logistique de pièces de rechange générique (Huiskonen <i>et al.</i> (2001))	37
Figure I.7 : Les types de pièces de rechange	39
Figure II.1: Les divisions de General Electric.....	42
Figure II.2: Structure et équipes du service de maintenance.....	44
Figure II.3: Les flux des pièces de rechange chez GEHC.....	49
Figure II.4: La structure d'un centre de réparation GEHC.....	51
Figure II.5 : Processus haut niveau de décision de planification	53
Figure III.1: Le processus de filtrage des publications des revues de littératures.....	60
Figure III.2: Nombre de publication en IA appliquée à la CL (Baryannis <i>et al.</i> (2019))	61
Figure III.3: Répartition des travaux en termes des variables d'entrée et de capacité à traiter des données de masse (Baryannis <i>et al.</i> (2019)).....	63
Figure III.4: Répartition des travaux par les capacités des modèles de prise de décision, de prévision et d'apprentissage (Baryannis <i>et al.</i> (2019)).	64
Figure III.5: Illustration de la phase de « training » d'un modèle Machine Learning (Gillian <i>et al.</i> (2010)).....	66
Figure III.6: La décomposition du problème et le type de modèles ML adaptés.....	66
Figure III.7: Combinaison entre les modèles Machine Learning et la Simulation	67
Figure IV.1: Domaines sémantiques d'UML (Kautz <i>et al.</i> (2018)).....	71
Figure IV.2: Processus de modélisation ASCI (Lemoine <i>et al.</i> (2008)).....	76
Figure IV.3: Présentation la distribution géographique et de l'expérience des participants au sondage (Torre <i>et al.</i> (2018))	77

Figure IV.4: Nombre de participants utilisant chaque langage de modélisation, repartis en fonction du domaine d'activité (académique et industriel) (Torre <i>et al.</i> (2018))	78
Figure IV.5: Interaction entre les trois sous-systèmes (Lemoine <i>et al.</i> (2008)).....	79
Figure IV.6: Le sous-système physique	81
Figure IV.7: Modélisation UML du système médical et les différentes entités qui y sont associe dans le sous-système physique	82
Figure IV.8: Modélisation UML (Diagramme de classe) de pièce de rechange et les entités qui y sont liées	84
Figure IV.9: Les centres de distribution	85
Figure IV.10: Les entités de transport des pièces de rechange.....	86
Figure IV.11: Les centres de réparation	87
Figure IV.12: Le sous-système logique	87
Figure IV.13: Vue globale des différentes règles de gestion du sous-système décisionnel ...	88
Figure IV.14: Les règles de gestion du modèle DRP pour une chaine logistique de pièces de rechange réparables.....	89
Figure IV.15: Les règles de gestion logistique	91
Figure IV.16: Les règles de gestion de réparation et d'interventions.....	92
Figure IV.17: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 1).....	94
Figure IV.18: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 2).....	95
Figure IV.19: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 3).....	96
Figure IV.20: Etapes d'exécution du modèle de simulation.....	105
Figure IV.21: L'outil de simulation en cours d'exécution. Le modèle simule les opérations pour toutes les pièces existantes dans la base de données, et offre la possibilité de visualiser l'évolution des différentes variables pour une pièce en temps réel.	105
Figure IV.22: Résultat du modèle de simulation sur la catégorie I des pièces de rechange (NB)	107
Figure IV.23: Résultat du modèle de simulation sur la catégorie II des pièces de rechange (RP).....	108
Figure IV.24: Niveau du BOH en baisse dû au taux d'échec des réparations.....	109
Figure IV.25: Résultat du modèle de simulation sur des pièces avec une combinaison RP, NB comme solution adéquate	110

Figure IV.26: Comparaison du résultat de la simulation avec l'historique des opérations de réparation de GEHC, pour une référence de pièces.....	116
Figure IV.27: Les volumes de réparation créées par le modèle de simulation en comparaison avec l'historique, par semaine (A) et par mois (B).....	117
Figure IV.28: Comparaison du volume de réparation cumulé entre la simulation et l'historique.....	118
Figure IV.29: Comparaison du résultat de la simulation avec l'historique des opérations de réparation de GEHC.....	118
Figure IV.30: Visualisation des écarts entre la vraie demande des pièces de rechange de GEHC et le résultat de la simulation.....	119
Figure IV.31: Visualisation de l'écart cumulé entre la vraie demande des pièces de rechange de GEHC et le résultat de la simulation.....	120
Figure V.1: Les applications principales de Machine Learning (Gao <i>et al.</i> (2020)).....	124
Figure V.2: Les principaux types de modèles Machine Learning (Gao <i>et al.</i> (2020)).....	125
Figure V.3: Les différents modèles d'apprentissage supervisé (Gao <i>et al.</i> (2020)).....	126
Figure V.4: Mode de fonctionnement des modèles de classification.....	127
Figure V.5: Grandeurs élémentaires pour l'évaluation des performances (Matrice de confusion) (Sauvanaud <i>et al.</i> (2016)).....	128
Figure V.6: Principe simplifié des modèles de régression.....	129
Figure V.7: Méthodologie d'implémentation de modèles d'apprentissage.....	131
Figure V.8: Application de la méthodologie d'implémentation des modèles d'apprentissage dans notre contexte.....	132
Figure V.9: Les processus de décision d'approvisionnement en se basant sur les données.....	134
Figure V.10: Volume en fonction de la solution d'approvisionnement avant la mise en forme.....	137
Figure V.11: Volume en fonction de la solution d'approvisionnement après la mise en forme.....	137
Figure V.12: La « boîte à moustache » et les points aberrants.....	138
Figure V.13: Résultat d'une régression polynomiale avec plusieurs degrés :1, 2 et 5.....	140
Figure V.14: Cycle de vie du produit, évolution de la base installée et demande de pièces de rechange (Dekker <i>et al.</i> (2013)).....	142
Figure V.15: Présentation des 500 références représentant 80% de la demande globale des pièces de rechange de GEHC.....	144

Figure V.16: Distribution des temps de chargement de chaque modèle sur l'échantillon des 500 références	145
Figure V.17: Résultats de la prévision du modèle de régression polynomiale, applique sur les données de GEHC.....	146
Figure V.18: La distribution de l'erreur de prévision du modèle de régression polynomiale.	146
Figure V.19: Résultats de la prévision du modèle de régression ridge, applique sur les données de GEHC.....	147
Figure V.20: La distribution de l'erreur de prévision du modèle de régression Ridge.....	147
Figure V.21: Résultats de la prévision du modèle de régression Lasso, applique sur les données de GEHC.....	148
Figure V.22: La distribution de l'erreur de prévision du modèle de régression Lasso.....	148
Figure V.23: Résultats de la prévision du modèle Machine Learning, applique sur les données de GEHC.....	150
Figure V.24: Distribution de l'écart entre la prévision et l'historique.....	151
Figure VI.1: Structure du modèle hybride (Banihashemi, Saeed <i>et al.</i> (2016)).....	156
Figure VI.2: Structure du modèle hybride proposé par Nguyen <i>et al.</i> (2020).....	156
Figure VI.3: Subdivision du processus de décision en plusieurs sous-décisions	157
Figure VI.4: Exemple d'une classification « multi-classe » avec une modèle KNN 2 dimensions, en prenant en considération k voisins, avec k allant de 1 à 6 (distance Euclidienne).....	159
Figure VI.5: Exemple d'une classification avec un modèle SVM binaire	160
Figure VI.6 : Structure basique d'un arbre de décision.....	161
Figure VI.7 : Illustration d'une foret aléatoire, composée de multiples configurations d'un arbre de décisions.....	161
Figure VI.8 : Processus d'implémentation du modèle Machine Learning Hybride.....	163
Figure VI.9 : Résultat de la prévision du modèle Machine Learning Hybride par rapport à l'historique	166
Figure VI.10 : Distribution de la valeur absolue de l'écart entre l'historique et les prévisions du modèle Machine Learning Hybride	166
Figure VII.1 : Dashboard d'accueil de la plateforme.....	171
Figure VII.2 : Dashboard d'accueil de la plateforme.....	171
Figure VII.3 : Visualisation de l'historique de la charge des différents ateliers du centre de réparation	172

Figure VII.4 : Module de gestion d'équipe du centre de réparation 173

Liste des algorithmes

Algorithme 1– Boucle de traitement de la demande dans les centres de distribution	97
Algorithme 2– Fonction : Traitement de l'expédition des commandes des techniciens	98
Algorithme 3 – Fonction : Actualisation des prévisions de la demande.....	98
Algorithme 4 – Fonction : Simulation des décisions d'approvisionnement	100
Algorithme 5 – Fonction : Traitement du transport des pièces de rechange	101
Algorithme 6 – Fonction : Simulation des commandes en cours, d'achat en neuf.....	103
Algorithme 7 : Calibration des modèles sur les différentes références de pièces.....	143

Liste des tableaux

Tableau III.1: Revues de littérature sur l'utilisation des modèles d'apprentissage en chaîne logistique.....	60
Tableau III.2: Distribution des méthodes utilisées dans la revue de littérature	62
Tableau IV.1: Présentation des différentes combinaisons des différentes parties du modèle de simulation avec leur performance (MSE, RMSE et R2).....	115
Tableau V.1: Exemple de lignes des données d'apprentissage	134
Tableau V.2: Exemple de données non représentatives de la problématique	135
Tableau V.3 : Nombre des observations dans l'échantillon de test avant et après les opérations de nettoyage.....	136
Tableau V.4: structure des données d'apprentissage.....	136
Tableau V.5 : Le nombre de références pour lesquelles chaque modèle a été choisi comme le meilleur.	149
Tableau V.6: Le minimum, maximum et la moyenne des scores de chaque modèle de prévision.....	149
Tableau VI.1 : Performance du modèle Machine Learning Hybride, avec les scores de chaque sous-modèle pour la phase de calibration et de test.....	165
Tableau VI.2 : Récapitulatif des résultats des modèles classiques, modèles Machine Learning et les modèles Machine Learning Hybride.....	167

Introduction Générale

Problématique

Dans le domaine de service après-vente, et particulièrement dans la maintenance qu'elle soit préventive ou corrective, l'intervention et la réparation du bien du client d'une façon rapide et efficace garantissent **la satisfaction client** et participent à la création d'une **image de marque** dans le marché de l'entreprise distributrice. La disponibilité et la qualité des pièces de rechange sont des éléments clé de cette stratégie tout en garantissant des coûts de gestion minimaux.

La réutilisation des pièces de rechange récupérées des systèmes clients est une pratique de plus en plus présente dans la stratégie de maintenance et impacte ainsi la chaîne logistique traditionnelle de pièces de rechange. Cette réutilisation est motivée, premièrement, par l'extension de la vie économique des biens, qui étaient initialement considérés comme des déchets et donc sans valeur ajoutée, en les transformant en des pièces de rechange de valeur utilisables par d'autres clients ; deuxièmement, par des raisons environnementales ou de réglementation, imposant une responsabilité de traitement de produits en fin de vie ; et, troisièmement, par l'amélioration de la disponibilité des pièces pour la maintenance notamment les pièces que l'organisation n'arrive plus à acheter en neuf ou bien qui sont impactées par la problématique d'obsolescence.

La réutilisation de pièces de rechange implique une analyse de leur état et leur possible remise en état de fonctionnement, vu qu'elles sont récupérées dans un état défectueux des systèmes chez les clients. Ce processus d'analyse et de réparation représente un challenge pour les industriels au vu du **nombre généralement important des références** de pièces de rechange dans chaque système, **l'interchangeabilité des pièces** entre différents modèles de systèmes ainsi que la **multitude des pannes possibles** pour chaque référence de pièces.

Les centres de réparation font ainsi face à un challenge de gestion de toutes ces variables afin de couvrir un nombre maximal de références et avoir une bonne capacité de réparation, tout en proposant une offre économiquement intéressante pour l'entreprise, c'est-à-dire offrir des pièces réparées moins chères que les pièces neuves, et ce en intégrant les aspects temporels de gestion. A ceci, on doit rajouter la gestion des compétences des équipes de réparation, les relations avec les fournisseurs des composants nécessaires pour la réparation de chaque panne et référence, ainsi que la gestion et la répartition de l'espace entre les ateliers.

Une des difficultés principales rencontrées dans notre contexte est **le manque de visibilité de la future charge dans les centres de réparation**, qui est un élément important dans la prise

de décision sur un horizon à moyen terme au vu des temps d'approvisionnement (plusieurs mois), la gestion des ressources humaines qui demande également une anticipation au préalable, la planification de la capacité de production, ainsi que l'ordonnancement des tâches dans les ateliers. Le manque de visibilité augmente un ensemble de risques financiers, comme l'obsolescence des composants et des matériaux de réparation : des dates d'obsolescence existent pour les produits (liquides chimiques, huiles...) et les composants utilisés dans la réparation. D'un côté, des décisions d'approvisionnement qui génèrent un excès sont la cause principale de l'obsolescence et de l'immobilisation d'inventaire représentant une perte financière pouvant atteindre des millions de dollars, pour une entreprise telle que General Electric Healthcare. D'un autre côté, un manque de matériaux ou de composants signifie une augmentation importante du temps de réparation et par conséquent un risque d'insatisfaction des clients.

Les modèles existants dans la littérature se rapportant à notre contexte, qui est la chaîne logistique de pièces de rechange en boucle fermée, ne permettent pas d'avoir des résultats fiables. Leur application pour la prévision de la charge des centres de réparation n'a pas permis de limiter leurs pertes financières importantes dans le passé. Ceci est dû à la nature de ces modèles basés sur des séries temporelles qui se reposent principalement sur la tendance de l'historique de la demande pour construire la prévision, qui n'est pas une approche adaptée à notre cas où la décision de réparation des pièces est basée sur un ensemble de paramètres, qui peuvent changer dans le temps. Ces algorithmes de prévisions ont également un ensemble de limites pour la détection des pics de la demande ou chutes brusques de la charge.

Un ensemble de modèles d'apprentissage est présent dans la littérature. Ils sont appliqués à des problématiques semblables à la nôtre et présentant des résultats plutôt prometteurs. Ces modèles ne se basent pas uniquement sur les tendances de la demande mais arrivent à intégrer les principaux paramètres qui influencent les données, en construisant une connaissance qui permet de prévoir à la fois la décision de réparation et sa quantité. Dans ce projet de thèse, nous proposons l'application des modèles d'apprentissage en les adaptant au contexte de la prévision de la charge en chaîne logistique de pièces de rechange, à travers une hybridation de différents types de modèles. Ainsi, une des contributions qui peut être vue dans cette thèse est l'application de ces approches avancées en apprentissage statistique dans un contexte opérationnel industriel et plus particulièrement pour la chaîne logistique de pièces de rechange chez General Electric Healthcare.

Contributions

Ce travail de recherche vise à mettre en place un modèle de prévision de la charge dans les centres de réparation, en se basant sur les algorithmes d'apprentissage (Machine Learning), dans une chaîne logistique de pièces de rechange, en boucle fermée. Il doit déboucher sur l'élaboration d'une démarche applicable dans les centres opérationnels de General Electric

Healthcare et ainsi garantir ce niveau d'opérationnalité, de déploiement et de prise en main dans l'entreprise.

Les contributions principales de ce travail sont les suivantes :

- **Une modélisation conceptuelle (par langage UML) de la chaîne logistique de pièces de rechange en boucle fermée, en se basant sur la méthodologie ASCI :** cette première étape de modélisation a permis de bien formaliser l'ensemble du processus de gestion des pièces de rechange, d'en appréhender sa complexité pour mieux comprendre l'ensemble des données et de leurs flux. Enfin, cette modélisation a aussi permis l'identification de premiers leviers d'amélioration principalement sur le champ des pratiques internes.
- **La mise en place d'un système de simulation pour la chaîne logistique de pièces de rechange en boucle fermée :** la deuxième étape de notre travail consiste à mettre en place un modèle de simulation pour comparer la logique d'approvisionnement théorique (type DRP) à la réalité des ordres de réparation placés pour les ateliers. Ceci a permis de constater, par une analyse systématique et quantitative, des écarts importants avec la simulation, et a ainsi motivé l'implémentation des modèles d'apprentissage.
- **La modélisation et l'implémentation d'un modèle Machine Learning de prévision de la charge dans les centres de réparation :** l'objectif est ici d'analyser le potentiel en termes d'applications et de performances des modèles classiques d'apprentissage par rapport aux données fournies par General Electric. L'analyse de performance nous a montré le bénéfice de ces approches par rapport aux approches de prévision classiques, mais aussi leur limite en termes de détection des pics et de chutes de la charge. Ceci a motivé l'élaboration d'un nouvel algorithme reposant sur une hybridation entre différents types de modèles.
- **La modélisation et l'implémentation d'un modèle Hybride Machine Learning pour la prévision de la charge de réparation :** dans cette partie nous proposons une nouvelle approche permettant de décomposer notre problème et appliquer un ensemble de modèles de classification et de régression pour prévoir à la fois la décision de réparation et la charge dans les centres de réparation dans une chaîne logistique de pièces de rechange.

Pour toutes les différentes contributions, nous avons l'avantage d'avoir des données réelles à disposition, qui sont les données historiques du service de réparation de General Electric Healthcare. L'ensemble de nos applications numériques, tests, simulations et évaluations de performance est effectué sur ces données.

Organisation du manuscrit

Ce manuscrit débute par la description du contexte industriel ayant motivé ce travail de thèse. Cette introduction permettra de classer le problème de gestion de pièces de rechange chez General Electric Healthcare par rapport aux problèmes classiques de chaînes logistiques, notamment les chaînes en boucle fermée introduites principalement dans le contexte de logistique inverse. Ainsi, une présentation rapide de ces chaînes sera proposée allant de la chaîne logistique classique, puis nous allons couvrir les aspects de la logistique inverse et les chaînes logistiques en boucle fermée (CL-BF), ses aspects et enjeux, pour ensuite passer à la chaîne logistique de pièces de rechange (CL-PR). Enfin, nous allons arriver à la chaîne logistique de pièces de rechange réparables qui combine les enjeux de la CL-BF et la CL-PR. Nous allons présenter sa structure, ces différents flux et ses particularités.

Dans le deuxième chapitre, nous allons introduire le contexte d'application industrielle, qui est le service de réparation des pièces de rechange de GEHC, entité accueillant ce projet de recherche. Nous allons voir le conglomérat GE et sa branche médicale GEHC et son service de maintenance après-vente pour enfin arriver au service de réparation. Nous présentons ensuite la formalisation de la problématique industrielle et l'impact de celle-ci sur les opérations de l'entreprise et sur la satisfaction des clients.

Un troisième chapitre va présenter la revue de littérature de cette problématique, qui présente les différents modèles de prévision existants et leurs limites, avant de faire une étude bibliographique des modèles qui semblent appropriés à notre problématique et proposer une adaptation et une application de ces derniers dans notre contexte.

Les 4^{ème}, 5^{ème} et 6^{ème} chapitres vont ensuite présenter nos contributions et plus précisément vont introduire respectivement la modélisation UML et le modèle de simulation, le modèle Machine Learning de prévision, et enfin le modèle Machine Learning Hybride.

Le 7^{ème} chapitre sera une conclusion qui reprend les contributions et les résultats principaux en termes de performance des différents modèles rencontrés dans cette thèse. Ces contributions ont été formalisées dans un outil d'aide à la décision mis en place chez GEHC, outil qui sera présenté en tant qu'application industrielle. Enfin, nous soulignerons quelques perspectives sur le champ industriel et scientifique de ce travail de thèse.

Chapitre I : Contexte de la problématique industrielle : La chaîne logistique de pièces de rechange réparables en boucle fermée

I Introduction

Les critères de disponibilité des systèmes de production deviennent prépondérants en termes d'optimisation industrielle tant les autres domaines tels que la gestion des stocks, la planification, l'ordonnancement ou encore la distribution des produits ont été l'objet de recherche et d'amélioration de performance. Aujourd'hui, ces aspects de disponibilité ont même des fins concurrentielles pour la vente de produits. On ne vend plus seulement le produit mais on cherche à contractualiser les aspects Service Après-Vente allant bien au-delà de la « simple » période de garantie. Ceci est d'autant plus vrai dans les secteurs industriels où les enjeux de disponibilité ou continuité de service qu'ils soient liés à une rentabilité économique, de sécurité ou de criticité de service – tels que le domaine médical – sont forts.

Les facteurs clé constitutifs de la disponibilité opérationnelle (en phase d'utilisation chez le client) sont, pour le fournisseur, les aspects maintenance et par-delà la logistique associée. On parle de Soutien Logistique ou encore de Soutien Logistique Intégré lorsqu'on analyse le produit au travers de ses performances intrinsèques (généralement associées à sa fiabilité et maintenabilité) et des performances globales organisationnelles. Comme souligné ci-dessus, la contractualisation en termes de niveau de disponibilité assurée chez le client amène de véritables enjeux économiques pour le fournisseur allant de la conception des produits et le choix de ses composants à la définition de l'ensemble de la chaîne logistique principalement des pièces de rechange en passant par la mise en place de stratégies de maintenance préventive. L'objet de cette thèse portera sur les aspects de chaîne logistique des pièces de rechange et particulièrement chez General Electric Healthcare (GEHC).

Les produits de GEHC jouent des rôles prépondérants dans le diagnostic et le traitement de patients, les rendant critiques en termes de continuité de service aux patients mais aussi en termes de rentabilité économique pour les organisations, généralement des hôpitaux et des centres de soins. Il en est de même pour les pièces qui les composent. Une stratégie de réhabilitation des composants défectueux offre de multiples avantages tels que la gestion de

l'obsolescence des différents sous-systèmes, une économie de renouvellement de stocks et une politique éco-responsable de l'entreprise. La notion de « réparabilité » au niveau du produit global et aussi au niveau des différents composants est alors à intégrer dans la gestion globale de la chaîne de soutien. On parle alors de chaîne logistique de pièces de rechange en boucle fermée (CL-PR-BF).

Dans ce chapitre, nous allons chercher à mettre en évidence les particularités de la chaîne logistique de pièces de rechange en boucle fermée (CL-PR-BF) par rapport aux chaînes logistiques traditionnelles de distribution de produits (CL) et aussi aux chaînes logistiques en boucle fermée normales (CL-BF). Rappelons ici que, bien que l'ensemble des coûts ne soit pas spécifié par souci de confidentialité, son optimisation est clairement l'objet permanent de nos recherches.

Ainsi, l'organisation de ce chapitre est comme suit. La première section va introduire une revue de littérature de la chaîne logistique normale. Dans la deuxième section, sera présentée la chaîne logistique en boucle fermée. La troisième section traitera de la chaîne logistique de pièces de rechange réparables et sa revue de littérature y sera présentée. La quatrième section va présenter la chaîne logistique des pièces de rechange en boucle fermée, qui est une combinaison des enjeux de la 2ème et la 3ème section.

II La chaîne logistique

Dans cette section, nous allons rapidement présenter un ensemble d'éléments constitutifs de la chaîne logistique. Ces définitions permettront ainsi au lecteur non spécialisé de positionner les enjeux élémentaires liés à la modélisation d'une chaîne logistique.

1 Définition

Le terme de chaîne logistique a été formellement défini par Oliver *et al.* (1982) comme « la collaboration systématique entre les personnes, les processus et les informations d'organisations pour créer des valeurs tangibles (Produit) ou intangibles (Service) et les livrer aux clients ».

Un ensemble d'autres définitions de la chaîne logistique est présent dans la littérature. Selon Fleischmann *et al.* (2003), une chaîne logistique est généralement définie comme le système consistant à déplacer un produit ou un service d'un fournisseur à un client. Boldoczki *et al.* (2020) définit une chaîne logistique comme : « un enchaînement d'organisations permettant de ramener des services ou produits au marché ».

Une publication visant à définir les particularités des chaînes logistiques par Mentzer *et al.* (2001) indique que la chaîne logistique est un ensemble de trois entités, ou plus, impliquées

dans les flux amont et aval de produits, services, finances et / ou informations d'une source à un client.

Lazrak *et al.* (2015) définit la chaîne logistique comme l'ensemble des activités et des flux physiques, informationnels et financiers permettant le transfert du bien ou du service, depuis le point de départ (fournisseur), jusqu'au point d'arrivée (client final). Le même travail précise qu'elle couvre également les activités et les flux inverses entre le client final et le fournisseur.

2 La structure de la chaîne logistique

La chaîne est un ensemble d'entités qui sont impliquées dans les activités et les fonctions supportant les flux des biens allant des fournisseurs originaux jusqu'au client final (Cooper M. C. *et al.* (1997)). Un processus de fabrication à une structure dans laquelle les matières premières sont transformées en produits finis, eux-mêmes livrés aux clients finaux (Benita M *et al.* (1998)).

Mentzer *et al.* (2001) propose trois types de structures de chaînes logistiques : « les chaînes logistiques directes », les « chaînes logistiques étendues » et les « chaînes logistiques ultimes » selon la complexité (Figure I.1).

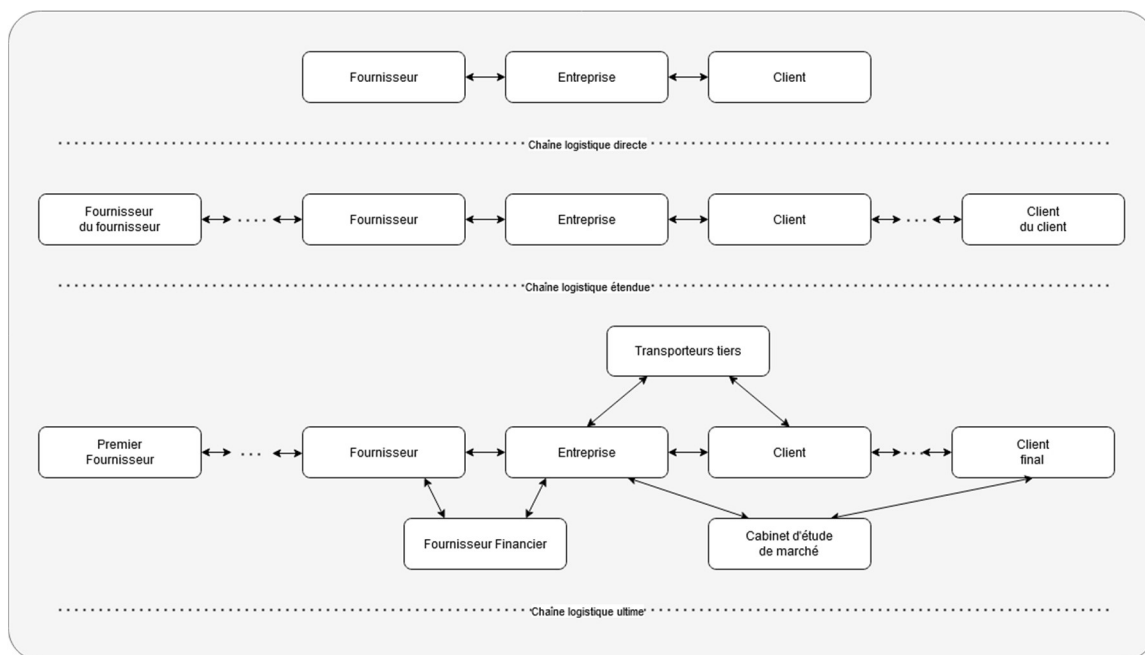


Figure I.1 : Les types de chaînes logistiques (Mentzer *et al.* (2001))

Des structures plus avancées comme par exemple la chaîne logistique en boucle fermée reposent sur ces architectures de base. Avant d'aborder ces structures avancées, nous proposons, dans la sous-section suivante, de présenter, à des fins pédagogiques, les différents flux de la chaîne logistique.

3 Les flux de la chaîne logistique

Il existe trois flux principaux dans une chaîne logistique : le flux physique, le flux informationnel et le flux financier. La figure I.2 présente ces différents flux et l'impact de leur optimisation sur le niveau d'inventaire de l'entreprise.

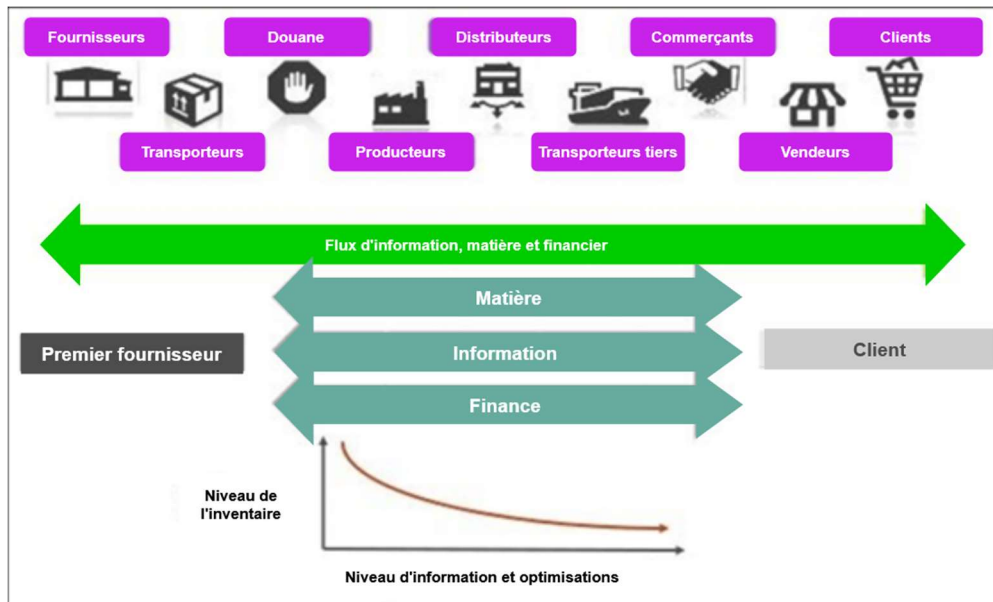


Figure I.2 : Les flux de gestion de la chaîne logistique (Tanyas *et al.* 2014)

3.1 Le flux Physique

Le flux physique est le premier élément constitutif d'une chaîne logistique et ainsi le plus analysé dans son contexte d'optimisation. Un flux physique représente la matière première, un produit semi-fini circulant d'un fournisseur vers un client intermédiaire, ou un produit fini en destination vers le client final.

Les principales opérations du flux physique sont :

- **L'approvisionnement** : Consiste à avoir des accords avec des fournisseurs afin de commander des biens ou services. Le processus implique la prise de décisions stratégiques d'achat adaptées à des conditions données.

L'objectif d'un approvisionnement optimal est de garantir une disponibilité des biens pour les clients, avec un inventaire minimal. Les opérations d'approvisionnement impliquent souvent des choix basés sur les délais de livraison et les prix de chaque fournisseur.

Les éléments clé (ou données d'entrée) de l'approvisionnement sont le niveau de stock actuel, le stock de sécurité, la prévision de la demande, la quantité minimale de commande et la taille de lot.

Pour les indicateurs de performance (ou éléments de sortie), ce sont des éléments permettant d'évaluer l'équilibre entre le niveau de stock et la satisfaction des besoins clients, comme le taux de couverture qui signifie le pourcentage des produits disponible aux clients actuellement. La rotation de stock est un autre indicateur de performance, représentant le nombre de jours moyen pour consommer l'inventaire. Plus le nombre de jours est faible, plus le stock est efficace.

- **La transformation** : Consiste à transformer des biens, qui peuvent être de la matière, des produits ou des services livrés par un fournisseur avant de les distribuer aux clients.

Un exemple de transformation est le traitement de la matière première tel que le bois, le fer ou autres pour produire des produits utilisables par les clients, (des meubles par exemple). La transformation peut être aussi présente dans le cas de services, tels que les services informatiques. Un hébergeur de fichier transforme les fichiers cryptés en les décryptant avant de les fournir au client final.

- **La distribution** : Représente les différentes opérations permettant le transport, le stockage, la distribution et l'entreposage des produits ou services dans le réseau logistique de l'entreprise.

La performance de la distribution des biens de l'entreprise a un rôle important dans la satisfaction client. Plusieurs indicateurs de la performance de distribution existent tels que le « Fill Rate » qui représente le taux de la demande client qui a été satisfaite à partir de l'inventaire local sans avoir besoin d'un réapprovisionnement.

3.2 Le flux informationnel

Le flux informationnel représente les données transitant dans la chaîne logistique du fournisseur vers le client ou dans le sens inverse.

Les données existent pour des objectifs de communication, coordination des activités de l'entreprise entre ses différentes entités ainsi que la planification et l'aide à la décision telle que les prévisions et la visualisation des métriques relatives à la performance.

Selon Botta-Genoulaz *et al.* (2005), on peut trouver dix catégories de données dans les systèmes d'informations des entreprises :

- **Les acteurs** (les caractéristiques, identification ou localisation d'une entité organisationnelle ou d'une ressource vue comme acteur)
- **Les règles** (les règles et les conditions du pilotage de l'activité)
- **Les plans** (prévisions, plannings et programmes qui déterminent un cadre de décision)

- **Les objets** (définition et caractérisation des produits, ressources, moyens, procédés et services conçus, réalisés et livrés)
- **Les ressources** (point de vue capacitaire des objets définis)
- **Les conditions** (les règles et les conditions du pilotage des relations)
- **Les décisions** (information caractérisant des décisions proposées, confirmées ou en cours de réalisation et impliquant des flux physiques ou financiers)
- **Les flux** (flux et événements liés aux relations : flux physiques, temps passés, anomalies, réceptions, livraisons, etc.)
- **Les journaux et la traçabilité** (données historiques, traces, listes d'événements datés et qui ne sont plus modifiés)
- **Les métriques** (données représentatives de l'activité aux niveaux stratégique, tactique et opérationnel)

3.3 Le flux financier

La coordination des flux physiques, des flux d'information et des flux financiers est aujourd'hui indispensable pour garantir la rentabilité économique, la satisfaction des clients et pour assurer la pérennité de l'entreprise (El Miloudi *et al.* (2015)).

Sans les flux financiers, les partenaires de la supply chain cesseraient de fonctionner, et la collaboration entre eux deviendrait impossible (El Miloudi *et al.* (2015)). Le flux financier (ou flux monétaire) circule dans le sens inverse du flux physique (Figure I.2). Il représente la valeur totale de ventes et d'achats dans une période comptable.

Dans leur publication, Ellram *et al.* (2004) ont précisé que dans une chaîne logistique, les flux financiers sont parallèles ou simultanés à tous les autres processus informationnels et physiques. Dans leur publication, les auteurs précisent que les paiements sont effectués périodiquement et que les échéances de paiement doivent être déterminées par le service gestion de la relation client ou fournisseur.

Les entreprises commencent de plus en plus à donner de l'importance pas seulement au chiffre d'affaire mais aux liquidités pures et au savoir de gestion de cette dernière afin d'arriver à leurs objectifs financiers, commerciaux et de production.

D'après Hofmann *et al.* (2011), cette liquidité, mesurée notamment par le ratio du fond de roulement, traduit l'aptitude de l'entreprise à transformer les actifs circulants (créances) en liquidité afin de faire face aux dettes à court terme. Le besoin en fonds de roulement (BFR) est donc calculé comme étant la différence entre l'actif circulant (AC) et le passif circulant (PC) : $BFR = AC - PC$ (Figure I.3).



Figure I.3 Les éléments du fonds de roulement (schéma inspiré du modèle Hofmann *et al.* (2011))

III La chaîne logistique en boucle fermée

Depuis ces dernières années, une chaîne logistique n'est plus uniquement considérée comme un flux où les produits sont créés, mais aussi utilisés et puis recyclés ou disposés dans un flux en boucle. Dans cette section nous allons présenter la chaîne logistique en boucle fermée, qui, en comparaison avec une chaîne logistique normale, prend en charge également le flux inverse de retours des produits.

1 Définition

Pour des raisons stratégiques et économiques pour les entreprises, ainsi qu'une prise de conscience croissante de l'impact environnemental dans l'industrie, les chaînes logistiques en boucle fermée attirent de plus en plus l'attention des industriels, ainsi que celle des académiques (Peng, Hui *et al.* (2020)). On appelle Chaîne Logistique en Boucle Fermée (CL-BF) la chaîne logistique combinant un flux direct, allant des fournisseurs vers les clients, ainsi qu'un flux inverse remontant les biens utilisés, soit sous forme de retours simples, des produits en panne ou en fin de vie.

L'économie en boucle fermée est une stratégie qui a émergé pour mettre fin au système conventionnel ouvert, visant à relever le défi de la rareté des ressources et de l'élimination des déchets dans une approche économiquement gagnant-gagnant (Homrich *et al.* (2018)). L'étude des chaînes logistiques en boucle fermée a montré l'importance de ces dernières dans la réduction de la consommation de la matière première, des émissions, de la consommation

énergétique et la production des déchets. Elle engage d'autres opérations telles que la réparation de produits défectueux, la réutilisation ou la remise en états des produits déjà utilisés mais qui ne sont pas forcément défectueux (Van Wassenhove *et al.* (2008)).

2 Structure de la chaîne logistique en boucle fermée

La chaîne logistique en boucle fermée est une structure intégrant la logistique inverse et la chaîne logistique classique. La chaîne logistique en boucle fermée consiste d'un flux direct allant des fournisseurs de matière première vers les usines d'ajout de la valeur jusqu'au client final, et d'un flux « inverse » retournant des clients vers l'entreprise, ou même les fournisseurs de matière première dans le cas de recyclage (Figure I.4).

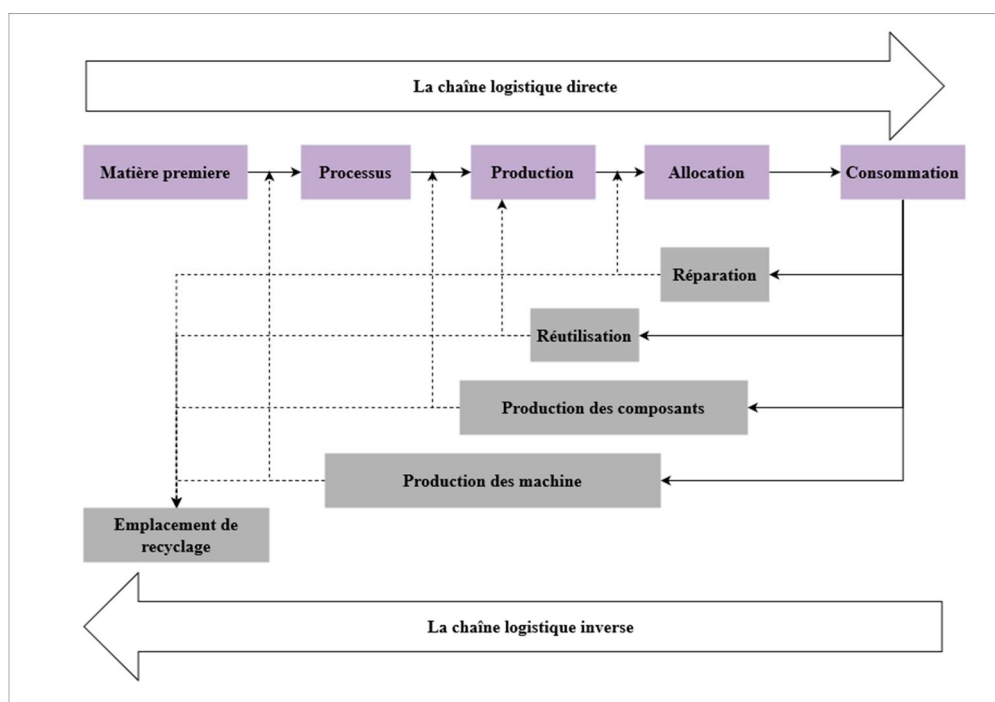


Figure I.4: La forme générale d'une CL-BF (Khor *et al.* (2012)).

Nous proposons dans la section suivante d'en préciser certaines caractéristiques et d'en souligner les spécificités par rapport aux chaînes logistiques classiques.

3 Objectifs de la logistique en boucle fermée

3.1 La réutilisation des biens

Une des motivations principales de la logistique inverse est l'extension de la vie économique des biens via leur récupération et éventuelle réutilisation en fin de vie. Ainsi, on assiste à une transformation d'un bien considéré initialement comme un déchet en un bien de valeur qu'il faudra chercher à réestimer. Il peut ainsi être économiquement intéressant pour des industriels de définir des stratégies commerciales orientées client afin de récupérer leurs produits en fin de vie. Parmi ces propositions commerciales, on note la remise à niveau (remplacement du

produit du client par une nouvelle génération avec prix réduit) ou le rachat du produit (ou l'entreprise rachète le système en fin de vie).

Les stratégies de réutilisation des entreprises diffèrent par domaine et par produit. La revente du produit reconditionné est une stratégie très présente par exemple, qui permet aux clients d'acquérir le produit reconditionné sans avoir à payer le prix original, ce qui permet d'élargir les bases clientèles intégrant des clients plus modestes et/ou à aspiration écologiste.

La réutilisation des pièces de rechange est aussi une stratégie assez classique dans laquelle l'entreprise récupère la totalité ou une partie du système, afin de le démonter et récupérer les composants importants y appartenant. L'entreprise remet les pièces de rechange en état de fonctionnement et les réinjecte dans son inventaire afin de pouvoir maintenir sa base installée pendant une durée plus longue.

Pour les systèmes en fin de vie économique, cette stratégie permet de compenser le retrait de certains fournisseurs de pièces de rechange. Ceci permet de gérer au mieux les problèmes d'obsolescence, d'accroître la durée de vie rentable des produits et par-delà accompagner progressivement l'entreprise dans ses transitions technologiques.

3.2 Le traitement et recyclage

Pour des raisons environnementales ou de réglementation, les industriels ont une responsabilité croissante du traitement de leurs produits en fins de vie (l'écotaxe sur les produits électroniques en est un exemple). Il est ainsi nécessaire de mettre en place des processus de récupération des biens et des produits pour ensuite les traiter. Pour des produits courants, on peut ainsi mettre en place des stratégies incitatives aux clients (généralement des particuliers) ou encore la mise à disposition de bacs de récupération. Lorsque les produits sont soumis à autorisation réglementaire ou encore conservent un intérêt technologique ou économique, il est nécessaire d'en optimiser sa collecte directement auprès des différents clients. Notons ici que, dans ce cas, l'opération associée peut être vue comme une source de coût à l'entreprise et non de véritable profit, sauf s'ils offrent une possibilité de réutilisation ou représentent de la valeur (par exemple, la récupération de l'or des microprocesseurs ou du cuivre des câbles électriques).

Sans chercher à faire une analyse bibliographique exhaustive à ce niveau, on peut dire que la prise en compte des flux inverses dans la définition et la gestion de la chaîne logistique vient clairement complexifier les modèles décisionnels avec :

- Une explosion des paramètres de la chaîne logistique ;
- L'interdépendance de certains de ces paramètres de contrôle, un niveau de stock pouvant dépendre de la stratégie de récupération et des traitements des biens et produits ;

- Une incertitude croissante sur ces paramètres liée à cette capacité de récupération et de traitement des produits fonction, pour cette dernière, de leur état de fonctionnement lui-même inconnu avant un diagnostic profond ;
- Une faible valeur ajoutée à court terme d'une telle approche avec un retour sur investissement à plus long terme.

Dans le contexte de notre projet de thèse chez General Electric Healthcare, les deux utilisations (réutilisation et recyclage) sont présentes, en fonction du type du produit. Cependant, au vu du service de réparation dans lequel s'effectue ce projet, nous allons nous focaliser sur le processus restreint à la réparation des pièces de rechange après la récupération de ces dernières, afin d'améliorer la performance des centres de réparation. Notons que ce processus a un impact direct sur la satisfaction client, priorité de l'entreprise, et offre par ailleurs un plus grand intérêt d'un point de vue scientifique au vu des verrous que nous présenterons dans le chapitre 3 de cette thèse.

Avant l'analyse de la problématique scientifique, nous proposons dans la section suivante, de souligner plus en détail les spécificités de la chaîne logistique de pièces de rechange.

IV La chaîne logistique de pièces de rechange

1 Définition

La chaîne logistique de pièces de rechange représente une composante importante dans la continuité de fonctionnement de nombreux systèmes industriels, en contribuant à l'exécution des opérations de maintenance dans les délais optimaux (Figure I.5). L'absence d'une chaîne logistique appropriée peut engendrer des MDT (Mean Down Time) élevés ainsi que des coûts importants associés à la non-disponibilité des systèmes chez les clients (Huiskonen *et al.* (2001)) mais aussi des coûts d'acheminement non négligeables.

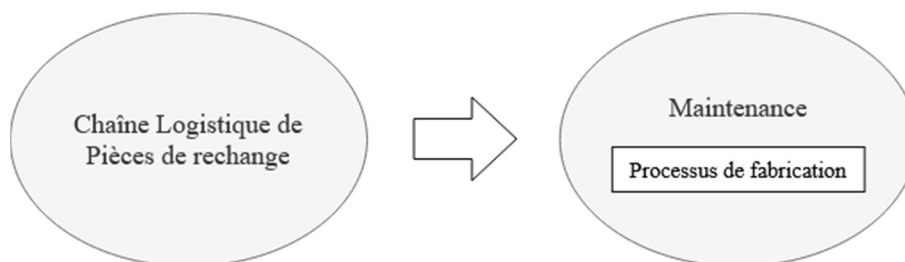


Figure I.5: Le lien entre la CL-PR et le service de maintenance (Huiskonen *et al.* (2001))

Selon Tapia-Ubeda *et al.* (2018), la chaîne logistique de pièces de rechange est une structure composée de fournisseurs, transporteurs, ainsi que des entrepôts de stockage internes et externes afin de garantir la disponibilité des pièces auprès des clients (Figure I.6). Notons ici que la demande en pièces de rechange est pilotée par les stratégies de maintenance mises en place chez le client. L'optimisation des maintenances préventives et correctives repose sur des

caractéristiques fiabilistes des systèmes, elles-mêmes pouvant dépendre des conditions d'exploitation et d'usage de ces mêmes produits. Ceci rajoute une couche de complexité et renforce l'aspect aléatoire du problème décisionnel.

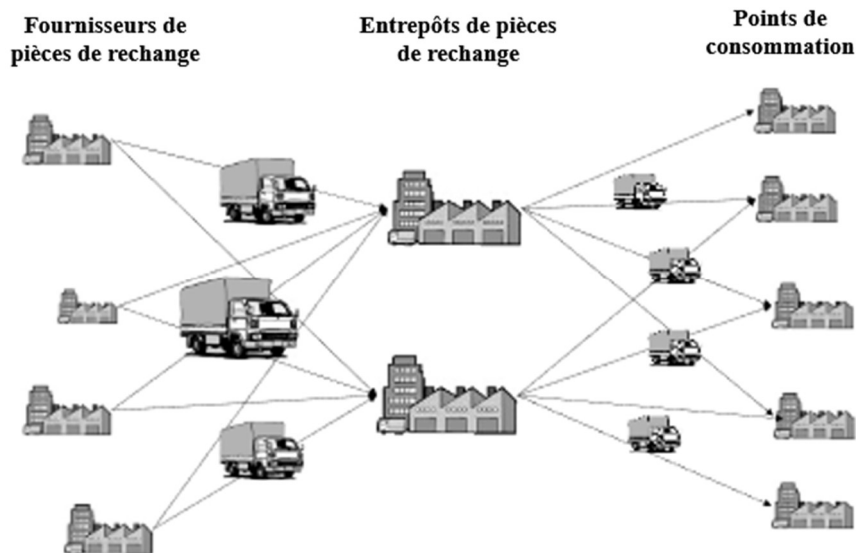


Figure I.6: La chaîne logistique de pièces de rechange générique (Huiskonen *et al.* (2001))

2 Les particularités de la chaîne logistique de pièces de rechange

La chaîne logistique de pièces de rechange a des particularités par rapport à la chaîne logistique classique, sur plusieurs niveaux :

2.1 La nature de la demande :

Dans une chaîne logistique de pièces de rechange, la demande est principalement liée à des opérations de maintenance corrective ou préventive. Alors que traditionnellement, les dates de maintenance préventive étaient déterminées préalablement en fonction de critères notamment économiques, la migration vers des approches dites conditionnelles ou prévisionnelles pour un remplacement au plus juste en fonction de l'état de santé du système ne permet plus de les connaître précisément, voire pire pour le fournisseur qui n'a généralement pas accès aux différentes données d'exploitation machine de les rendre totalement aléatoires. On pourra alors être amené à renforcer les niveaux de stocks pour améliorer la désynchronisation avec la demande. En revanche, le prix croissant de pièces de rechange pousse à limiter ces mêmes niveaux de stocks ayant par ailleurs des taux de rotation relativement faibles au regard de la taille de la base installée, les biens pouvant être considérés comme fiables. Enfin, il n'est pas possible pour le fournisseur d'augmenter « artificiellement » ce taux de rotation via du marketing ou de la publicité, ce qui est faisable pour d'autres types de produits.

Pour finir, la différence majeure est bien liée aux modèles de prévision de la demande devant chercher à capturer une sporadicité très élevée ainsi que d'éventuels pics de demande liés par exemple à des problèmes de qualité, ces pics étant très difficiles à prédire et donc à gérer par les industriels.

2.2 La structure de la chaîne logistique de pièces de rechange :

Au niveau de la structure, la chaîne logistique de pièces de rechange se compose d'une partie de fournisseurs qui peuvent être internes ou externes, offrant des pièces de rechange sur la demande du service d'approvisionnement. Ces pièces sont ensuite stockées dans la deuxième partie de la CL qui est les entrepôts centraux de stockage qui, au contraire d'une chaîne logistique classique, sont divisés en deux parties : (i) le stock en état de fonctionnement et (ii) le stock défectif.

Les entrepôts centraux de stockage alimentent les entrepôts régionaux qui eux livrent les pièces à la demande des techniciens de maintenance. Après une intervention, les techniciens retournent les pièces défectueuses récupérées des systèmes clients aux entrepôts centraux directement, et pas aux régionaux. Ces pièces défectueuses sont stockées dans la partie spécialisée pour stocker le stock BOH (Bad on Hand).

Au besoin, cet inventaire défectueux est ensuite envoyé à la réparation dans les centres de réparation internes ou externes suivant le cas et les avantages ou inconvénients pour chaque entreprise.

2.3 Le nombre de références de pièces de rechange :

L'une des particularités de la chaîne logistique de pièces de rechange est le nombre très important de références. Pour le cas du domaine de l'aviation par exemple, le nombre de pièces de rechange a géré varie entre 300.000 et 500.000 références (Janvier-James *et al.* (2012)).

V La chaîne logistique en boucle fermée de pièces de rechange réparables

La chaîne logistique en boucle fermée de pièces de rechange (CL-BF-PR) est une structure hybride combinant la chaîne logistique de pièces de rechange et la chaîne logistique en boucle fermée.

La réutilisation des pièces de rechange est de plus en plus présente dans l'industrie. Des stratégies visant à intégrer les retours de biens dans les opérations apparaissent plus fréquemment et les biens retournés sont considérés comme des ressources de haute valeur surtout à leur fin de vie afin de limiter les effets d'obsolescence.

En addition, aux opportunités financières pour les entreprises, selon la hiérarchie de gestion de déchets européenne, la réparation ayant pour objectif la réutilisation est préférable au recyclage vu l'impact positif que ce type de chaînes logistiques a sur l'environnement (Fleischmann *et al.* (2003)).

La CL-BF-PR combine la logistique inverse, la logistique inverse et la chaîne logistique de pièces de rechange afin de répondre aux exigences à la fois environnementales, économiques et pour supporter la base installée des équipements clients jusqu'à leur fin de vie.

1 Les pièces de rechange réparables et consommables

La réparabilité des pièces de rechange est une des différences les plus importantes entre une chaîne logistique de pièces de rechange normale et une chaîne logistique de pièces de rechange réparables. La figure I.7 illustre ces deux différents types et les différences entre les deux lorsqu'elles tombent en panne.

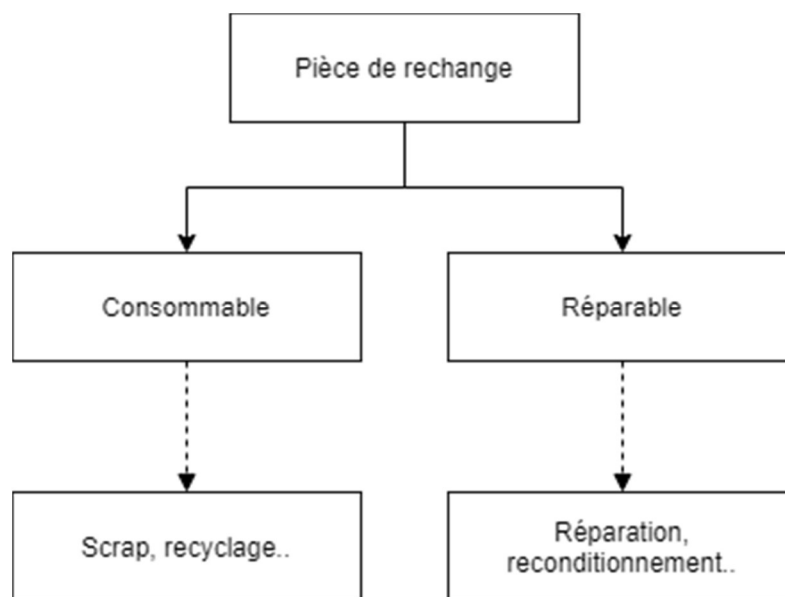


Figure I.7 : Les types de pièces de rechange

1.1 Les pièces de rechange consommables

Les pièces de rechange consommables sont des ressources destinées à un usage unique. Lors d'une intervention sur une pièce consommable, cette dernière est récupérée et retournée à l'entreprise d'origine pour des raisons de recyclage ou de gestion de matériaux tels que des batteries ou des câbles électriques...

Une pièce consommable peut se transformer en une pièce réparable via des programmes d'étude de viabilité technique et commerciale.

1.2 Les pièces de rechange réparables

À l'opposé des consommables, les pièces de rechange réparables sont des ressources qui peuvent être potentiellement remises en état, à minima de manière minimale, pour être ensuite réutilisées généralement dans d'autres systèmes que ceux dont elles provenaient initialement. Ces pièces sont généralement défectueuses, leur défaillance ayant généralement entraîné la défaillance du système complet, ou, dans le meilleur cas, dégradées lorsqu'elles sont liées à une maintenance préventive. Notons qu'une pièce récupérée sur un système défaillant ou non, est renvoyée à l'entreprise pour analyse et éventuellement réparation ou réhabilitation. Il est délicat de considérer qu'une pièce réparée présente les mêmes performances de fiabilité qu'une pièce standard. Des analyses de fiabilité des systèmes réparables peuvent être dans certains cas produites. Elles requièrent un suivi spécifique des pièces de rechange pour analyser leur durée de vie après réparation. Notons ici que, dans un contexte industriel de diversité de produits et de clients, cette traçabilité est loin d'être acquise et qu'il devient difficile, voire impossible de réaliser de telles études.

Les effets de maintenance dite imparfaite sont ainsi mesurables :

- Sur les lois de demande en termes de fréquence (augmentation de la demande liée à la réduction de la durée de vie d'une pièce de rechange) et aussi en termes d'incertitude (liée à la difficulté d'estimer leurs lois de durées de vie résiduelle) ;
- Sur les niveaux de stocks de ces mêmes pièces qui varient en fonction du taux de réparabilité lui-même variant en fonction du nombre de réparations que chaque pièce aura subi.

Afin de s'assurer des niveaux de stocks de pièces de rechange suffisants, il sera nécessaire de les réapprovisionner avec des pièces neuves. Cette analyse montre bien un niveau de complexité croissant par rapport à des chaînes logistiques inverses classiques pour lesquelles seuls des taux de retour des équipements ou composants en bon état sont à estimer.

La gestion de pièces de rechange réparables diffère des approches classiques, vu qu'il faut inclure plus de variables dans le processus de décision d'approvisionnement. L'inventaire n'est plus un inventaire simple, car pour une référence donnée, on devra différencier un niveau d'inventaire en état de fonctionnement et un deuxième niveau pour les pièces défectueuses.

Pour les décisions d'approvisionnement, il n'est pas suffisant de simplement considérer le niveau de stock de sécurité, la demande et niveau d'inventaire. Il faut également inclure d'autres éléments tels que le taux de réussite de réparation, le niveau d'inventaire défectueux, le délai de réparation en comparaison avec le délai d'approvisionnement, etc. Ceci est un des éléments les plus importants dans la problématique de cette thèse, et que nous allons traiter plus en détails dans le chapitre suivant.

VI Conclusion

Dans ce chapitre, nous avons introduit les notions théoriques de la chaîne logistique directe, la logistique inverse, la chaîne logistique en boucle fermée et la chaîne logistique de pièces de rechange en boucle fermée. Nous avons cherché à spécifier les caractéristiques de différenciation et souligner les sources de complexification dans le contexte de la chaîne logistique de pièces de rechange en boucle fermée, contexte de ce projet de thèse.

En plus du flux purement physique de la chaîne logistique, nous avons vu que les flux financiers et informationnels jouent des rôles clés dans le déroulement des opérations industrielles de l'entreprise. La grandeur du volume des opérations, la diversité des références et la sporadicité des opérations se reflètent sur la couche informationnelle de la chaîne logistique. Ce flux informationnel joue un rôle clé dans notre projet, vu que nous allons nous baser principalement sur ce dernier pour construire nos modèles et proposer des améliorations.

Le travail de cette thèse s'est effectué au sein de General Electric Healthcare (GEHC), qui est un leader mondial en production de systèmes médicaux. L'objectif du chapitre suivant est l'analyse de l'organisation de la gestion des pièces de rechange chez GEHC sur la base des éléments que nous venons de présenter dans ce chapitre.

Chapitre II : La chaîne logistique de pièces de rechange de General Electric Healthcare

I Le conglomérat General Electric

Le conglomérat General Electric (GE) tient ses origines du célèbre inventeur Thomas Edison qui créa en 1878 la société Edison Electric Light Company après avoir inventé l'ampoule électrique. La société a fusionné en 1892 avec Thomson-Houston Electric pour donner naissance à General Electric (GE), entreprise alors spécialisée dans les solutions d'installations électriques. General Electric est aujourd'hui un conglomérat multinational, regroupant une multitude d'entreprises dont les activités sont très variées. Le siège de l'entreprise se situe à Boston aux Etats-Unis et qui emploie environ 205 000 personnes dans plus de 100 pays.

Depuis sa création, General Electric est devenue l'une des entreprises les plus reconnues à l'échelle mondiale. En 2014, elle a été classée 27ème entreprise mondiale en termes de chiffre d'affaires et 7ème plus grande entreprise mondiale selon le célèbre classement du magazine Forbes. Actuellement, GE est un leader technologique dans les domaines des énergies renouvelables, des énergies thermiques, des réseaux électriques, de l'imagerie médicale, du numérique et des moteurs d'avions.



Figure II.1: Les divisions de General Electric

II La branche General Electric Healthcare

L'entreprise GE Healthcare (GEHC) constitue la branche médicale du conglomérat General Electric depuis le rachat en 1920 par General Electric de la Victor Electric Company, spécialiste des tubes à rayons X, utiles pour l'imagerie médicale. Certains sont d'ailleurs aujourd'hui produits sur le site de Buc. A la suite d'une multitude d'acquisitions de diverses entreprises de systèmes médicaux, l'entreprise n'a cessé de se développer jusqu'à constituer aujourd'hui un leader mondial dans la conception et la production d'équipements médicaux avec aujourd'hui une gamme de produits extrêmement large allant de systèmes de radiographie, scanners, IRM, ultrason, etc... à des solutions informatiques intégrées pour la revue d'images et de dossiers patients. Le siège de GEHC se situe à Chicago, aux Etats-Unis, après qu'il a déménagé du comté de Buckinghamshire, au Royaume-Uni, début 2016.

GE Healthcare emploie aujourd'hui plus de 54.000 personnes à travers le monde. La structure organisationnelle de GEHC peut être décrite comme s'articulant autour de trois principaux axes : un axe fonctionnel intégrant les différents métiers et services, un deuxième axe distinguant les différents types de produits et enfin un axe régional afin de prendre en compte les spécificités des différents pôles sur lesquels GEHC est implantée.

III Le service de maintenance de GE Healthcare

GE Healthcare propose à ses clients - des structures médicales telles que des hôpitaux - de nombreux services après-vente dont le service de maintenance, qui a pour rôle d'accompagner les clients au cours de l'utilisation de leurs produits tout en assurant leur disponibilité.

L'importance pour une entreprise comme GE Healthcare d'associer des prestations de services à son offre d'équipements médicaux n'est en effet plus à démontrer. D'abord, d'une manière générale, General Electric, en tant qu'acteur industriel de premier plan, doit faire face à une concurrence exacerbée, d'où l'importance de proposer une gamme de services après-vente toujours en phase avec les attentes des clients pour s'en démarquer.

On retrouve parmi ces prestations de service le service de maintenance, service clé pour GE Healthcare. En effet, le niveau élevé de technologie des équipements que propose GE Healthcare fait qu'il est nécessaire de mettre à disposition le savoir-faire et les compétences du concepteur auprès du client tout au long de la durée de vie du produit pour pouvoir résoudre au mieux les pannes.

L'importance d'avoir un service de maintenance pour une entreprise comme GE Healthcare tient aussi et surtout de l'exigence de ses clients pour la disponibilité de ses équipements, exigence inhérente au secteur médical. L'indisponibilité d'un équipement médical peut en effet avoir des conséquences particulièrement graves sur la capacité des structures médicales à traiter leurs patients et par conséquent présenter des risques pour leur santé.

1 La chaîne logistique de pièces de rechange réparables de GEHC

GEHC dispose d'un nombre très important de clients dans la majorité des pays du monde. La grandeur des opérations de l'entreprise et la dispersion géographique de la base installée font que la gestion des inventaires, la distribution et l'approvisionnement en termes de pièces de rechange sont des problématiques importantes pour l'entreprise. Afin de supporter les équipements pendant tout leur cycle de vie, GEHC dispose de plusieurs équipes, chacune responsable sur une catégorie d'activités de support.

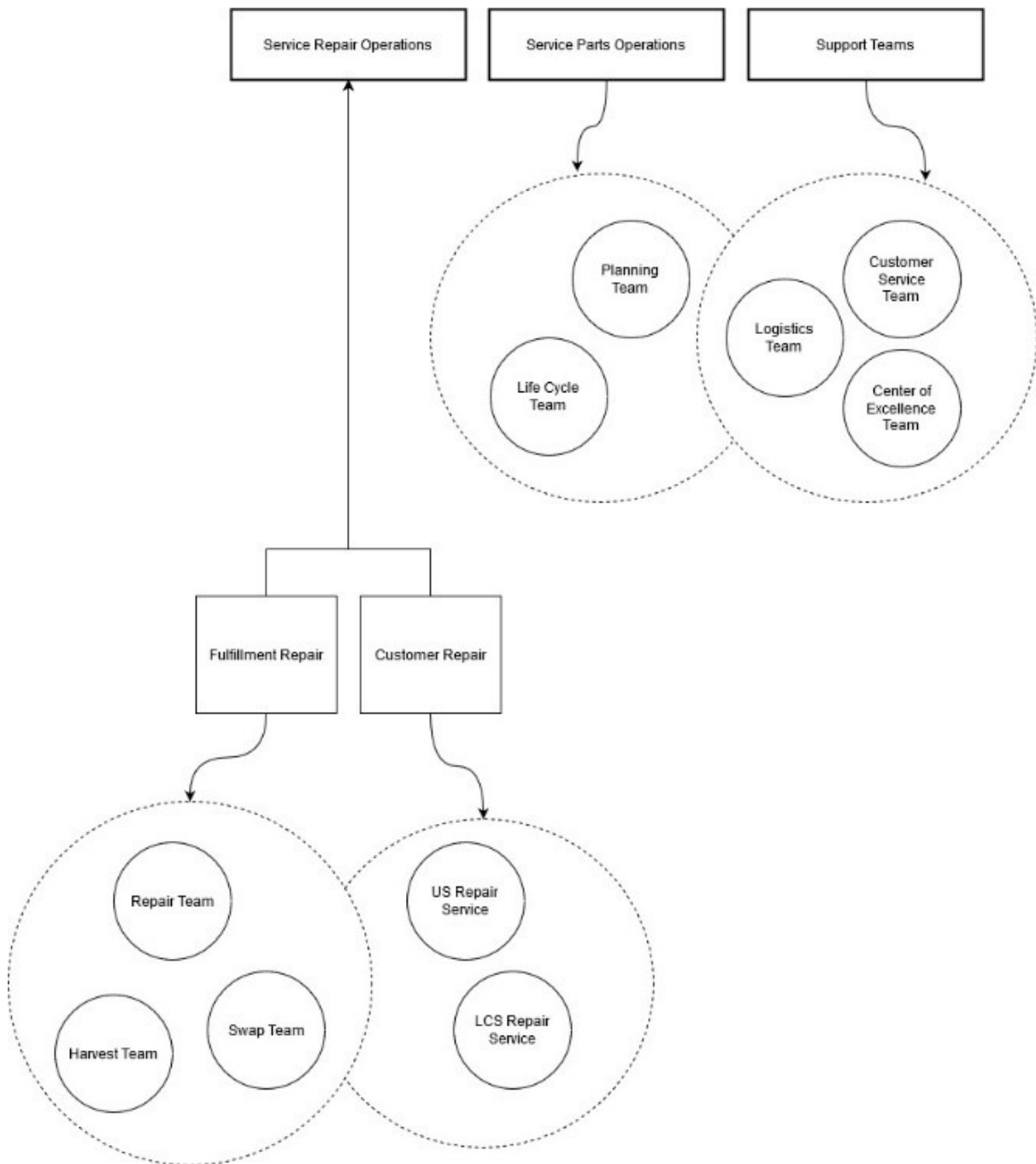


Figure II.2: Structure et équipes du service de maintenance

2 Le service de réparation des pièces de rechange de GEHC :

Le service de réparation est responsable de la remise en état de fonctionnement des pièces de rechange défectueuses. Il est reparti sous différentes équipes, chacune avec une activité différente :

○ *Réparation pour approvisionnement :*

La réparation pour approvisionnement (Fulfillment Repair) permet la réparation de pièces de rechange défectueuses pour les transformer en pièces reconditionnées, puis les remettre sur étagère en disponibilité pour les clients.

Ce type de réparation est considéré comme une source d'approvisionnement, et les réparateurs, internes ou externes, sont considérés comme des fournisseurs de pièces réparées.

▪ L'équipe de réparation :

L'équipe de réparation représente les différents centres de réparation de pièces de rechange de GEHC. Pour se rapprocher des clients, les centres de réparation sont présents dans toutes les régions du monde.

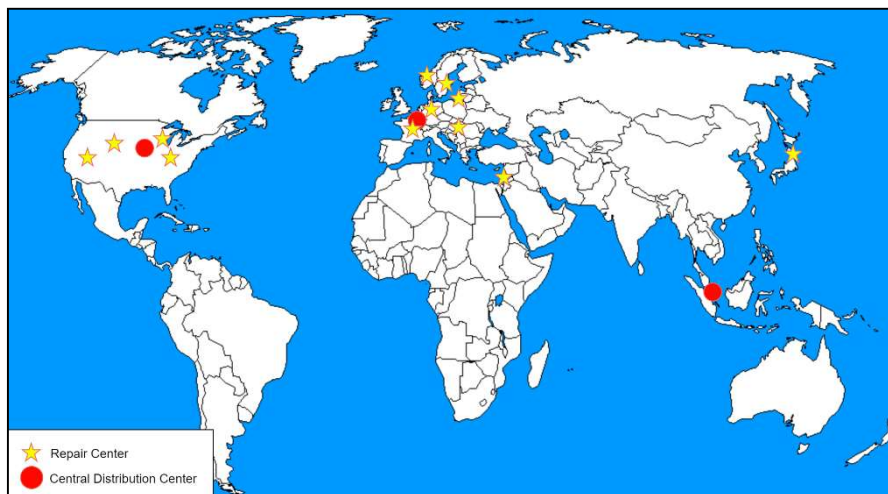


Figure 9 : Localisation des centres de réparation du monde

Les pièces de rechange sont envoyées des entrepôts de stockage vers les centres de réparation, puis, elles sont retournées pour être mises sur étagères.

▪ L'équipe « Harvest » :

Comme décrit dans les sections précédentes, GEHC récupère les systèmes client quand ces derniers arrivent en fin de vie ou quand les clients décident de les changer. L'équipe Harvest s'occupe de tester puis démonter les systèmes reçus, de récupérer les pièces de valeur et de recycler les déchets.

L'équipe Harvest s'occupe également de l'étude de viabilité technique et économique des nouveaux programmes pour des nouveaux systèmes.

- L'équipe « Swap » :

L'équipe Swap s'occupe du traitement des pièces qui sont encore sous garantie. Quand une pièce revient sur site d'un client en état défectueux, et si elle est sous garantie, l'équipe Swap s'occupe de contacter le fournisseur original et du traitement du cas afin d'avoir une réparation ou remplacement de la pièce de rechange.

- ***Réparation à la demande :***

Au contraire de la réparation pour approvisionnement, la réparation à la demande est disponible pour les clients sans contrat qui désirent réparer leurs systèmes chez GEHC. Pour ceci, ils entrent en contact avec le service de maintenance, par téléphone ou email, puis, soit un technicien se déplace chez le client, soit le client envoie le système à l'entreprise en fonction de la taille et la nature de ce dernier.

- L'équipe de réparation à la demande Ultrason « Unisyn » :

Créée en 2009 et rachetée par GEHC, Unisyn est une entreprise de réparation d'équipements ultrason tels que les sondes utilisées pour l'échographie de grossesse.

- L'équipe de réparation à la demande d'équipements de support de vie (LCS : Life Care Solutions) :

Comme Unisyn, l'équipe de réparation LCS fait des réparations sur demande pour des équipements de support de vie, comme les systèmes de support respiratoire.

3 Particularités de la chaîne logistique de pièces de rechange de GEHC : complexité, volume et variabilité

3.1 Diversité du portfolio des pièces de rechange de GEHC :

La taille et la complexité de la chaîne logistique de GEHC sont des éléments qui contribuent à la problématique qu'on traite dans ce projet de thèse. L'entreprise GEHC dispose de dizaines de centres de distribution dans les différentes régions du monde. L'entreprise gère un ensemble de plus de 200.000 références différentes de pièces chacune avec une demande et une gestion différente. En termes de volume, l'entreprise dispose d'un inventaire de plus de 10 millions de pièces de rechange dans le monde.

En termes de diversité, les pièces de rechange peuvent appartenir à plus de 10 familles différentes : Mammographie, Vasculaire, Ultrason, Support de vie, etc... Les pièces aussi varient également en termes de criticité, valeur et taille.

Ces pièces de rechange transitent dans le réseau logistique de l'entreprise d'une façon continue afin d'ajuster les inventaires régionaux et satisfaire la demande de chaque région.

La base installée de l'entreprise est également en évolution continue, de nouveaux systèmes sont introduits au portfolio de l'entreprise d'une façon continue, proposant de nouvelles fonctionnalités et de nouvelles technologies. Ceci implique de nouvelles pièces que l'entreprise doit gérer et mettre à disposition des clients.

3.2 La variation de la demande en fonction des pièces rechange et des régions

En termes de demande des pièces de rechange, il est important de noter que la demande est généralement différente d'une référence de pièces à une autre, même dans la même famille. Ceci ajoute un niveau de complexité pour les centres de réparation en termes de prévision et d'approvisionnement.

La demande varie également en fonction de la région dans le monde et l'entreprise doit ajuster les niveaux d'inventaire d'une façon continue pour répondre à la demande dans des délais optimaux. Le nombre de transactions internes entre les entrepôts de l'entreprise est de plus de 10.000 transactions par mois.

En addition, le service de gestion de pièces de rechange de GEHC doit s'adapter aux différentes réglementations de chaque pays dans le monde, notamment en termes de lois d'import-export et de douane. L'entreprise doit gérer les contraintes imposées par des pays qui n'acceptent que des pièces de rechange neuves et non des pièces réparées, ou bien qui acceptent uniquement l'importation des pièces et non l'export de ces dernières. Il existe également des pays qui acceptent la réparation, mais uniquement dans le territoire du pays. Dans ce cas, l'entreprise doit étudier la faisabilité technique et financière d'ouvrir un centre de réparation dans ce pays.

3.3 La compatibilité entre les pièces de rechange

3.3.1 La rétrocompatibilité et la compatibilité ascendante

Afin d'optimiser et de simplifier la conception des systèmes médicaux de GEHC, le service de production tend à prioriser la compatibilité en termes de pièces de rechange entre plusieurs générations de systèmes. Les systèmes médicaux peuvent avoir une compatibilité ascendante : compatibilité avec les futures générations, et/ou une rétrocompatibilité : compatibilité avec les anciens systèmes.

Ceci permet à l'entreprise de profiter de la chaîne logistique existante et de l'inventaire qu'elle a déjà pour maintenir de nouveaux systèmes, ce qui réduit les coûts financiers d'une façon considérable.

3.3.2 Les changements de conception

Un autre type de compatibilité est mis en place pour répondre à des problèmes de qualité. Quand l'entreprise fait face à un problème de qualité sur un modèle donné de systèmes, une étude est faite afin de définir la criticité du problème et identifier les pièces de rechange dans le système qui causent le problème.

Après l'étude du problème, l'entreprise peut soit :

- Retirer tous les systèmes de la base installée et livrer les nouveaux systèmes avec la nouvelle conception. Cette option est d'une criticité élevée ;
- Faire des interventions préventives sur toute la base installée et changer la pièce concernée dans les systèmes avec le nouveau modèle ;
- Attendre que les systèmes tombent en panne et changer les pièces avec le nouveau design même si ce n'est pas la cause de la panne traitée.

Dans tous les cas cités ci-dessus, un lien de compatibilité (CU : Change-up) est créé entre deux pièces ou plus. Lors d'une maintenance chez un client, le changement d'une pièce avec un CU suit une des politiques suivantes :

- Si la pièce est défectueuse :
 - La pièce peut être remplacée par le même modèle jusqu'à la consommation de l'inventaire entier, puis utiliser le nouveau modèle ;
 - L'inventaire du modèle concerné est obsolète et il faut impérativement utiliser uniquement le nouveau ;
 - Les deux modèles peuvent être utilisés, c'est au technicien de maintenance de choisir la pièce à commander en fonction du pays, des délais de livraison et de l'état de l'inventaire.
- Si la pièce n'est pas défectueuse :
 - La pièce doit être changée avec le nouveau modèle ;
 - La pièce doit être changée avec le nouveau ou l'ancien modèle ;
 - Ne rien faire.

Toutes ces options doivent être modélisées pour être intégrées dans les modèles de gestion de stock et de prévision de la demande de pièces de rechange ainsi que pour les demandes de réparation. A titre d'illustration de l'importance d'une bonne communication, supposons une

forte demande pour une référence de pièces de rechange donnée. L'entreprise pourrait revoir ses priorités de réparation pour répondre à ce besoin alors que la législation du pays de réception impose des pièces neuves. Une telle décision peut entraîner un pic d'activité dans les services de réparation et un potentiel retard de livraison au vu des stocks disponibles de pièces neuves en étudiant leur compatibilité, qu'il faudrait donc réapprovisionner d'urgence.

L'ensemble de ces contraintes augmente le nombre de variables de décision à intégrer dans le processus d'approvisionnement et le rend extrêmement complexe et quasi impossible à gérer manuellement ou avec des modèles classiques de gestion d'inventaire. GE Healthcare est depuis plusieurs années dans un processus d'amélioration continue de gestion de ses pièces de rechange au vu de son engagement de qualité de service et de maîtrise des coûts. Elle développe ses solutions propres que nous allons chercher à analyser au travers de ses processus de décision. Cette analyse nous permettra de définir nos axes de recherche.

IV Description du processus de gestion du point de vue des centres de réparation et du point de vue du client

Nous proposons ici de détailler le mode de fonctionnement de la chaîne logistique de GEHC.

Le flux de pièces de rechange forme une boucle fermée allant des fournisseurs intégrant les centres de réparation vers les centres de distributions puis vers le client. La Figure 10 schématise les différents flux de matière de cette chaîne.

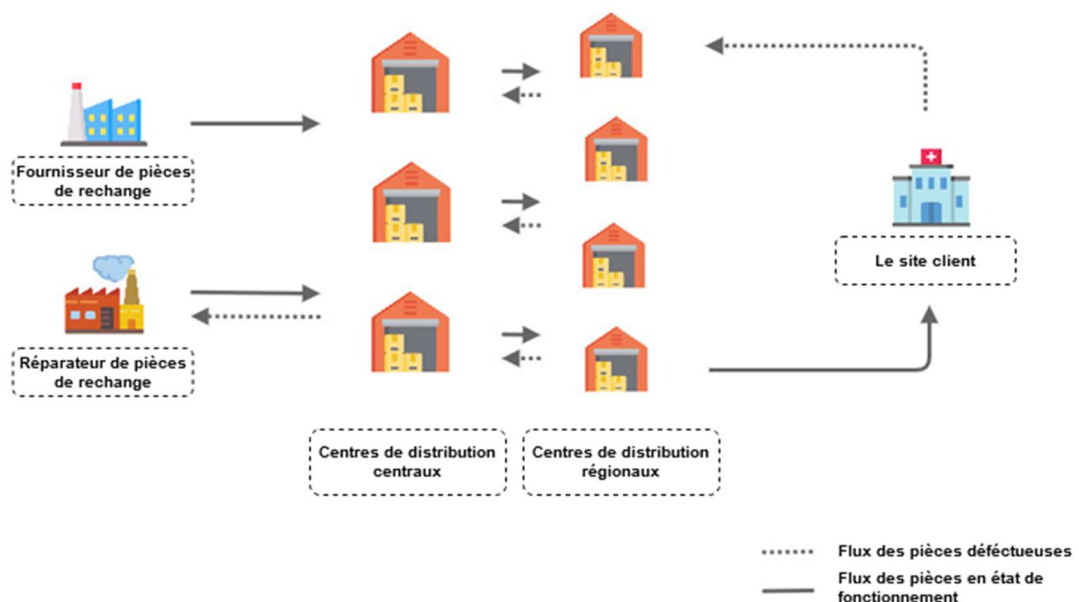


Figure II.3: Les flux des pièces de rechange chez GEHC

1 Le retour des pièces de rechange

Lors d'une maintenance sur un système client, le technicien commande la pièce de rechange nécessaire pour faire l'intervention. La pièce est envoyée du centre de distribution régional le plus proche du site client.

A la réception de la pièce, le technicien remplace la pièce défectueuse par la nouvelle pièce, puis, la pièce défectueuse est ensuite retournée vers le même centre de distribution.

Le transport des pièces de rechange est pris en charge par des transporteurs spéciaux. Le technicien généralement regroupe un lot de pièces avant de les envoyer à la fin de la journée ou la semaine. Ce choix dépend du technicien, de la région et des pièces.

A la réception dans le centre de distribution, la pièce est saisie dans le système d'information de l'entreprise et puis elle est tiquetée avec un « label » contenant le modèle de la pièce, le numéro de série et les autres informations nécessaires pour son identification. La pièce est mise sur étagère en état défectueux, en attente d'un besoin exprimé d'approvisionnement pour le lancement du processus de réparation.

2 La réparation des pièces de rechange

Comme abordé dans les sections précédentes, les pièces de rechange sont approvisionnées de plusieurs manières : soit avec un achat de pièces neuves, en réparant des pièces défectueuses, en récupérant les pièces d'anciens systèmes ou bien en profitant de la garantie de fournisseurs pour remplacer ou réparer des pièces défectueuses.

Quand une commande de réparation est créée et approuvée pour une référence donnée, une pièce défectueuse est envoyée au centre de réparation approprié. Pour le cas de GEHC, chaque pièce a un seul fournisseur de réparation, qui peut être interne (un centre de réparation GE) ou externe.

Pendant le processus de réparation, la pièce passe par plusieurs étapes :

i. Création et validation du PO (Purchase Order)

La création de la commande de réparation est effectuée par le système de planification automatique (DRP) et validée par le responsable de planification de la référence concernée, appelé « planner ». La commande en « deux étapes » est importante, vu que le planner doit prendre en considération un ensemble d'indicateurs non présents dans un modèle DRP classique. Ces indicateurs sont, par exemple, le taux d'échec de réparation, l'évolution de la base installée de l'entreprise, la possibilité d'avoir une pièce d'un ancien système au lieu de l'acheter ou la réparer, l'historique des retards ou de ponctualité des fournisseurs. En fonction de sa propre analyse, il prendra la décision ou non de créer une commande d'approvisionnement et avec quelle solution.

ii. Transport de la pièce vers le centre de réparation

GEHC dispose d'un ensemble de collaborations avec des transporteurs d'équipements médicaux. Les fournisseurs dépendent de la nature de la commande (urgente ou normale), de la région dans le monde et de la nature de la pièce.

iii. Réparation de la pièce

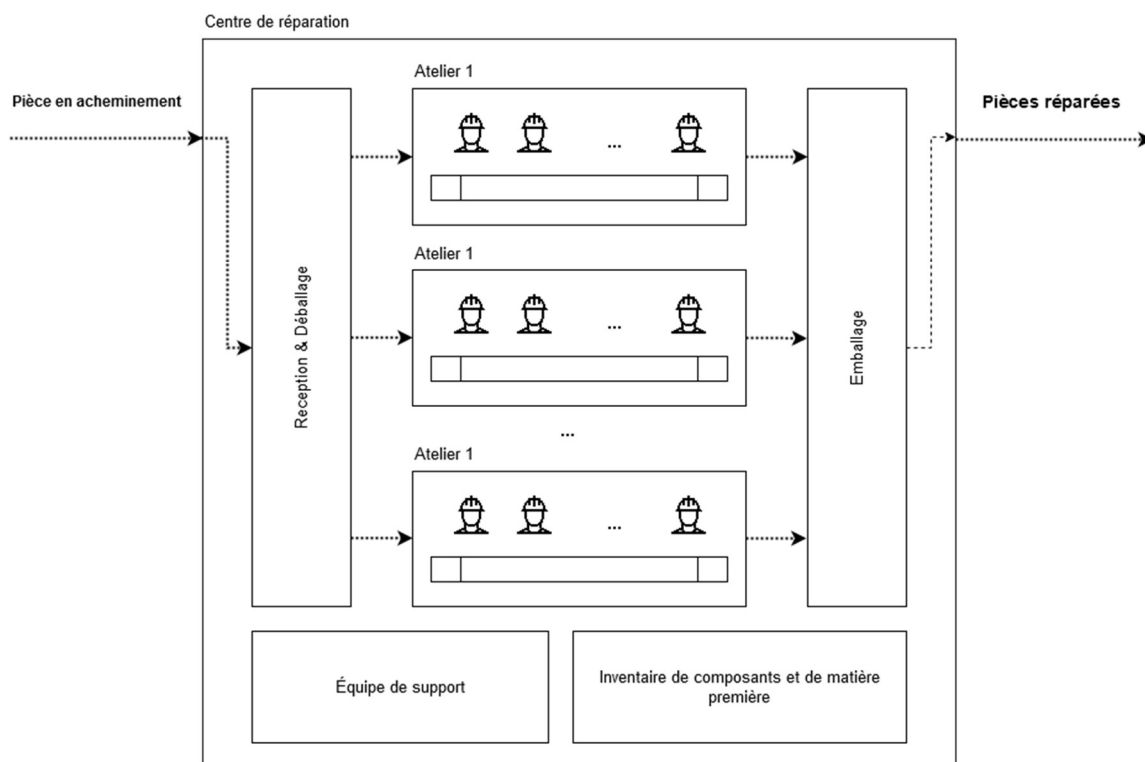


Figure II.4: La structure d'un centre de réparation GEHC

A l'arrivée dans le centre de réparation, la pièce passe par un ensemble d'étapes (Figure II.4).

- Réception de la pièce : A la réception de la pièce, une équipe s'occupe du déballage de la pièce, saisit la date de réception et les informations de la pièce dans le système d'information de l'entreprise. Elle est ensuite mise sur étagère pour les techniciens de réparation.
- Diagnostic de la pièce : Le processus consiste à faire un ensemble de tests qui sont documentés et liés à chaque modèle de pièces. Le résultat de test est la localisation de la partie défectueuse de la pièce et oriente le technicien vers les composants potentiellement défectueux.
- Réparation de la pièce : La procédure de réparation est dans la plupart des cas un remplacement direct d'un ou plusieurs composants défectueux par un nouveau.

Dans quelques cas, les techniciens peuvent être amenés à réparer un composant (absence de stock, opérations simples, ...).

- Procédure de test : Après chaque réparation, une procédure de test est obligatoire pour s'assurer du bon fonctionnement de la pièce de rechange. Il existe plusieurs types de procédures de test dépendant de la nature de la pièce. Pour une boîte d'alimentation par exemple, la procédure consiste à une mise sous tension et de stress pendant huit heures.
- L'emballage : L'emballage consiste à protéger la pièce de rechange dans un emballage antistatique et antichoc.

iv. Envoi de la pièce à l'entrepôt d'origine.

Après le déroulement de la réparation, les pièces sont mises en attente puis transportées vers les entrepôts de stockage. Il est important de noter que les pièces sont retournées au même centre de stockage d'où elles proviennent.

3 Le processus d'approvisionnement

Le processus d'approvisionnement est l'un des plus critiques au vu de sa complexité et des coûts et délais qu'il engage.

Comme stipulé auparavant, GEHC dispose d'un système de planification semi-automatique implémentant un modèle DRP. A la création d'un PO (Purchase Order) par ce système, le planner de la pièce peut approuver, modifier ou annuler la commande en fonction de son expertise et des critères mentionnés précédemment et non inclus dans le système de gestion automatique (DRP). Un exemple de critères est donné ci-après :

- La durée d'approvisionnement pour chaque solution ;
- Les niveaux d'inventaire « neuf », « réparé » et « défectueux » ;
- L'évolution de la base installée ;
- L'historique et la réputation des fournisseurs (retard, ponctualité) ;
- La prévision de la demande ;
- La disponibilité de la pièce dans des anciens systèmes qui seront récupérés par l'entreprise ;
- Le nombre de « backorders » ;
- L'historique de la demande ;
- Le taux d'échec / succès de réparation ;

- Etc...

Une commande d’approvisionnement se traduit par un ordre d’achat de neuf ou de réparation. Quand le planner reçoit une commande, il doit déterminer quelle solution choisir et ensuite quel volume commander (Voir figure II.5).

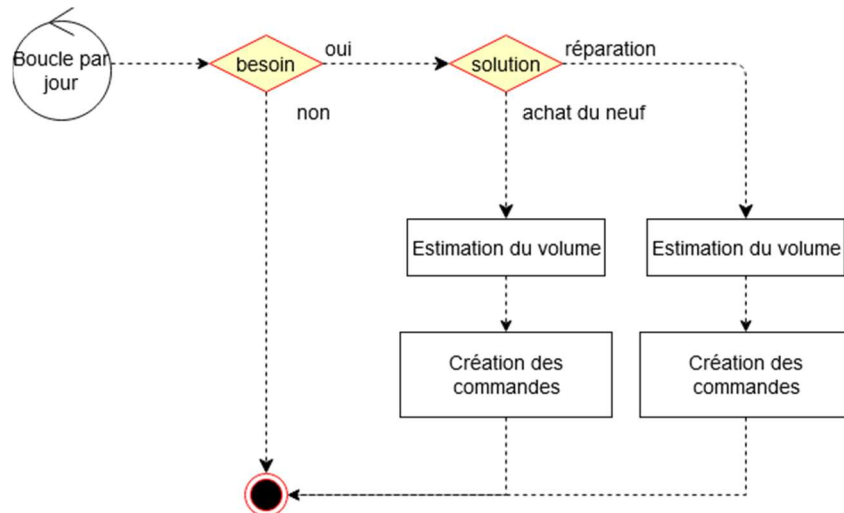


Figure II.5 : Processus haut niveau de décision de planification

Il est à noter que cette liste n’est ni exhaustive, ni formalisée. Elle est propre à chacun des planners en fonction de son propre niveau d’expertise et sensibilité. Ce constat est lié à une analyse de pratiques conduite durant cette thèse auprès de 10 planners (la population des planners est d’environ 80). Notons que, même si cette étude n’est pas statistiquement représentative et ne permet donc pas d’assurer des « bonnes pratiques » des planners, elle montre cependant des comportements très hétéroclites à la fois entre planners mais aussi entre produits. Par ailleurs, notons qu’un planner va réaliser en moyenne environ 30 opérations par jour.

Vu ces éléments auxquels s’ajoutent le nombre et la diversité des opérations, des références de pièces (+200.000 références), des solutions d’approvisionnement possibles et la complexité de la chaîne logistique de GEHC, il est délicat voire non envisageable de chercher à proposer des solutions automatisées pour la prise de décision. Il en est, au vu de l’impact sur la variabilité de la charge de travail des centres de réparation et des coûts engendrés, qu’il semble nécessaire de renforcer les outils d’aide à la décision.

4 Présentation de la problématique industrielle

Actuellement, GEHC dispose d’un réseau de dizaines de centres de réparation distribués partout dans le monde : Amérique du nord, Amérique du sud, Europe, Moyen-Orient et l’Asie ; réparant plusieurs milliers de pièces de rechange chaque jour et un chiffre d’affaire de plusieurs millions de dollars chaque mois. Ces centres de réparation reçoivent des pièces défectueuses récupérées lors de la maintenance et les remettent en état de fonctionnement à travers des lignes de réparations dédiées à chaque type d’équipement.

Dans son souci d'amélioration permanente, GEHC conduit un ensemble d'analyses de la performance de sa chaîne logistique de pièces de rechange. Actuellement, un des constats est la nécessité d'améliorer la gestion des centres de réparation qui peuvent présenter de forts pics d'activité, liés à des effets butoirs de demandes sur des pièces de rechange réparées. Ces pics d'activité ont des influences importantes sur la qualité de service en termes de délai de disponibilité des pièces et sur les coûts. Pour compenser ces retards, les mesures actuelles sont le renforcement ponctuel des équipes et le réapprovisionnement en pièces neuves. L'une comme l'autre de ces solutions ne sont pas satisfaisantes, la première engendrant des délais liés à l'identification et au recrutement ponctuel de personnels qualifiés et la seconde augmentant le nombre de pièces de rechange dans la chaîne, allant à l'encontre des objectifs de l'entreprise.

Plus spécifiquement, plusieurs conséquences critiques ont été identifiées liées à ces pics d'activité.

La première est liée aux stocks de matières premières. Pour répondre à une charge donnée, un centre de réparation doit être préparé en termes de matières premières. Les délais d'approvisionnement imposés par les fournisseurs sont généralement de plusieurs mois et ne permettent donc pas de compenser une augmentation subite d'une demande. Ceci conduit alors les gestionnaires de centres de réparation à sur-dimensionner leurs stocks augmentant de facto le risque d'obsolescence et une perte financière (immobilisation d'inventaire).

Comme souligné précédemment, la gestion des ressources humaines (taille de l'équipe et des compétences) ainsi que celle de l'espace dans les lignes de réparations dépendent aussi de la précision de cette prévision afin que le centre de réparation soit capable de gérer ses ressources et ses compétences d'une manière optimale.

Les opérations du service de maintenance de GEHC sont divisées en deux parties :

- L'équipe gestion des pièces de rechange, qui est répartie partout dans le monde et contient des dizaines de personnes.
- Les centres de réparation qui ont des spécialités différentes et qui sont distribués partout dans le monde.

Comme expliqué dans les précédentes sections, les indicateurs de performance des centres de réparation sont : (i) la longueur du cycle de réparation (Repair Cycle Time ou RCT) (ii) le taux de pièces défectueuses à la réception par les techniciens (Fail on Arrival ou FOA) (iii) le volume des réparations en cours (Backlog).

Afin d'améliorer la performance des centres de réparation, les équipes de support mettent en place plusieurs projets d'optimisation des opérations. Cependant, il est nécessaire d'avoir les ressources nécessaires en termes de matières premières, de personnes et d'espace pour absorber la charge dans les centres de réparation.

Notre analyse approfondie de la chaîne vis-à-vis de cette situation conduit à la conclusion d'un manque de fiabilité des modèles actuels utilisés chez GE Healthcare en termes de prévision de charge de ses centres de réparation. Les modèles existants sont dédiés uniquement à la prévision de la demande des techniciens de maintenance en pièces de rechange et non pas de la charge de réparation (la différence des critères à considérer entre les deux problématiques sera expliquée dans la rédaction). En plus, les modèles classiques de prévision (de type séries temporelles) ont des limites notamment quand la demande de réparation se caractérise par une variabilité élevée ainsi que des pics non réguliers. Un des objectifs est alors l'identification des causes potentielles de ces pics et leur modélisation en termes de fréquence, d'effet sur la demande et à plus long terme sur la charge des centres. Une analyse systématique des causes pour chaque référence nous semble tout à fait inenvisageable au vu de la complexité de la chaîne et du caractère multifactoriel et dépendant de ces causes. De plus, la construction d'un modèle prévisionnel basé sur ces causes ne nous semble pas possible au vu des pratiques des planners et de l'importance de leurs décisions dans cette gestion.

Cependant, les managers des centres de réparation ont un besoin primordial d'avoir une visibilité sur la future charge qui sera reçue afin de s'organiser pour avoir des opérations de qualité et de répondre dans des délais optimaux.

La difficulté d'avoir un modèle de prévision fiable est liée à la nature des opérations et à la complexité du processus d'approvisionnement. Comme décrit dans la section précédente, les planners se basent principalement sur leur expérience pour prendre des décisions, ce qui fait qu'il n'y a pas d'algorithme déterministe qu'on peut implémenter pour prévoir les futures décisions. En addition, des tentatives d'implémenter des programmes qui simulent le comportement des planners ont été réalisées sans succès dû à la complexité et au nombre très élevé de combinaisons et de situations possibles.

V Conclusion

Dans le secteur médical, la chaîne logistique de pièces de rechange joue un rôle crucial dans la maintenance des systèmes des clients afin d'éviter les interruptions et d'offrir un service stable aux clients dans les hôpitaux. GEHC a des contrats de maintenance avec de milliers de clients autour du monde, présents dans plus de 100 pays. Les prix des pièces de rechange varient énormément mais peuvent atteindre des montants de 100.000 euros pour une seule pièce.

L'analyse des points d'amélioration du processus de prévision existant de la charge dans les centres de réparation montre un défaut dans ce dernier. Pour améliorer la fiabilité de tels modèles, les approches classiques tendent à des études systématiques d'identification et de modélisation des causes de variabilité en termes d'effet et de prédiction. Ces études reposent généralement sur une connaissance précise et de la maîtrise complète de la chaîne logistique ainsi que l'acquisition de données fiables et en grande quantité.

Vu la complexité de la chaîne de GEHC, du nombre d'acteurs impliqués, un grand nombre d'incertitudes qui réside et ne permet pas de présager une modélisation fine mettant en évidence des relations explicites de cause-conséquence, renforcé par la nature multifactorielle de ces causes, nous nous proposons de retenir comme objectif de thèse d'élaborer une méthodologie pour l'amélioration des prévisions de charge des centres de réparation des pièces de rechange.

Ces prévisions permettront aux planners de mieux ajuster leur stratégie de réapprovisionnement et de réparation et pour les centres de mieux adapter leurs ressources par anticipation des pics ou de chutes brusques de charge.

Dans le chapitre suivant, nous allons d'abord étudier la littérature relative aux **problématiques de prévision qui sont déjà traitées dans la chaîne logistique de pièces de rechange**. Puis, nous allons présenter une **revue de littérature des méthodologies utilisées pour répondre à ce type de problèmes**, les avantages et inconvénients de chacune et les raisons pour lesquelles nous avons fait le choix des modèles Machine Learning.

Chapitre III : Positionnement scientifique de la problématique et contributions

Dans ce chapitre nous allons positionner le problème par rapport à une revue de littérature. Ensuite, nous allons présenter les verrous scientifiques et les contributions de notre travail qui y sont associés.

I Positionnement scientifique

Notre problématique principale est : La prévision de la charge de réparation dans une chaîne logistique de pièces de rechange en boucle fermée.

Nous allons voir la littérature traitant des problèmes similaires, et plus précisément les travaux employant des modèles d'apprentissage dans le contexte de la chaîne logistique, puis, nous allons identifier les opportunités de contributions, en analysant les modèles de la littérature, leurs avantages et leurs limites.

1 Littérature de la problématique de prévision dans la chaîne logistique de pièces de rechange

Concernant la prévision de la charge de réparation, et après une revue de la littérature, nous n'avons pas trouvé de travaux traitant la problématique exacte de « prévision de la charge dans un centre de réparation dans une chaîne logistique de pièces de rechange ». Cependant, des travaux en liaison avec notre problématique sont présents.

La publication de Hyung Tae Kim *et al.* (2019) traite la prévision des défaillances des pièces de rechange utilisées par la défense Coréenne. Le travail utilise une approche Data Mining et prévoit la quantité de pièces de rechange en utilisant des modèles tels que les « Random Forests », « Regression Trees » et des « Neural Networks ». La performance des modèles a été évaluée avec RMSE (Root Mean Squared Error) et MAE (Mean Absolute Error), le résultat montre que le modèle « Random Forests » a la meilleure précision. En comparaison avec notre problématique, le modèle ne cherche pas à prévoir la charge dans les centres de réparations mais les défaillances des pièces de rechange. La différence est que, dans notre contexte, à l'occurrence d'une défaillance, les pièces ne sont pas réparées, mais stockées jusqu'au besoin.

Ensuite, quand un besoin est exprimé pour cette référence de pièces, une évaluation par des experts est faite pour décider si la réparation est la solution optimale en prenant en considération un nombre important d'informations à ce moment donné.

Dans notre cas, la réparation n'est pas le résultat de défaillance mais d'une décision d'approvisionnement, qui fait partie d'un ensemble de solutions d'approvisionnement de pièces de rechange tels que : l'achat du neuf, la couverture de garantie des fournisseurs, la récupération des pièces d'anciens systèmes. La problématique traitée, le contexte ainsi que les données d'entrées et de sortie sont différents de notre travail.

Shinohara *et al.* (2003) propose un travail pour la prévision de la demande de réparation de pièces de rechange. Le système possède une base de données qui recueille diverses données et utilise une fonction de résumé qui extrait de la base de données les problèmes dus à la détérioration des pièces. Le système permet de prévoir, pour chaque année, un taux de panne de pièces de rechange, et en se basant sur ceci génère la prévision de la taille du marché des pièces de rechange. Comme le travail de Hyung Tae Kim *et al.* (2019), cette publication ne vise pas à prévoir la charge dans des centres de réparation.

Van Jaarsveld *et al.* (2011) présente une publication visant à optimiser la gestion de l'inventaire dans un atelier de réparation de pièces de rechange d'avions. Le modèle propose l'analyse des défaillances des pièces afin d'en tirer les composants nécessaires pour la réparation et en quelles quantités. Le travail descend au niveau des composants des pièces de rechange qui est important pour les techniciens de réparation. Le système proposé a été testé sur un vrai atelier en montrant comment on peut récupérer automatiquement les pièces et comment le système peut tourner d'une façon autonome afin de supporter les décisions de gestion de l'inventaire de composants de pièces de rechange. Le contexte du travail est le même que dans notre projet (CL-PR-BF), en revanche l'objectif est l'aide à la décision dans les centres de réparation pour l'optimisation de la gestion d'inventaire et pas la prévision de la charge. Cependant, la méthodologie d'implémentation est la même que celle qui a été suivie dans notre projet, qui est l'implémentation d'un outil informatique dans les ateliers de réparation et l'utilisation de vraies données pour calibrer le modèle.

Feng *et al.* (2014) propose un modèle à deux niveaux pour prévoir la demande de réparation des pièces de rechange utilisées dans le domaine d'aviation. Le modèle se différencie par le fait qu'il considère les raisons des pannes et pas uniquement les données historiques sur lesquels les modèles classiques de la littérature se basent. En comparaison avec notre travail, le modèle ne traite pas la charge sur les ateliers des réparations et considère qu'une pièce est envoyée en réparation directement après sa panne ce qui n'est pas le cas dans notre cas, où la réparation est demandée lors d'une décision d'approvisionnement pour remplacer un achat d'une pièce neuve.

Kareem *et al.* (2015) présente une approche pour la prévision des défaillances des pièces de rechange automobile. Le modèle utilise une méthodologie de classification afin d'intégrer

d'une façon dynamique le taux de défaillance des pièces et leur criticité dans les systèmes. Ensuite, pour chaque défaillance, le modèle se base sur des modèles classiques de gestion d'inventaire pour proposer des interventions. De la même façon que la publication de Hyung-Tae *et al.* (2019), le travail vise à prévoir la défaillance des pièces et non l'impact sur les centres de réparation.

En conclusion : Les travaux de prévision dans la chaîne logistique sont répartis en deux catégories. La première est la prévision de la demande qui couvre une grande partie de la littérature. La deuxième est la prévision des défaillances des produits (systèmes ou pièces de rechange). Les deux problématiques sont largement traitées dans la littérature et sont différentes de celle traitée dans ce projet de thèse. Dans notre contexte, ni la prévision de la demande, ni les défaillances des systèmes ne répondent à notre besoin qui est la prévision de la charge dans les centres de réparation. Une pièce qui tombe en panne n'est pas envoyée en réparation directement, et peut même ne jamais être réparée (dans le cas où le prix des pièces neuves baisse par exemple). Après une défaillance, la pièce est stockée jusqu'à ce qu'un besoin soit exprimé. Quand un besoin est signalé pour cette référence, une réévaluation est faite par des experts pour voir si la réparation est une solution optimale pour la pièce, ceci en prenant un nombre important d'informations, telles que le délai de réparation, le délai de l'achat du neuf, les prix du neuf et les coûts de réparation, la localisation du besoin dans le monde, la localisation des pièces défectueuses, le taux d'échecs de réparation etc...

Pour toutes ces raisons, les travaux existants dans la littérature représentent des premières approches mais qui ne permettent pas de prendre en compte la complexité de la chaîne logistique étudiée dans notre travail.

II Méthodologies de prévision dans la littérature et l'utilisation des modèles d'apprentissage dans le contexte de la chaîne logistique

La gestion de la chaîne logistique est un domaine traitant le contrôle et la gestion des flux des biens et de services en prenant en compte tous les processus de transformation et d'ajout de la valeur aux matières premières afin d'avoir un produit final répondant au besoin du client (Baryannis *et al.* (2019)). La nature du domaine et de ses problématiques le rend adapté à des applications de modèles d'intelligence artificielle.

Nous avons étudié 5 revues de littératures (Tableau III.1) traitant l'utilisation des modèles d'apprentissage dans le domaine de chaîne logistique, allant de 1986 à 2018. Notre revue de littérature prend en compte un total de 693 publications.

Reference	Nombre d'articles	Années	Objectifs
Baryannis <i>et al.</i> (2019)	276	1996-2018	Etude des travaux se focalisant sur la gestion des risques dans le domaine de la chaîne logistique, employant des modèles d'intelligence artificielle pour des fins de prévision, automatisation, aide à la décision afin de minimiser les incertitudes liées à la gestion de risques.
Ngai <i>et al.</i> (2014)	77	1994-2009	Revue des travaux d'aide à la décision dans le domaine du textile, proposant l'application de modèles d'intelligence artificielle
Giri <i>et al.</i> (2019)	149	1989-2018	Revue de littérature se focalisant sur les travaux employant des modèles d'intelligence artificielle sur des chaînes logistiques de produits de mode.
Min, Hokey <i>et al.</i> (2010)	28	1986-2008	Revue de littérature visant à présenter des opportunités scientifiques et des domaines adaptés à des applications d'intelligence artificielle.
Ko Mark <i>et al.</i> (2010)	163	1990-2008	Une revue de l'utilisation des méthodes informatique dans le domaine de la chaîne logistique

Tableau III.1: Revues de littérature sur l'utilisation des modèles d'apprentissage en chaîne logistique

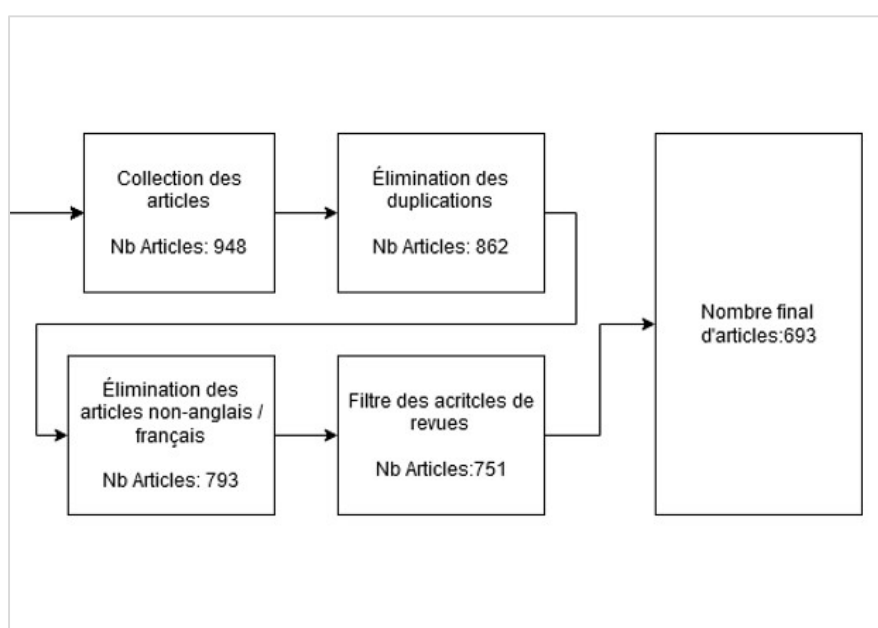


Figure III.1: Le processus de filtrage des publications des revues de littératures

Afin de suivre une méthodologie commune entre les différentes revues sur lesquelles on s'est basée, notre analyse va considérer uniquement les publications en anglais ou français, et seulement les publications en revues (Figure III.1).

1 La tendance croissante des contributions en application d'intelligence artificielle en chaîne logistique

Comme illustrée par la figure ci-dessous, dans le début des années 90s, seulement quelques travaux proposant des approches IA en CL ont été présents. Ceci peut être lié aux limites en termes de puissance des ordinateurs et de serveurs nécessaires pour faire tourner des modèles d'apprentissage à l'époque.

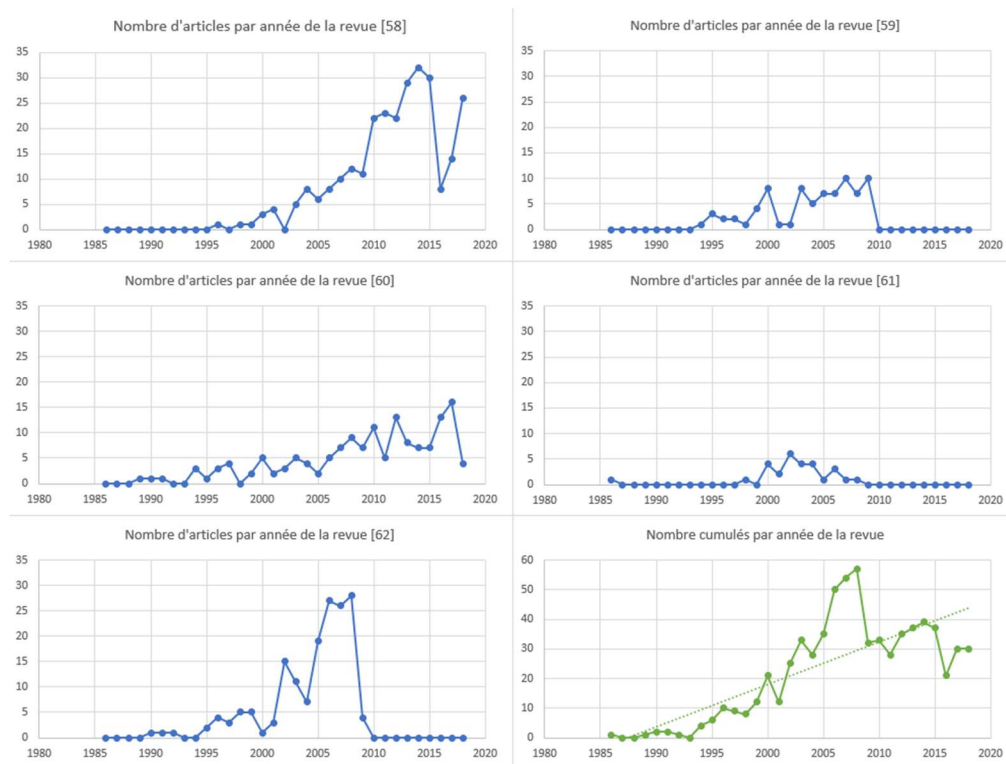


Figure III.2: Nombre de publication en IA appliquée à la CL (Baryannis *et al.* (2019))

A partir des années 2000s, nous remarquons une tendance fortement croissante du nombre de contributions. Pendant la même période, des progrès très importants ont été faits en termes de technologies telles que des cartes graphiques plus puissantes capables de tourner des modèles en temps réel, ainsi que des offres de Cloud IA qui ont commencé à se commercialiser dans la même période.

Ceci a permis aux laboratoires de recherche et aux entreprises d'appliquer les modèles théoriques de l'IA dans des contextes réels et sur de vraies données de masse, ce qui a donné naissance à des innovations telles que les voitures autonomes, assistants mobiles, etc...

2 Analyses des approches présentes dans la littérature :

En termes d'approches employées pour répondre aux problématiques de la chaîne logistique, une multitude de modèles est présente. Dans cette section, nous présentons les catégories les plus importantes.

Le tableau ci-dessous présente l'occurrence des différents types de modèles dans les revues étudiées. Nous remarquons une forte présence de modèles de programmations stochastiques avec un pourcentage de 20%. Selon Baryannis *et al.* (2019), ce genre de méthodes est très adapté aux problèmes de gestion de risque et d'incertitude.

Reference	Systèmes Expert	Programmation Linéaire	Programmation Stochastique	Modèles Mathématiques Hybrides	Algorithmes aide à la décision	Système de logique floue	Algorithmes Génétiques	Réseaux de neurones	Image Recognition	Machine Learning
Baryannis <i>et al.</i> (2019)	15	30	139	15	0	53	0	24	0	0
Ngai <i>et al.</i> (2014)	10	0	0	19	2	5	15	29	0	0
Giri <i>et al.</i> (2019)	41	20	0	0	13	0	0	0	12	63
Min, Hokey <i>et al.</i> (2010)	19	0	0	0	0	2	5	2	0	0
Ko Mark <i>et al.</i> (2010)	12	0	0	0	0	37	88	41	0	0
Total	97	50	139	34	15	97	108	96	12	63

Tableau III.2: Distribution des méthodes utilisées dans la revue de littérature

Ensuite, on trouve les algorithmes génétiques (15%), puis les systèmes de logique floue, les réseaux de neurones et les systèmes experts (14%), suivis de modèles Machine Learning (9%), programmation linéaire (7%) et enfin les modèles de programmation mathématique hybride (5%), systèmes de reconnaissance d'image et algorithmes classiques d'aide à la décision (2%).

La catégorie qui nous intéresse est les modèles d'apprentissage (Machine Learning). Ce genre de modèles se base sur des données afin d'apprendre des « patterns » en analysant ces données d'entrées. Les modèles Machine Learning peuvent avoir une base de connaissance (ou étiquettes) sur lesquelles les modèles décident la donnée de sortie (Apprentissage Supervisé) ou sans aucune donnée de support (Apprentissage non-supervisé).

Le type de modèle n'est pas suffisant pour identifier les verrous scientifiques. Un modèle peut être utilisé dans le même contexte, de plusieurs façons et paramétrages. Par exemple, un modèle d'apprentissage peut être adapté à un maximum de 2 paramètres d'entrée et pour des raisons d'aide à la décision, comme il peut être utilisé pour des fins de prévision avec un nombre élevé de paramètres.

Pour présenter les différences entre les implémentations présentées dans la littérature, la figure III.3 se focalise sur la catégorie Machine Learning et présente les pourcentages des travaux en fonction de leurs caractéristiques.

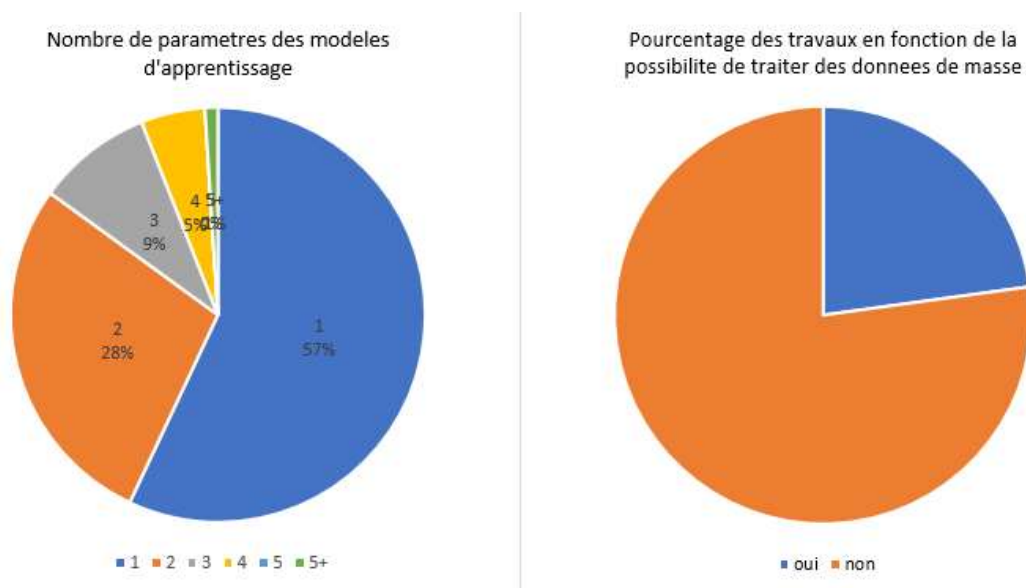


Figure III.3: Répartition des travaux en termes des variables d'entrée et de capacité à traiter des données de masse (Baryannis *et al.* (2019))

On remarque que la plupart des implémentations dans la littérature représente des modèles supportant uniquement un seul paramètre (57%). Dépendant du problème, une seule variable peut être suffisante pour que le modèle ait une « bonne » précision, mais selon Baryannis *et al.* (2019), les modèles d'apprentissage montrent une performance meilleure avec un nombre élevé de variables disponibles pendant la phase d'apprentissage.

En se basant sur les revues de littérature, une raison récurrente qui apparaît dans la plupart des articles est la difficulté de collecter des données fiables.

3 Analyses des différents cas d'utilisation des modèles :

Afin d'arriver à situer la contribution de notre travail, il est important de différencier les différentes utilisations des modèles d'apprentissage.

En analysant les articles des revues que nous avons étudiés, les modèles Machine Learning ont été utilisés pour l'aide à la décision ou pour la prévision. Parmi la première catégorie, seulement

6% des modèles est conçu pour prendre ou proposer des décisions automatiques, la majorité des travaux (51%) se focalise sur le développement des modèles qui prennent en considération des contraintes et incertitudes avec un focus moins important sur la prise de décision.

La problématique que nous traitons dans ce projet de thèse se focalise sur la prévision, pour ceci, nous employons des modèles Machine Learning. Comme le montre la figure suivante, l'utilisation des modèles d'apprentissage dans la littérature, pour des objectifs de prévision, ne dépasse pas les 4%. Selon Min, Hokey *et al.* (2010) et Ko Mark *et al.* (2010), la majorité des travaux se focalise sur l'aide à la décision et la gestion des risques.

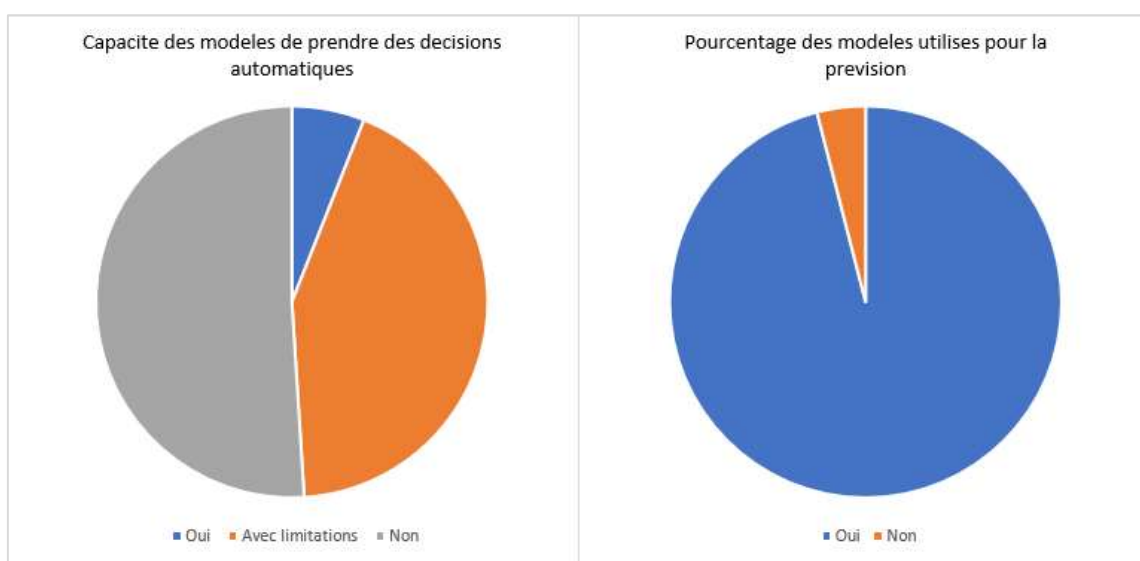


Figure III.4: Répartition des travaux par les capacités des modèles de prise de décision, de prévision et d'apprentissage (Baryannis *et al.* (2019)).

4 Contributions de la recherche

Dans la section précédente, nous avons présenté les différentes implémentations existantes dans la littérature des modèles d'apprentissage en chaîne logistique. Ceci nous a permis d'identifier des limites des modèles existants et le besoin d'implémentation d'un modèle qui prend en considération les contraintes additionnelles de notre problématique, telles qu'un nombre élevé de variables d'entrée, la capacité de traiter des données de masse en temps réel, la capacité de générer des prévisions continues, et d'autres points que nous allons voir en détail dans la deuxième partie. Ces limites représentent également des opportunités de contributions scientifiques, motivation ce travail.

Dans cette section, nous allons voir les différentes contributions pour répondre à chaque limitation que nous avons identifiée.

4.1 Modélisation et mise en place d'un modèle Machine Learning de prévision de la charge dans les centres de réparation dans une CL-PR-BF

Une contribution scientifique de ce projet de thèse est la mise en place d'un modèle de prévision prenant en considération toutes les variables de décision décrites dans les sections précédentes et qui permet de prévoir la future charge de réparation que les centres de réparation vont recevoir.

La contrainte principale de notre problématique est la complexité importante et le nombre très élevé de variables qui doivent être prises en considération pour prendre des décisions de réparation, les modèles d'apprentissage étant particulièrement adaptés à ce type de problématiques.

De plus, en cas de changement – qui sont fréquents dans notre domaine d'étude –, et contrairement aux modèles analytiques classiques où les changements doivent se faire manuellement dans l'implémentation, les modèles d'apprentissage se caractérisent par une flexibilité en termes d'adaptation à des nouvelles données d'entrée. Par exemple, si une certaine variable n'est plus prise en compte dans les décisions, les modèles ML ont la capacité d'apprendre ce changement et de s'adapter pour les futures décisions sans intervention manuelle.

Dans notre étude, nous allons proposer l'utilisation de l'apprentissage supervisé (Supervised Machine Learning) en deux types : Les modèles simples et les modèles hybrides.

4.1.1 Les modèles Machine Learning simples :

La première contribution de ce travail est l'implémentation, l'analyse et la comparaison de la performance des modèles machines Learning supervisés, avec les modèles classiques existants dans la littérature et qui sont utilisés à GEHC.

Les modèles Machine Learning sont bien adaptés pour ce genre des problématiques et sont appliqués à des problématiques similaires dans d'autres domaines. Notre but est d'appliquer ce type des modèles dans notre contexte et d'étudier leur performance.

Les modèles sont « entraînés » sur les données historiques de plusieurs années, qui peuvent être présentées par un couple $\langle x_i, t_i \rangle$ où x_i est le vecteur de données d'entrée et le t_i est la charge associée. Après la phase d'apprentissage, les modèles sont testés sur les données historiques et leur performance est calculée (Figure III.5).

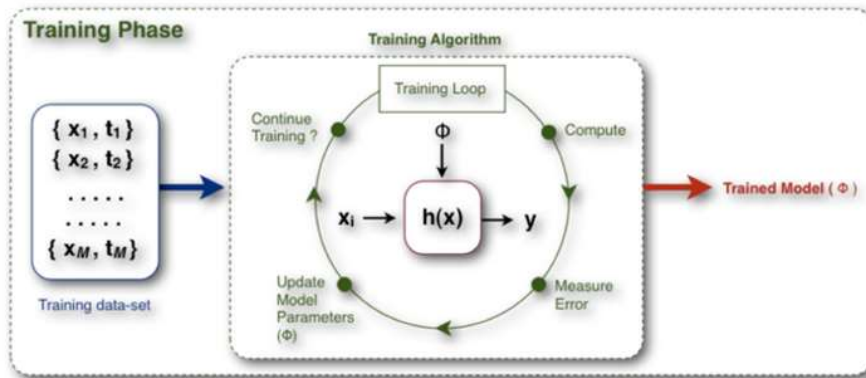


Figure III.5: Illustration de la phase de « training » d'un modèle Machine Learning (Gillian *et al.* (2010))

4.1.2 Les modèles Machine Learning Hybrides :

Une deuxième contribution de ce travail est l'utilisation de modèles d'apprentissage supervisés hybrides. L'idée derrière le concept est, au lieu d'appliquer directement un modèle, nous allons diviser notre problème de prévision en un ensemble de sous problèmes auxquels on applique des modèles Machine Learning indépendamment.

Comme nous allons le voir dans les prochaines sections, l'avantage de cette méthode est la possibilité d'utiliser pour chaque sous problème un modèle adapté. Le processus de décision de réparation qu'on essaie de prévoir par exemple, se décompose en plusieurs étapes ; dans certaines étapes, on essaie de prévoir le type de décision et dans d'autres on essaie de prévoir des quantités. Avec un modèle hybride, on a la possibilité d'utiliser des modèles de classification pour prévoir les décisions et des modèles de régression pour prévoir les quantités (Figure III.6).

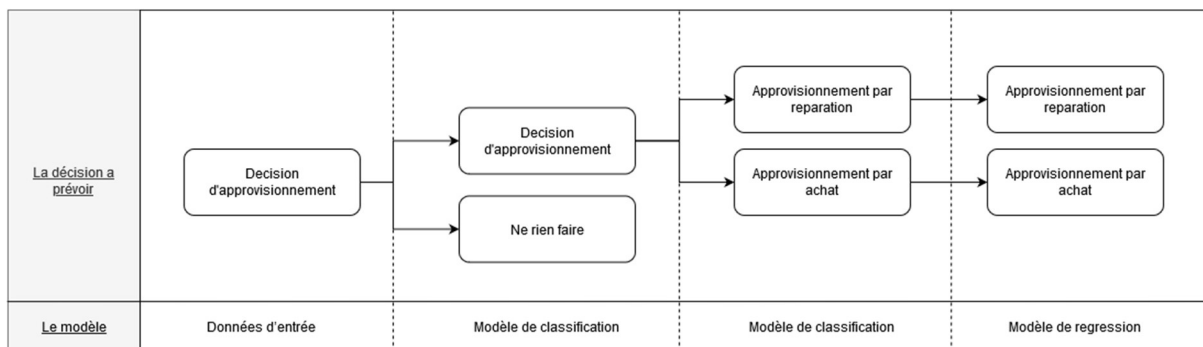


Figure III.6: La décomposition du problème et le type de modèles ML adaptés

Comme le montre la figure, nous divisons le problème de prise de décision d'approvisionnement en 3 sous-problèmes : deux classifications et une régression.

4.2 Implémentation d'un système de prévision de la charge, combinant des modèles Machine Learning Hybrides et de la simulation

Les modèles ML ne sont pas suffisants pour prévoir la charge d'un centre de réparation d'une façon continue, vu qu'ils permettent uniquement de prévoir la charge à un instant donné en fonction des données d'entrée de cet instant (état de l'inventaire, la demande, les pièces en transit...). C'est sur cette partie que la simulation joue un rôle important de transformation des données de sortie des modèles ML à de nouvelles données d'entrée via la simulation des opérations de la chaîne logistique, tels que le transport, les transactions entre les entrepôts et la consommation des pièces.

L'utilisation principale des modèles ML est la prévision de la décision des planners, afin de prévoir quelles décisions vont être prise en termes d'approvisionnements. La simulation prend les prévisions de ces décisions et enchaîne un ensemble d'opérations tels que l'envoi des pièces qu'on a prévu de réparer, la création des commandes d'achat et le transport, tout en prenant en compte les délais en se basant sur l'historique. La simulation génère un nouveau vecteur de données qui est utilisé par les modèles ML pour prévoir les prochaines décisions pour les prochains jours ou semaines en se basant sur ce qui s'est passé.

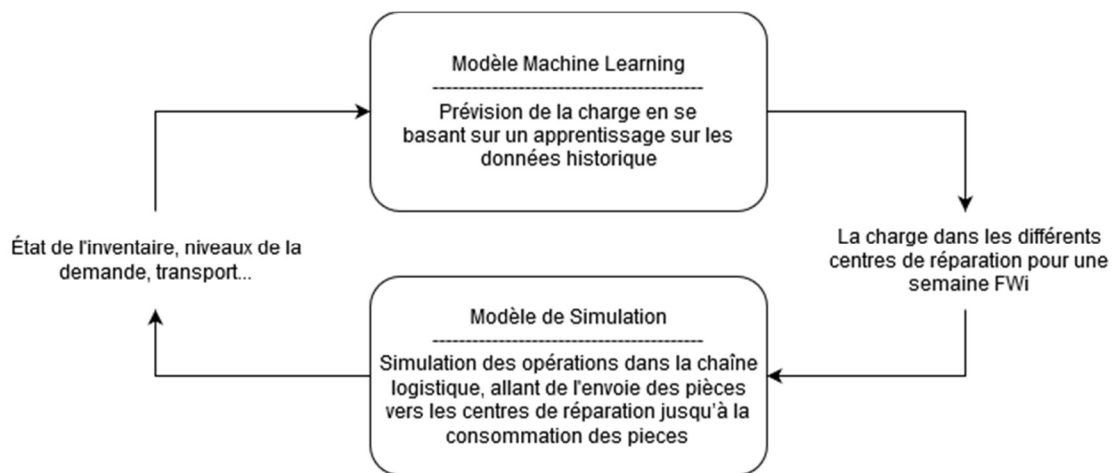


Figure III.7: Combinaison entre les modèles Machine Learning et la Simulation

4.3 Mise en place d'un système d'aide à la décision et de prévision pour le service de réparation de GEHC

Afin de mettre les résultats de ce projet de thèse en utilisation et d'en tirer le maximum de valeur, les modèles développés sont intégrés dans une plateforme d'aide à la décision qui a été mise en place également pendant cette thèse.

La plateforme met des prévisions de la charge de réparation à disposition de l'équipe de réparation, ainsi que de la visibilité en temps réel sur l'état de la chaîne logistique.

La plateforme contient également des modules responsables sur la synchronisation, « training » et d'évaluation des modèles Machine Learning en temps réel. Une fois que des nouvelles données sont disponibles, la plateforme relance le processus d'apprentissage en alimentant les nouvelles données et réévalue leur performance. Ensuite, les modèles avec les meilleures performances sont choisis pour générer les futures prévisions.

III Conclusion

Après notre revue de littérature, nous constatons un intérêt croissant à l'utilisation des méthodes Machine Learning en chaîne logistique, par les entreprises et aussi dans les publications scientifiques.

Aussi, les méthodes IA employées proposent plutôt des modèles de programmation stochastique et de la programmation linéaire. Des modèles Machine Learning existent dans la littérature, mais ils représentent des limites en termes du nombre de variables d'entrée à prendre en considération.

Les modèles Machine Learning présents dans la littérature -dans le contexte de la chaîne logistique- ne sont quasiment pas utilisés pour de la prévision mais pour la gestion de risque, l'optimisation ou pour l'aide à la décision. Pour la dernière catégorie, les modèles existants ne supportent la prise de décisions automatique que dans 6% des articles.

Considérant que la problématique de notre projet de thèse soit **la prévision**, le contexte est **la chaîne logistique de pièces de rechange en boucle fermée**, dans lequel des travaux de ce type sont inexistantes au moment du déroulement de ce projet de thèse. En plus, le nombre **d'indicateurs dans les modèles existants dans la littérature sont très limités** par rapport à notre problématique, vu qu'on est amené à prendre des dizaines de variables en considération.

De même, les applications des modèles Machine Learning dans d'autres contextes donnent des résultats prometteurs, ce qui **motive l'application des mêmes modèles dans notre contexte**.

Pour toutes les raisons citées ci-dessus, nous identifions l'application de modèles Machine Learning, pour faire de la prévision des charges dans des centres de réparation, dans une CL-PR-BF, un verrou scientifique à explorer dans notre projet.

Chapitre IV : Modélisation UML de la chaîne logistique de pièces de rechange en boucle fermée

I Introduction

Afin de répondre à une problématique donnée, la compréhension du système étudié est une étape primordiale avant de passer au processus de résolution. La modélisation est une représentation abstraite d'un système physique ou virtuel permettant d'exprimer schématiquement sa structure et son comportement dans son environnement, afin de la communiquer, la partager ou pour des raisons de documentation.

Dans ce chapitre, nous allons brièvement introduire le langage de modélisation unifié UML ainsi que les raisons de son utilisation et ses avantages. Nous allons présenter les différents diagrammes que le langage offre et les utilisations de chacun.

Ensuite nous allons présenter la modélisation de la chaîne logistique de pièces de rechange en boucle fermée, sa structure (aspect statique) et son mode de fonctionnement (aspect dynamique).

Enfin, nous allons implémenter un modèle de simulation à événements discrets de la chaîne logistique étudiée, ceci en se basant sur la modélisation faite. Ce système sera une brique de notre modèle de prévision, et va permettre de générer des données d'entrée des modèles d'apprentissage que nous allons voir dans le prochain chapitre.

II Le langage de modélisation UML

1 Définition

Le langage de Modélisation Unifié (ou UML) est un standard de spécification et de définition graphique proposé par l'organisme OMG (Object Management Group) pour construire et documenter les éléments d'un système d'objets distribué (Kautz *et al.* (2018)).

Le langage de modélisation UML a pour objectif de fournir aux architectes système et aux ingénieurs logiciels des outils d'analyse, de conception et de mise en œuvre de systèmes logiciels ainsi que de modélisation des processus métier et similaires.

L'utilisation d'un standard de conception comme UML permet l'interopérabilité entre les outils de modélisation. Pour faciliter l'échange d'informations et de projets entre les outils, un standard sémantique de syntaxe est nécessaire. UML suit les exigences suivantes :

- **Une définition formelle** d'un méta-modèle commun qui spécifie la syntaxe abstraite de l'UML. La syntaxe abstraite définit l'ensemble des concepts de modélisation UML, leurs attributs et leurs relations, ainsi que les règles de combinaison de ces concepts pour construire des modèles UML partiels ou complets.
- **Une explication détaillée** de la sémantique de chaque concept de modélisation UML. La sémantique définit, de manière indépendante de la technologie, comment les concepts UML doivent être interprétés par la partie logicielle.
- **Une spécification graphique** intuitive pour représenter les concepts de modélisation UML individuels ainsi que des règles pour les combiner.

2 Les différents diagrammes UML

UML propose différents diagrammes pour les spécifications visuelles des systèmes logiciels, qui peuvent être classés en diagrammes statiques et dynamiques. Les diagrammes statiques concernent les structures du système tandis que les diagrammes dynamiques concernent les changements de comportement du système.

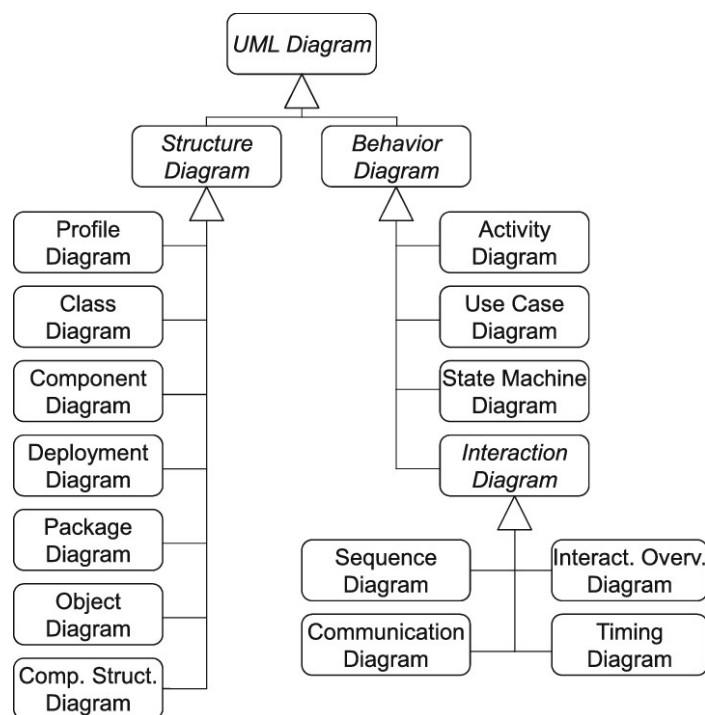


Figure IV.1: Domaines sémantiques d'UML (Kautz *et al.* (2018))

2.1 Les diagrammes de structure

Les diagrammes de structure (aussi appelés les modèles statiques) définissent la structure du système étudié, en mettant le point sur les entités principales du système, ses classes et ses objets. Les diagrammes statiques UML sont :

- **Le Diagramme de classes :**

Les diagrammes de classes sont le bloc de construction principal de toute solution orientée objet. Il montre les classes d'un système, les attributs, les opérations de chaque classe et la relation entre chaque classe.

Dans la plupart des outils de modélisation, une classe comprend trois parties. Nom en haut, attributs au milieu et opérations ou méthodes en bas. Dans un grand système avec de nombreuses classes connexes, les classes sont regroupées pour créer plusieurs diagrammes de classes. Différentes relations entre les classes sont illustrées par différents types de flèches.

- **Diagramme de packages :**

Comme son nom l'indique, un diagramme de package montre les dépendances entre les différents packages d'un système. Les diagrammes de packages peuvent utiliser des packages qui représentent les différentes couches d'un système logiciel pour illustrer son architecture.

- **Diagramme d'objets :**

Les diagrammes d'objets, parfois appelés diagrammes d'instance, sont très similaires aux diagrammes de classes. Comme les diagrammes de classes, ils montrent également la relation entre les objets mais ils utilisent des exemples du monde réel. Ceci aide à illustrer à quoi ressemblera un système à un moment donné.

- **Diagramme de composants :**

Un diagramme de composants affiche la relation structurelle des composants d'un système logiciel. Ils sont principalement utilisés lorsqu'on travaille sur des systèmes complexes avec de nombreux composants. Les composants communiquent entre eux à l'aide d'interfaces. Les interfaces sont liées à l'aide de connecteurs.

- **Diagramme de structure :**

Les diagrammes de structure composite sont utilisés pour montrer la structure interne d'une classe.

- **Diagramme de déploiement :**

Un diagramme de déploiement montre le matériel du système et les logiciels de ce matériel. Les diagrammes de déploiement sont utiles lorsque la solution logicielle est déployée sur plusieurs machines, chacune ayant une configuration unique.

2.2 Les diagrammes dynamiques

Les diagrammes de comportement définissent des déclarations sur la façon dont les éléments du domaine modélisé changent et interagissent au fil du temps. Les modèles de comportement sont aussi appelés des diagrammes dynamiques. Les diagrammes dynamiques UML sont :

- **Diagramme de cas d'utilisation :**

En tant que type de diagramme le plus connu des types UML comportementaux, les diagrammes de cas d'utilisation donnent un aperçu graphique des acteurs impliqués dans un système, des différentes fonctions nécessaires à ces acteurs et de la façon dont ces différentes fonctions interagissent.

- **Diagramme d'activités :**

Les diagrammes d'activités représentent les workflows de manière graphique. Ils peuvent être utilisés pour décrire le flux de travail métier ou le flux de travail opérationnel de n'importe quel composant d'un système. Parfois, les diagrammes d'activité sont utilisés comme alternative aux diagrammes de machine à états.

- **Le diagramme de machine à états :**

Les diagrammes de machine à états sont similaires aux diagrammes d'activité, bien que les notations et l'utilisation changent un peu. Ceux-ci sont très utiles pour décrire le comportement d'objets qui agissent différemment selon l'état dans lequel ils se trouvent à un moment donné.

- **Diagramme de communication :**

Les diagrammes de communication en UML montrent comment les objets interagissent entre eux et l'ordre dans lequel ces interactions se produisent. Il est important de noter qu'ils montrent les interactions pour un scénario particulier. Les processus sont représentés verticalement et les interactions sont représentées par des flèches.

- **Diagramme de séquence :**

Les diagrammes de séquence en UML montrent comment les objets interagissent entre eux et l'ordre dans lequel ces interactions se produisent. Il est important de noter qu'ils montrent les interactions pour un scénario particulier. Les processus sont représentés verticalement et les interactions sont représentées par des flèches.

- **Le diagramme de temps :**

Les diagrammes de temps sont très similaires aux diagrammes de séquence. Ils représentent le comportement des objets dans un laps de temps donné. S'il ne s'agit que d'un seul objet, le diagramme est simple. Mais, si plusieurs objets sont impliqués, un diagramme temporel est utilisé pour montrer les interactions entre les objets au cours de cette période.

- **Diagrammes d'interactions :**

Les diagrammes de vue d'ensemble des interactions sont très similaires aux diagrammes d'activité. Alors que les diagrammes d'activité montrent une séquence de processus, les diagrammes de vue d'ensemble des interactions montrent une séquence de diagrammes d'interaction.

Ils sont une collection de diagrammes d'interaction et l'ordre dans lequel ils se produisent. Comme mentionné précédemment, il existe sept types de diagrammes d'interaction, de sorte que chacun d'entre eux peut être un nœud dans un diagramme de présentation d'interaction.

3 Le langage UML en pratique

Le langage UML est utilisé dans une grande variété de domaines, allant de la modélisation de systèmes de grande échelle tels que des chaînes logistiques multinationales, des réseaux de télécommunication ou de grands systèmes tels que des navires ou avions, à la modélisation de systèmes de petites tailles tels que des capteurs ou des systèmes embarqués.

UML est pris en charge par un grand nombre d'outils de modélisation payants et open-source, comme Enterprise Architect, StarUML, UML Designer, Visual Paradigm, Magic Draw, Modelio et IBM Rational Rhapsody. À l'aide de ces outils, les praticiens peuvent modéliser leurs systèmes en UML et les mettre en pratique avec la simulation et analyser les résultats avec les fonctionnalités mises à disposition dans les différents outils.

III Approche méthodologique proposée : La méthodologie ASCI

La méthodologie ASCI (Analyse, Spécification, Conception et Implémentation) qui a commencé en génie logiciel, a été proposée la première fois par Gourgand et Kellert en 1991 et a été appliquée aux systèmes hospitaliers depuis 1994. Aujourd'hui, la méthodologie est présente dans de nombreux domaines.

L'approche ASCI s'appuie sur quatre étapes principales : Analyse, Spécification, Conception et Implémentation permettant d'avoir un processus intuitif allant de l'analyse haut niveau d'un système jusqu'à l'implémentation technique de ce dernier.

Un des points forts de la méthodologie est le fait qu'elle est orientée objet par défaut, ce qui facilite l'utilisation de langages de modélisation tels que UML, que nous avons aussi employé dans les sections suivantes. Les avantages de la Modélisation Orientée Object (MOO) ne s'arrêtent pas à la phase de modélisation, mais continuent jusqu'à la phase d'implémentation, vu qu'un ensemble d'outils informatique permettent la conversion d'un projet de modélisation orienté objet, à un code exécutable complet, et supportant plusieurs technologies telles que : Java Enterprise Edition, C#, C++, Python et autres.

1 L'approche ASCI dans la littérature

Ouzayd *et al.* (2011) présente différentes utilisations de la méthodologie ASCI dans la littérature et dans l'industrie. Di Martinelly *et al.* (2009) a utilisé la méthodologie ASCI pour la conception d'un modèle de pilotage permettant l'optimisation de la chaîne pharmaceutique et des flux patients, ensuite le modèle a été utilisé pour implémenter un outil d'optimisation automatique.

Moussa *et al.* (2009) a également employé la méthodologie ASCI pour l'étude des flux des clients dans un hôpital. Le modèle de connaissance a été implémenté par la suite à un modèle de simulation afin d'évaluer le taux d'utilisation des salles, des radiologues ainsi que les temps d'attentes des patients. Lemoine *et al.* (2008) a aussi employé la méthodologie ASCI pour mettre en place des modèles génériques et des méthodes de résolution pour la planification tactique mono-site et multi-sites.

Huet *et al.* (2010) s'est également basé sur la méthodologie ASCI pour structurer le processus de modélisation. Leurs objectifs étaient de mettre en place un modèle de connaissance et de l'utiliser afin de concevoir un outil d'aide à la décision permettant la planification des ressources humaines dans un hôpital.

Belkadi *et al.* (2010) a adopté la démarche ASCI pour mettre en place une modélisation et conduire des simulations de l'unité d'Hémobiologie du Laboratoire Central en Algérie. La structuration du processus de modélisation a permis aux auteurs d'évaluer le taux d'occupation des salles et de médecins en fonction de plusieurs paramètres d'entrée.

2 La démarche méthodologique

La méthodologie ASCI est composée de deux grands segments principaux, le premier est composé de deux étapes : l'Analyse et la Spécification, permettant la construction de modèle de connaissance qui représente l'aspect statique du système (la structure) ainsi que son aspect dynamique (le mode de fonctionnement).

La deuxième partie de la démarche se compose également de deux parties qui sont la conception et l'implémentation, permettant la mise en place du modèle d'action. Ce dernier représente la « traduction » du modèle de connaissance en un algorithme, un langage de programmation, un formalisme mathématique ou un outil de simulation (Lemoine *et al.* (2008)). Le modèle d'action permet par la suite d'avoir le modèle de résultats, via sa mise en œuvre et son exploitation (Figure IV.2).

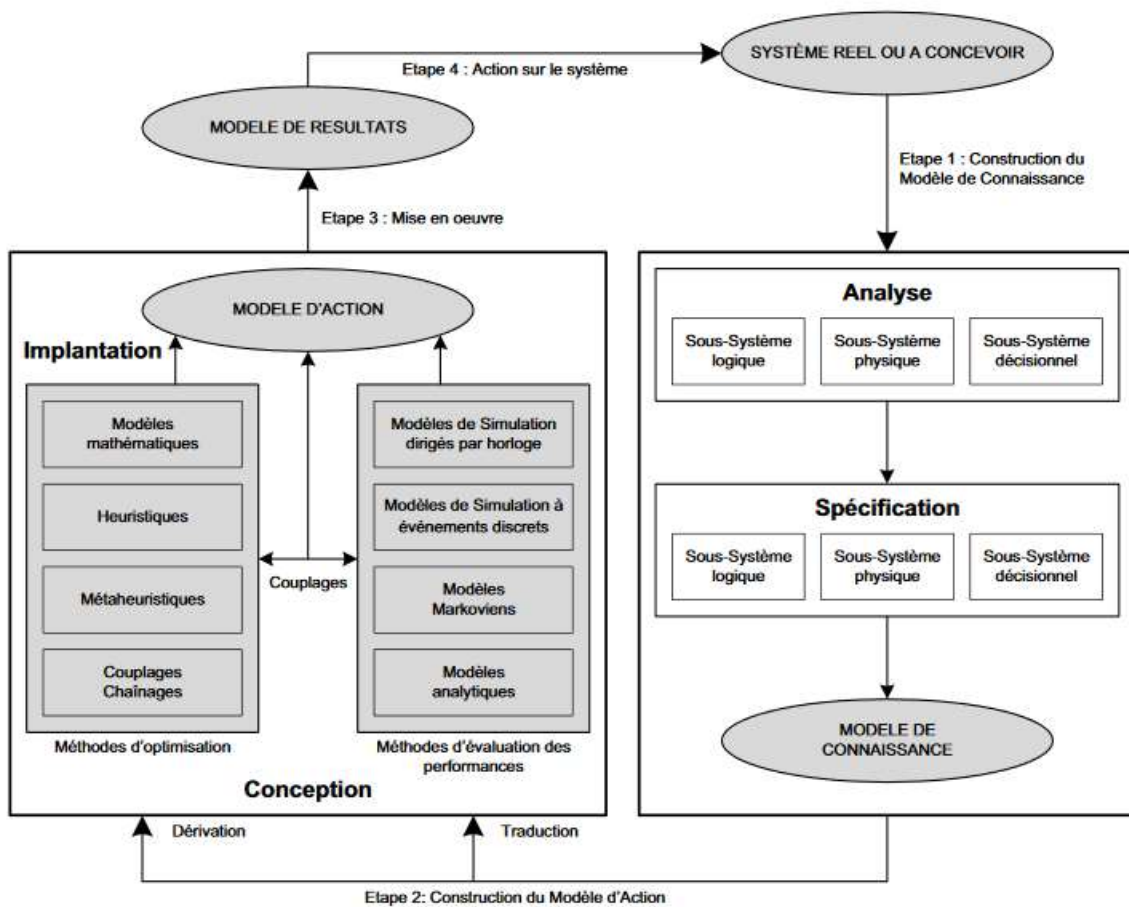


Figure IV.2: Processus de modélisation ASCI (Lemoine *et al.* (2008))

2.1 Analyse

L'Analyse représente la première étape de la méthodologie ASCI. Elle consiste à faire un « état de lieu » complet du système étudié. Selon Ouzayd *et al.* (2011), cette étape préconise de mettre une analyse fonctionnelle et structurelle du système.

Dans l'analyse fonctionnelle, le concepteur et les différents acteurs définissent les objectifs et les modes de fonctionnement du système avec un niveau de détails progressif, allant d'une vue globale à une qui est plus précise (Ouzayd *et al.* (2011)).

Pour l'analyse structurelle, la méthodologie ASCI a pour objectif de définir sans ambiguïté la composition du système, ses différentes entités principales, les attributs de chacune, leurs comportements (méthodes) ainsi que les interrelations (Ouzayd *et al.* (2011)).

2.2 Spécification

La spécification est la mise en place du modèle conceptuel proprement dit, représentant le système étudié en détails et sans ambiguïté. Cependant, avant de commencer la mise en place du modèle, un choix de langage de modélisation ainsi que les outils informatiques appropriés est nécessaire. Pour ceci, et selon Ouzayd *et al.* (2011), le concepteur doit faire une étude programmatique afin d'opter pour l'outil le plus convenable.

Un ensemble de langages de modélisation sont présents dans la littérature. La référence (Torre *et al.* (2018)) présente les résultats détaillés d'un sondage dans 28 pays et avec 124 participants pratiquant la modélisation dans des domaines différents, avec des expériences différentes allant d'un niveau débutant à un niveau « expert » avec plus de 10 ans d'expérience en modélisation (Figure IV.3).



Figure IV.3: Présentation la distribution géographique et de l'expérience des participants au sondage (Torre *et al.* (2018))

Les résultats (Figure IV.4) montrent que le langage le plus utilisé est UML avec 105 utilisateurs parmi 124 (86%), ce qui le met en première place, suivi par « autres DSL » qui représente des langages spécifiques à leurs domaines.

L'avantage de UML par rapport à ce type de langages réside dans le fait qu'il s'agit d'un standard (d'où le terme « unifié » dans le nom du langage), qui peut être compris par différentes personnes ou équipes avec des domaines d'expertises différents. Comme le précise la même étude de Torre *et al.* (2018), ceci facilite la collaboration entre les équipes techniques de développement et les équipes d'étude du projet, marketing et même le client final.

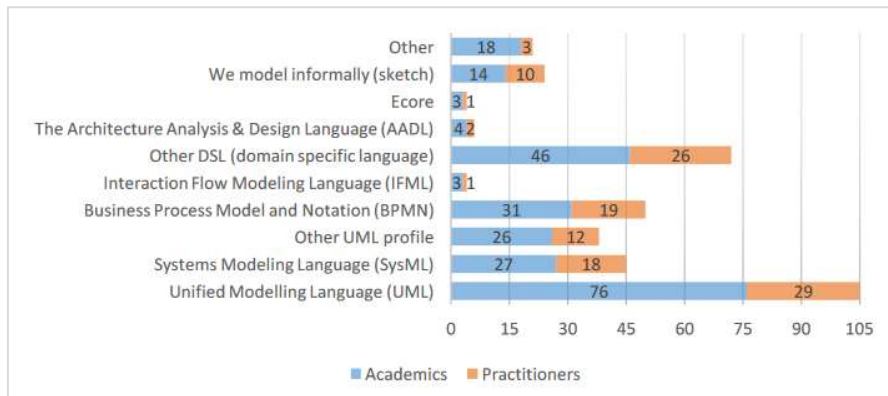


Figure IV.4: Nombre de participants utilisant chaque langage de modélisation, repartis en fonction du domaine d'activité (académique et industriel) (Torre *et al.* (2018))

Après une étude telle que celle présentée par Torre *et al.* (2018), le concepteur doit choisir le langage de modélisation et les outils à utiliser pour la phase de conception et aussi pour communiquer avec les différents collaborateurs.

Après le choix des outils, le modelleur doit mettre en place la spécification. Il s'agit de faire représenter l'analyse fonctionnelle déjà définie sous forme de modèle et de la faire compléter par les spécifications dynamiques, c'est-à-dire l'évolution du système dans le temps. Il s'agit alors de deux spécifications :

- **La spécification fonctionnelle** (statique ou structurelle) : représente d'une façon structurée les entités principales, leurs définitions (nom, attributs, titres des comportements...) ainsi que leurs interrelations et leurs types (un vers un, un vers plusieurs ou plusieurs vers plusieurs).
- **La spécification dynamique** qui met le point sur l'aspect dynamique du système, en formalisant les interactions entre les différentes entités du système, et leurs évolutions dans le temps.

Selon Lemoine *et al.* (2008), la formalisation du modèle de connaissance d'un système complexe est un processus qui doit être conduit avec de la rigueur. En plus de l'analyse structurelle et fonctionnelle, il faut également spécifier les activités et les flux qui traversent ce système, ainsi que les règles de fonctionnement qui gèrent ces flux.

Pour ceci, la décomposition du système en trois sous-systèmes est nécessaire (Figure IV.5) :

- **Le sous-système physique (SSP)** : contient les différentes entités physiques du système. Dans notre cas, ce sont les entités physiques de la chaîne logistique de pièces de rechange réparable, avec leurs différentes caractéristiques et interrelations.

- **Le sous-système logique (SSL)** : contient les différents types de pièces de rechange présentés dans notre contexte ainsi les règles permettant la transformation de ce dernier (GOH vers BOH et l'inverse).
- **Le sous-système décisionnel (SSD)** : contient l'ensemble des règles de gestion contrôlant le fonctionnement du système complet, tels que les décisions d'approvisionnement, la réparation et la gestion du transport des pièces de rechange. Selon Lemoine *et al.* (2008) le SSD est formé d'un ensemble de règles de gestion qui agissent sur les sous-systèmes logique et physique.

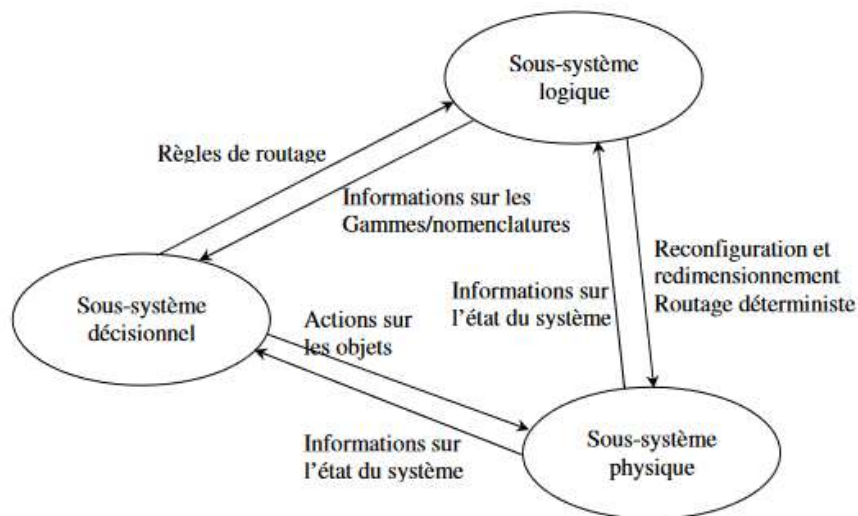


Figure IV.5: Interaction entre les trois sous-systèmes (Lemoine *et al.* (2008))

2.3 Conception

L'étape de conception consiste à exploiter la modélisation faite (et validée) dans l'étape précédente, afin de présenter les différents scénarios ou stratégies de décision.

Ceci est fait via la formalisation des modèles sous forme de modèles mathématiques ou en l'implémentant sous forme d'un modèle informatique, comme l'algorithmique ou bien en utilisant directement un langage de programmation.

2.4 Implémentation

L'implémentation est l'étape finale de réalisation de l'outil désiré en se basant sur la conception déjà faite. La méthodologie de réalisation varie d'un domaine à un autre et d'un système à un autre. Pour notre cas, nous allons implémenter les algorithmes de simulation sous forme d'un outil logiciel de simulation, permettant l'exécution des algorithmes et l'extraction des résultats.

Nous allons utiliser les langages comme Python et Pandas et nous allons intégrer l'outil de simulation final dans une plateforme d'aide à la décision pour le service de réparation des pièces de rechange de GEHC.

IV Modélisation UML de la CL-PR-BF

Afin d'implémenter des modèles de simulation, des algorithmes Machine Learning ou des outils de gestion, la modélisation du système étudié est recommandée surtout dans le cas de systèmes complexes comme le cas dans notre contexte.

Dans cette section nous proposons une modélisation de la chaîne logistique de GEHC, en prenant en compte les aspects statiques et dynamiques.

1 Le sous-système physique

L'objectif de cette section est de représenter les différentes entités physiques prenantes dans notre système, leurs différentes interrelations, ainsi que leurs attributs.

La figure IV.6 présente notre système physique complet qui est formé de 15 entités. Afin de l'analyser, nous allons détailler cette décomposition dans les sections suivantes :

- Les systèmes médicaux
- Les pièces de rechange
- Les centres de distribution
- Les moyens de transport
- Les centres de réparation

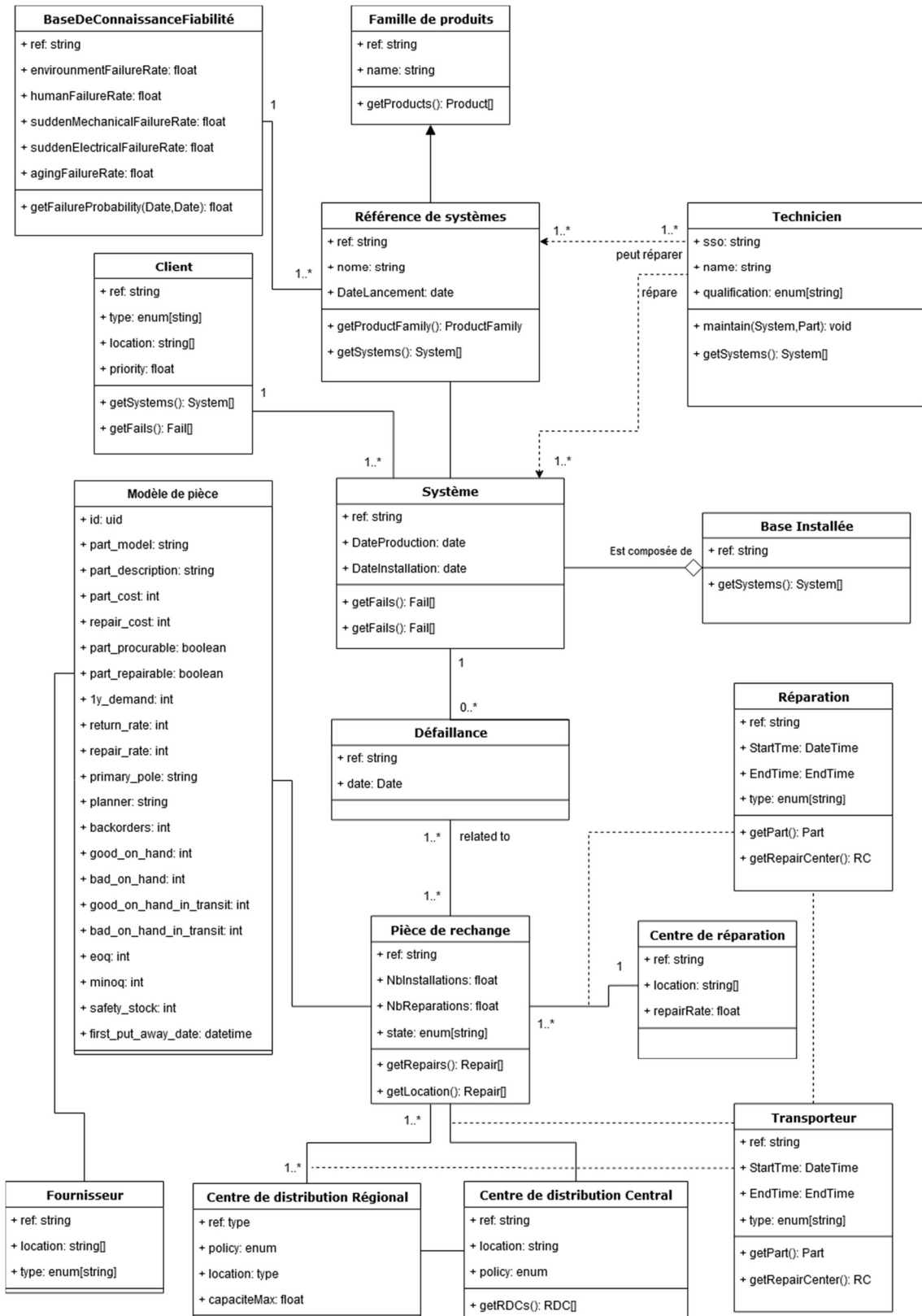


Figure IV.6: Le sous-système physique

1.1 Les systèmes médicaux

Le système médical est l'entité principale dans notre étude, c'est le produit que l'entreprise développe pendant des mois ou des années, puis qu'elle produit partout dans le monde et distribue à ses clients.

La maintenance de ces systèmes médicaux installés chez les clients, d'une façon rapide et avec des coûts optimaux est l'objectif principal de la chaîne logistique des pièces de rechange de l'entreprise.

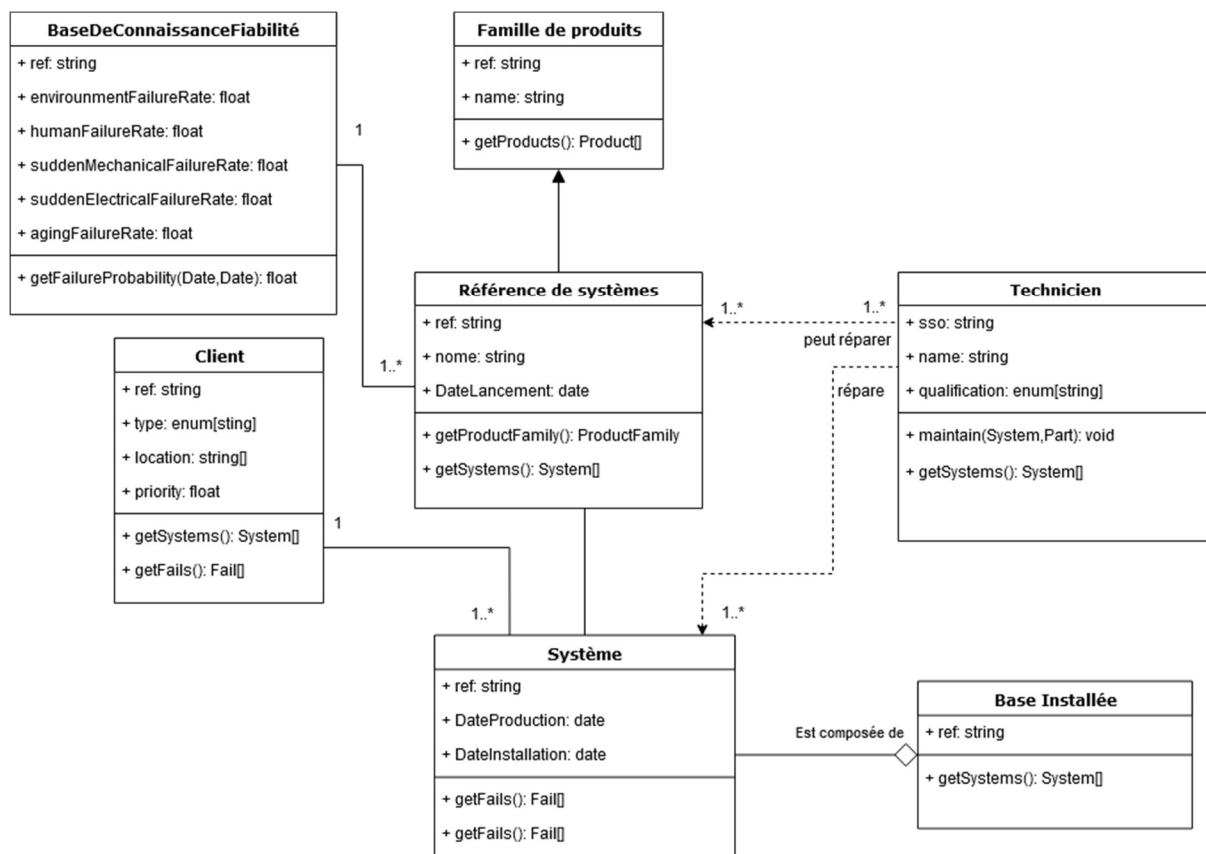


Figure IV.7: Modélisation UML du système médical et les différentes entités qui y sont associées dans le sous-système physique

La base installée de l'entreprise est composée d'un grand nombre de systèmes. Un système médical physique se caractérise par :

- Un numéro de série unique : ref (Par exemple : 435-H4-421)
- Une date de production : DateProduction (Par exemple : 01/01/2017)
- Une date d'installation : DateProduction (Par exemple : 01/01/2017)

Chaque système médical est physiquement installé chez un client, et fait partie d'une référence de systèmes, et donc « hérite » de ses attributs. Par exemple, un système avec un numéro de

série unique : 435-H4-42, installé dans un hôpital en Espagne est lié à une référence de systèmes qui est : Optima-900. Ce type de relations entre l'objet physique et son modèle (Le système physique et son modèle de référence dans notre cas) est nécessaire afin d'avoir une modélisation « normalisée », ce qui facilite le passage aux diagrammes d'objets et à l'implémentation.

De la même façon, un modèle de systèmes est lié à une famille de produits, par exemple : Systèmes mammographie, systèmes vasculaire, systèmes à rayons X ou systèmes ultrason... qui sont catégorisés en fonction de leurs fonctionnalités et technologies.

L'entreprise dispose d'un nombre de techniciens partout dans le monde, prêts à intervenir sur les systèmes client pour les inspecter et les remettre en état de fonctionnement. En fonction de sa formation, le technicien a la capacité de réparer un ensemble de références de systèmes, et lors de son activité, répare des systèmes physiques de ces références.

1.2 La pièce de rechange

La pièce de rechange (Figure IV.8) est une entité importante dans notre étude. Notre objectif final est de prévoir le volume qui sera reçu dans chaque centre de réparation dans l'entreprise. Une pièce de rechange physique est caractérisée par :

- Un numéro de série unique
- Un « état » binaire de la pièce. Défectueux, ou en état de fonctionnement.
- Une date de fabrication
- Une date d'installation dans le système client
- Le nombre de réparations de la pièce physique

Une pièce physique appartient à une référence de pièces et hérite de ces attributs. Une référence de pièces (aussi appelée modèle de pièces) est caractérisée par :

- Un prix d'achat neuf, qui est défini par le fournisseur de la pièce
- Un coût de réparation, qui est défini par les centres de réparation, et actualisé chaque année.
- La disponibilité d'achat de la pièce neuve
- La réparabilité de la pièce
- Le taux de retour de la pièce par les techniciens de maintenance

- Le taux de réparabilité de la pièce (le pourcentage des réparations effectuées avec succès).
- La demande annuelle pour cette référence
- Le niveau d'inventaire défectueux et réparable de la référence
- Le niveau d'inventaire en état de fonctionnement, prêt à être utilisé de la référence

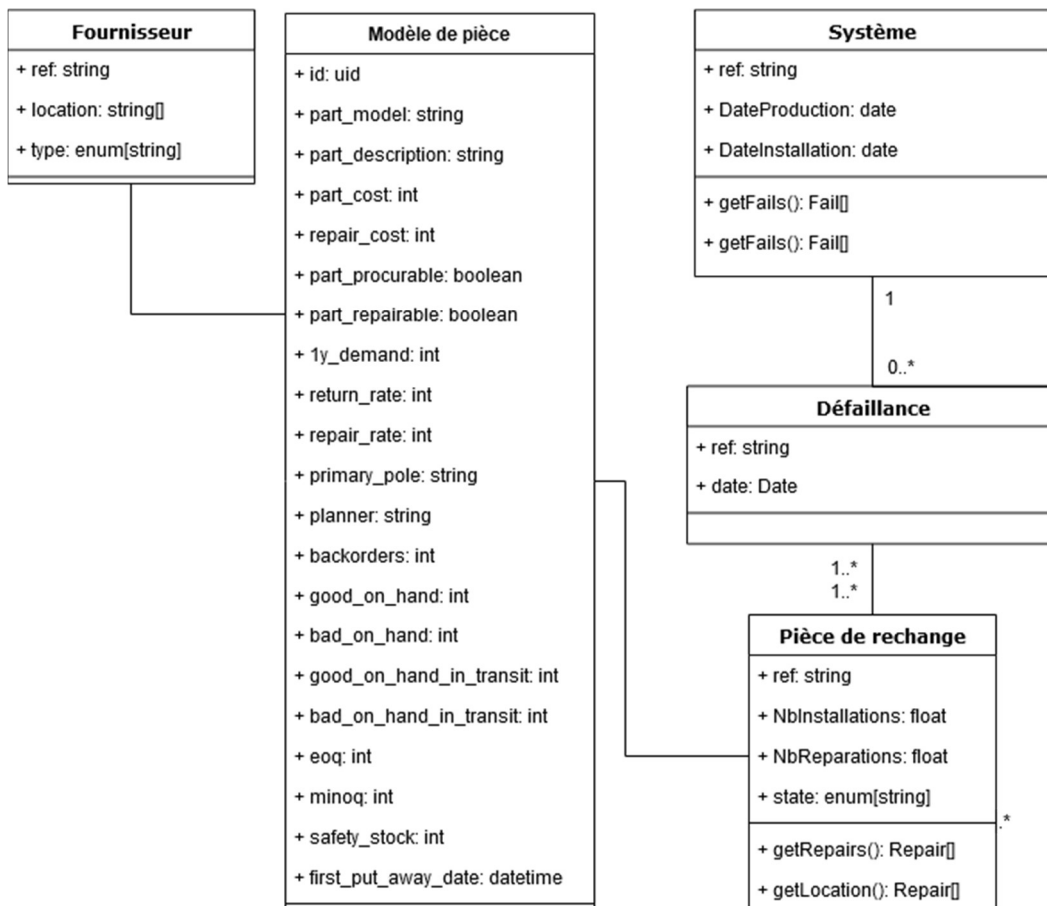


Figure IV.8: Modélisation UML (Diagramme de classe) de pièce de rechange et les entités qui y sont liées

1.3 Les centres de distribution

Les centres de distributions (Figure IV.9) représentent des entrepôts de stockage où les pièces de rechange (défectueuses ou en état de fonctionnement) sont stockées. Pour le cas de GEHC, il en existe plusieurs partout dans le monde, organisés d'une façon à réduire le temps de livraison des pièces aux clients.

Il existe deux types de centres de réparation :

- **Les centres de distribution centraux** : Il en existe trois au monde, présents dans les trois pôles : Asie, Europe et Amérique. Ces centres sont le point de contact entre les différents pôles et c'est à partir de ces derniers que les pièces sont envoyées d'un pôle

à un autre. Le calcul des prévisions de la demande finale est également fait au niveau de CDC.

- **Les centres de distribution régionaux** : Il existe plusieurs CDR par pôle, chacun est attaché à un CDC auquel la demande est remontée d'une façon hebdomadaire pour recalculer les prévisions pour chaque région. Quand un technicien de maintenance commande une pièce, c'est le CDR qui est responsable pour l'expédier, et aussi pour réceptionner les pièces défectueuses retournées du terrain.

Les centres de distribution sont caractérisés par une capacité maximale de stockage, une localisation et une référence.

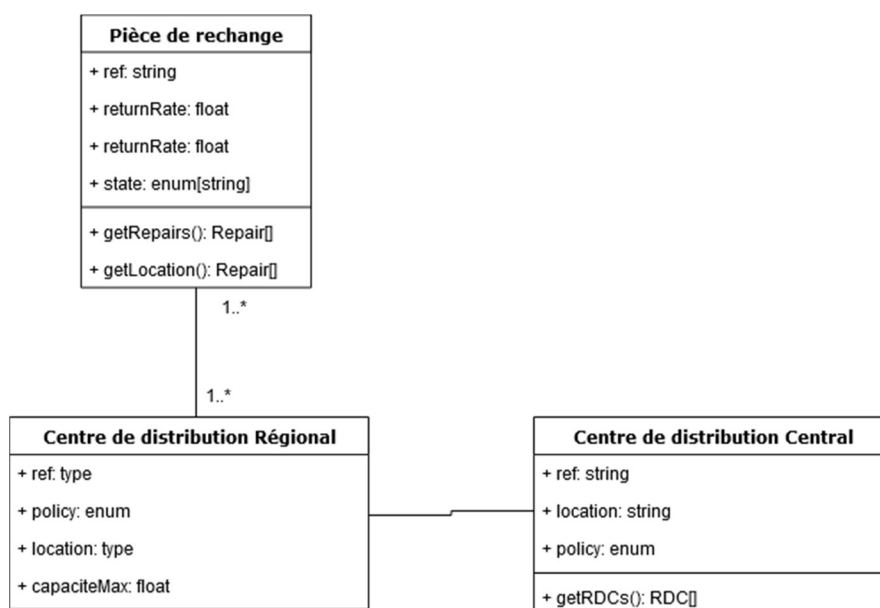


Figure IV.9: Les centres de distribution

1.4 Le transport de pièces de rechange

La classe « transport » (Figure IV.10) représente les expéditions des pièces de rechange entre les différentes infrastructures de la chaîne logistique. La classe fait le lien entre les différentes entités suivantes :

- Des fournisseurs vers les centres de distribution et l'inverse
- Du terrain (site du client) vers les centres de distribution et l'inverse
- Des centres de réparation vers les centres de distribution et l'inverse

Chaque expédition d'une pièce de rechange est caractérisée par une date de départ et une date de fin à laquelle la pièce va arriver à sa destination. Pour calculer ces dates, le modèle se base sur l'historique et calcule une estimation du nombre de jours pour le transport.

L'entité utilise trois façons différentes pour estimer le temps de transport :

- La durée fixée par les transporteurs
- La moyenne basée sur l'historique de toutes les transactions entre l'origine et la destination de livraison
- La moyenne en se basant sur l'origine, la destination et aussi la référence de pièces, ce qui fait sens pour les pièces qui ont une grande différence de taille ou de valeur.

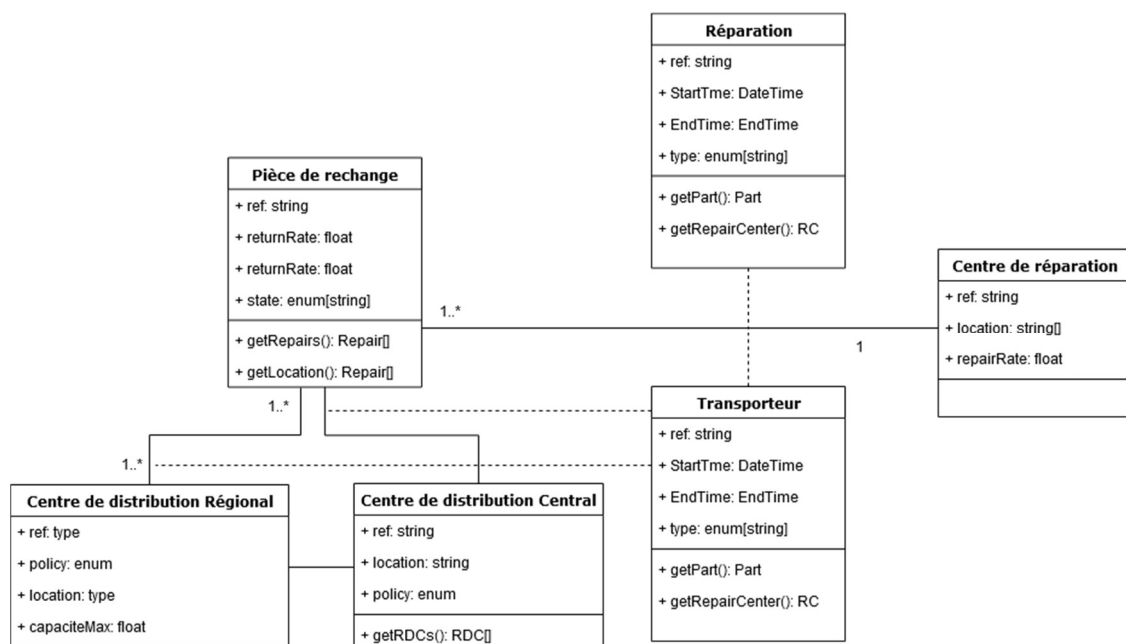


Figure IV.10: Les entités de transport des pièces de rechange

1.5 Les centres de réparation

Dans notre modèle de simulation, les centres de réparation (Figure IV.11) sont responsables de la transformation des pièces défectueuses en pièces utilisables par les techniciens de maintenance. Les centres de réparation sont caractérisés par les attributs suivants :

- Une référence unique : identifiant chaque centre de réparation et qui est utilisée dans les configurations définissant quel centre répare quelles références de pièces de rechange.
- Une capacité maximale de réparation : chaque centre de réparation est caractérisé par une capacité maximale de réparation, basée sur l'espace dans le centre et la taille de l'équipe. Par défaut, cette capacité est « infinie » dans notre modèle, mais

qu'on peut configurer et limiter. Quand cette capacité est atteinte, le centre de réparation ne peut plus réceptionner de nouvelles pièces et pour chaque lot qui arrive, les pièces sont mises en attente à la réception.

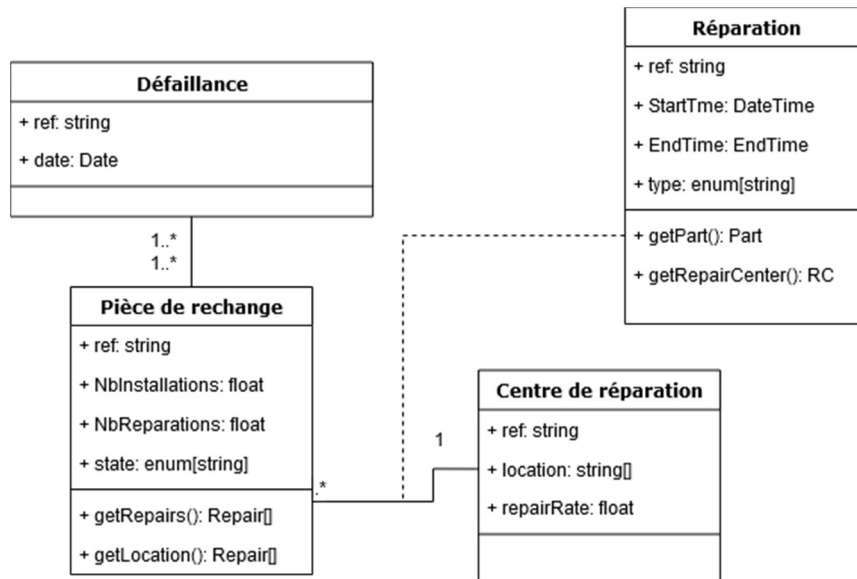


Figure IV.11: Les centres de réparation

2 Le sous-système logique

Dans cette section, nous présentons le sous-système logique contenant l'ensemble des produits présents dans notre chaîne logistique. Dans notre cas, ces produits sont les pièces de rechanges, dans deux différents états (Figure IV.12) :



Figure IV.12: Le sous-système logique

- Les pièces de rechange « en état de fonctionnement »
- Les pièces de rechange « défectueuses »

Les deux types de pièces héritent de la classe « pièces de rechange » que nous avons introduits dans le modèle physique. Une pièce défectueuse peut être transformée en une pièce utilisable via une réparation.

3 Le sous-système décisionnel

Le sous-système décisionnel (SSD) contient des éléments contrôlant le fonctionnement des entités du sous-système physique et logique. Le SSD contient la logique de fonctionnement et les règles de gestion à appliquer dans notre modèle de simulation.

Nous catégorisons les règles de gestion en 3 principales catégories (Figure IV.13) : (i) les règles de gestion DRP (ii) les règles de gestion logistiques (iii) les règles de gestion de réparation et d'interventions.

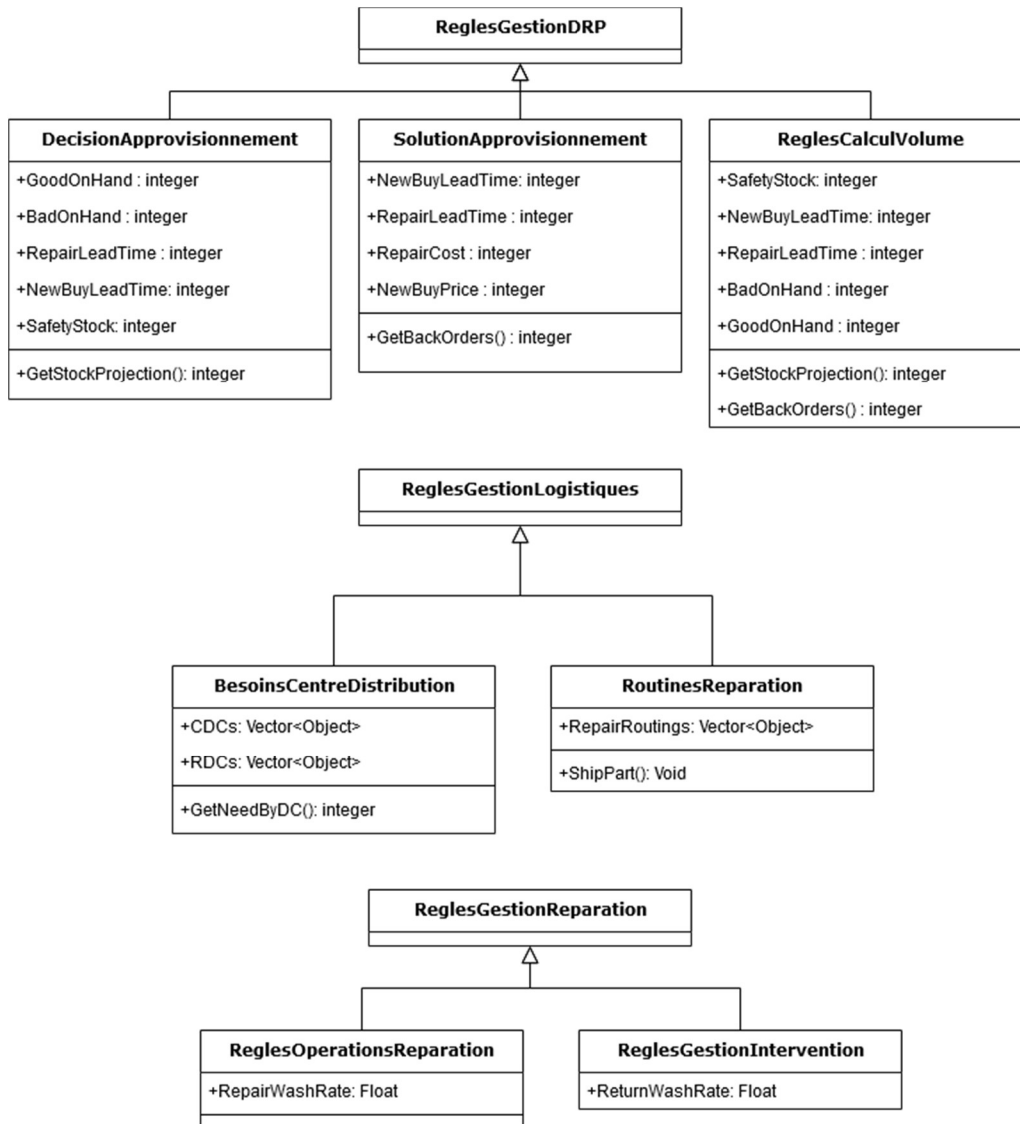


Figure IV.13: Vue globale des différentes règles de gestion du sous-système décisionnel

3.1 Les règles de gestion DRP

Le modèle de réapprovisionnement de base appliqué dans notre contexte est un DRP (ou Distribution Resource Planning en anglais). Ce modèle de base est adapté pour une chaîne logistique de produits finis et semi-finis. Afin de l'appliquer dans notre contexte, ce modèle a été adapté en termes de logique et des contraintes pour répondre aux besoins d'une chaîne logistique de pièces de rechange réparables.

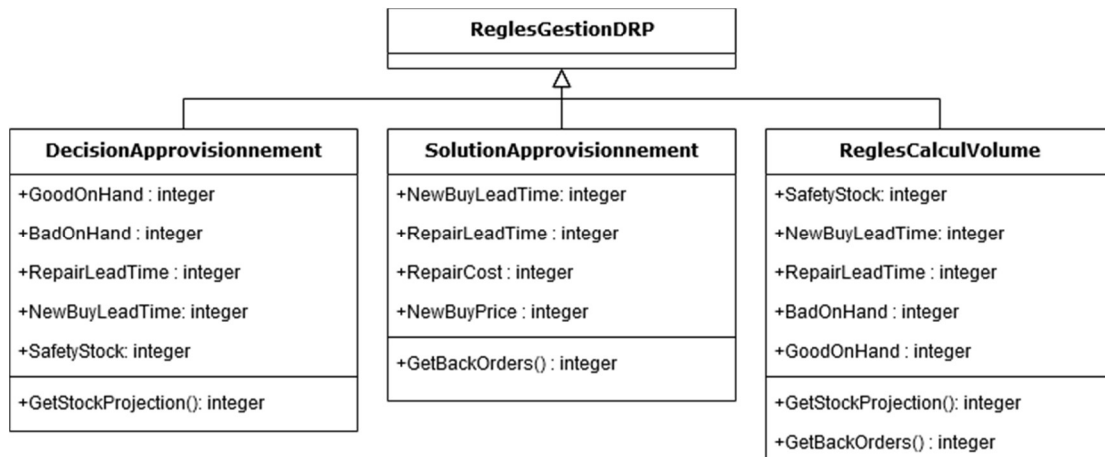


Figure IV.14: Les règles de gestion du modèle DRP pour une chaîne logistique de pièces de rechange réparables

Il existe cinq différences principales entre le modèle DRP de base et celui que nous appliquons dans notre modèle de simulation. Ces éléments sont :

- Les types de pièces en inventaire : au contraire d'un inventaire classique avec des produits de plusieurs références, un inventaire de pièces de rechange réparables contient deux catégories de pièces :
 - o Des pièces en état de fonctionnement : (Good On Hand ou GOH) représentent les pièces neuves ou bien utilisées et reconditionnées via la réparation soit par un centre de réparation interne ou un réparateur externe.
 - o Des pièces défectueuses : (Bad On Hand ou BOH) qui représentent les pièces qui doivent être réparées avant l'utilisation sur un système client. Ces pièces peuvent provenir de plusieurs sources, comme les systèmes clients lors des interventions de maintenance ou des anciens systèmes récupérées auprès des clients lors d'un achat d'un nouveau modèle de systèmes, par exemple.
- Les sources d'approvisionnement : un des points les plus importants dans notre modèle de simulation est le choix entre les sources d'approvisionnement. Au contraire du cas classique où on passe des commandes d'un même fournisseur, pour des pièces de rechange réparables, il existe deux possibilités :

- L'approvisionnement en neuf : (New Buy ou NB) représente le mode d'approvisionnement normal, par l'achat de pièces neuves auprès d'un fournisseur externe, en respectant la quantité de commande minimale et la taille de lots.
 - L'approvisionnement par réparation : (Repair ou RP) représente un mode d'approvisionnement spécifique pour les produits réparables, qui est la transformation des pièces défectueuses en pièces de rechange utilisables ou aussi appelle reconditionnée. Ce mode d'approvisionnement en GOH en utilisant du BOH a plusieurs conditions qu'on va voir par la suite, ce qui ajoute un degré de complexité. Cependant, ceci représente une opportunité financière pour les entreprises ainsi qu'un impact positif sur l'environnement, ce qui le rend intéressant malgré sa complexité.
- Les délais d'approvisionnement : vu que dans notre contexte on a plusieurs sources d'approvisionnement, les délais à prendre en considération pendant les décisions d'approvisionnement sont aussi nombreux :
- Le délai d'approvisionnement en neuf : (New Buy Lead Time ou NBLT) est le délai que le fournisseur impose pour livrer les commandes. Dans le domaine de pièces de rechange de systèmes médicaux, ce délai est généralement de plusieurs mois.
 - Le délai de réparation : (Repair Lead Time ou RLT) est le nombre de jours nécessaire pour transformer une pièce BOH en une pièce GOH prête à être utilisée par les techniciens de maintenance. Au contraire du NBLT, ce délai n'est pas fixé par les fournisseurs mais calculé annuellement en se basant sur le passé des opérations de réparation.
- Les prix des pièces de rechange : Comme pour les délais d'achat ou de réparation, les prix des pièces de rechange dépendent de la solution d'approvisionnement choisie et doivent être pris en considération dans les décisions d'achat.
- Le prix des pièces neuves : (New Buy Price ou NBP) représente le prix d'achat de la pièce neuve auprès du fournisseur.
 - Le coût de réparation : (Repair Cost ou RPC) représente le coût de réparation de la pièce de rechange. Comme pour les délais d'approvisionnement, ce coût n'est pas fixé par le centre de réparation mais calculé annuellement en fonction de plusieurs variables telles que le coût horaire de travail, l'espace alloué pour l'activité de réparation et le coût des composants utilisés pour une référence donnée pendant une année.

- Les taux de succès de réparation et d'installation : dans notre contexte, la réparation de pièces de rechange n'est pas toujours garantie, ceci dépend de l'état de la pièce ainsi que le déroulement des opérations de réparation elles-mêmes. Pour prendre ça en considération, un taux de succès de réparation est calculé annuellement pour chaque référence :
 - Le taux de réussite de réparation : (Repair Wash Rate ou RPWR) signifie le taux de réussite de réparation pour une référence de pièces de rechange.
 - Le taux de réussite de retour des pièces de rechange : (Return Wash Rate ou RTWR) représente le taux de retour des pièces de rechange par les techniciens de rechange, qui est le complément du taux de consommation (CONS + TRWR = 1)

Nous avons regroupé ces différences en trois parties (Figure IV.14) qui sont représentatives du processus d'approvisionnement dans notre contexte. Ces parties sont : (i) la décision d'approvisionnement, (ii) le choix de la solution d'approvisionnement et (iii) l'estimation du volume des commandes.

3.2 Les règles de gestion logistique

Les règles de gestion logistique (Figure IV.15) englobent les algorithmes gérant la logique de la simulation du transport des pièces de rechange dans notre chaîne logistique.

Les pièces en transit sont représentées par une « file » contenant toutes les commandes en route. A chaque unité de temps t_i , le modèle parcourt toutes les pièces et vérifie si la date actuelle est égale à la date estimée d'arrivée de la pièce. Si ce n'est pas le cas, le modèle continue le parcours, sinon il « livre » la pièce à la nouvelle « file » de destination et supprime la commande de la liste en transit.

La destination de la pièce peut être le site du client où elle est « consommée » instantanément, mais peut aussi être une deuxième file où elle est mise en attente, comme celle de la réparation.

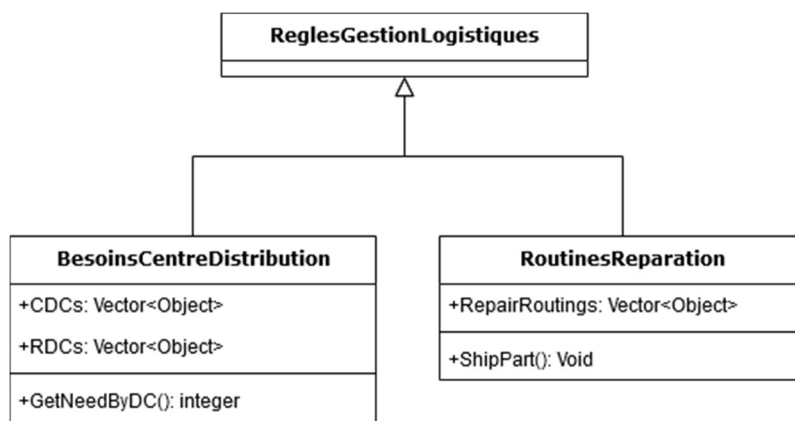


Figure IV.15: Les règles de gestion logistique

3.3 Les règles de gestion de réparation et d'intervention

Pour les règles de gestion de réparation, elles représentent deux aspects principaux (Figure IV.16) :

- **La logique de contrôle de la durée nécessaire de réparation** : Un vecteur de couples (référence, durée) est créé au lancement du modèle à partir de l'historique, pour définir la durée nécessaire pour réparer chaque pièce dans chaque centre de réparation. Cette durée est ensuite utilisée pour décider quand finir la réparation et expédier les pièces aux centres de distribution.
- **Les taux d'échec de réparation (ou le « Repair Wash Rate » RWR)** : Les opérations de réparation ne sont pas toujours effectuées avec succès, à cause de l'état des pièces qui se dégrade dans le temps et avec plusieurs réparations, des pièces peuvent ne plus être réparables. Ceci peut être aussi causé simplement par des problèmes dans les opérations. Cette règle de gestion contient un vecteur de couples (référence, taux d'échec de réparation) qui est comme la durée de transport et de réparation, initialisé au lancement du modèle à partir de l'historique.

Le taux d'échec est situé entre 0 et 1, et pour chaque opération de réparation, il est comparé avec une variable aléatoire suivant une loi uniforme entre 0 et 1, ensuite une décision est prise pour dire si la réparation a été effectuée avec succès ou pas.

Dans le cas où elle est faite avec succès, la pièce est renvoyée en centres de distribution, sinon elle est jetée (simplement supprimée de la file d'attente).

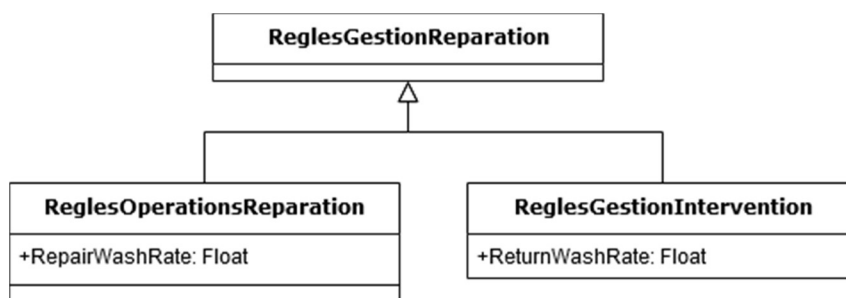


Figure IV.16: Les règles de gestion de réparation et d'interventions

V Le modèle de simulation de la CL-PR-BF – Cas d'application : GEHC

Dans cette section, nous allons nous baser sur les règles de gestion expliquées précédemment pour formuler le modèle de simulation dans le langage algorithmique.

Les figures IV.17, IV.18 et IV.19 présentent une vue globale de notre modèle de simulation, qui est formée par cinq processus parallèles :

- **Le traitement de la base des références de pièces de rechange (Traitement PR) :** qui parcourt le catalogue complet des références des pièces de GEHC, puis pour chacune lance les différents processus de simulation, d'une façon parallèle.
- **Le service de planification :** C'est le processus qui simule les décisions de planification pour chaque pièce, pour décider le moment des commandes, par quel moyen (achat ou réparation) et les quantités.
- **Le transport :** Ce processus est responsable de la simulation du transfert des pièces entre les différentes entités de la chaîne logistique et des durées de transport nécessaire.
- **Le service de réparation (RP) :** Responsable de la simulation des opérations de réparation et leur taux de succès.
- **Les fournisseurs des pièces neuves (NB) :** Ce processus simule les commandes d'achat de pièces neuves, leurs délais d'attente et leur expédition.

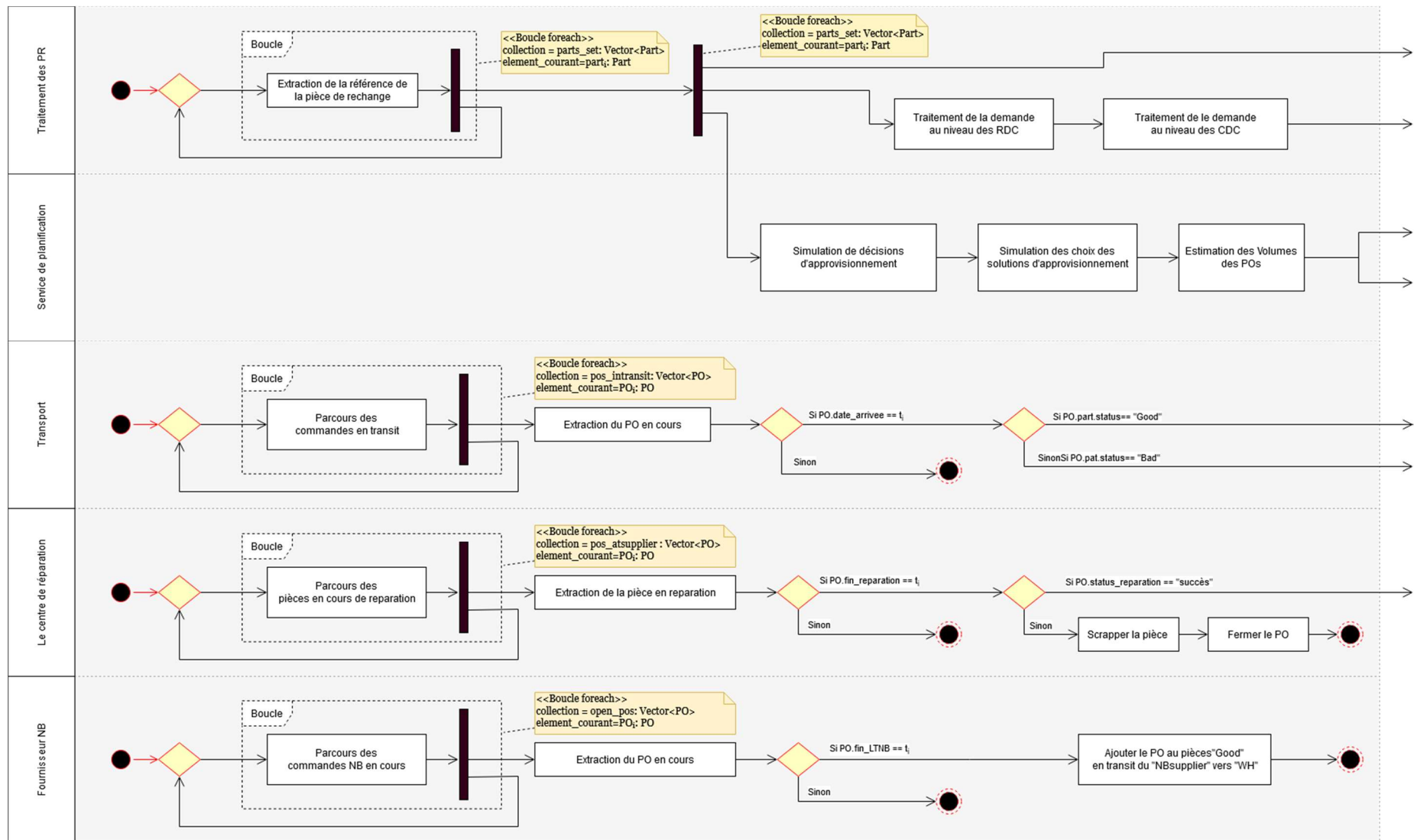


Figure IV.17: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 1)

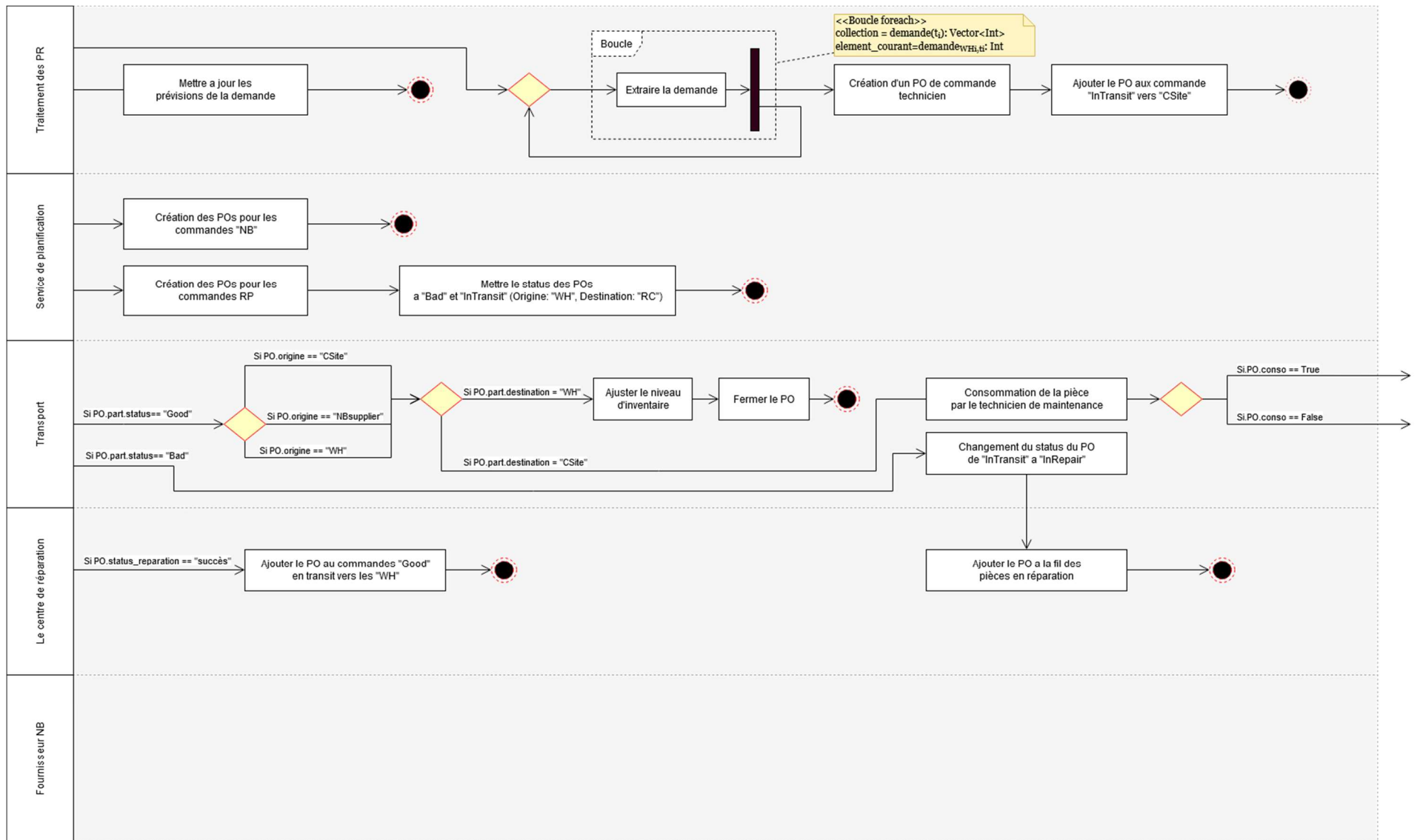


Figure IV.18: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 2)

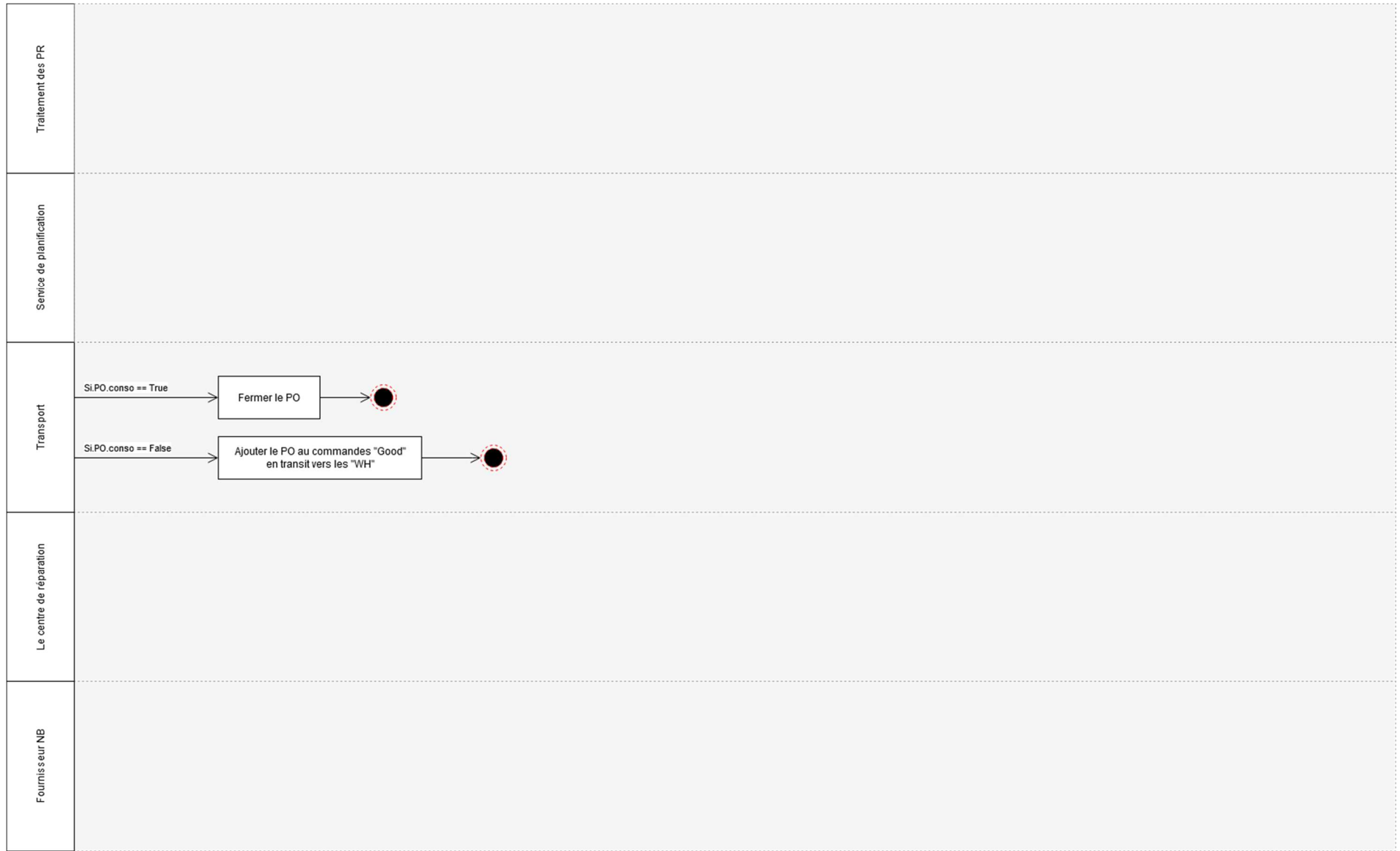


Figure IV.19: Diagramme d'activité (UML) de la vue globale du processus de simulation (partie 3)

L'algorithme 1 présente le processus global qui parcourt les pièces de rechange et lance les différentes fonctions de notre modèle de simulation.

C'est le programme principal qui est lancé au début de la simulation et qui tourne en boucle pendant toute la période concernée. Cet algorithme lance trois fonctions pour chaque pièce pour chaque unité de temps. Après ça, le modèle exécute quatre autres fonctions pour traiter les éléments en attente des itérations précédentes.

Algorithme 1– Boucle de traitement de la demande dans les centres de distribution

```
1.       $t_0 \leftarrow 0$ 
2.       $P_{set}[] \leftarrow getPSet()$ 
3.      Pour  $i$  allant de 0 à N Faire :
4.          Pour chaque  $P_i$  dans  $P_{set}[]$  Faire :
5.              Expedier_PO_Csite( $t_i$ )
6.              Actualiser_Prev_Demande()
7.              Decisions_Approvisionnement( $P_i, t_i$ )
8.          Fin de Pour
9.          Taiter_InTransit( $t_i$ )
10.         Traiter_Pos_NB( $t_i$ )
11.         Traiter_Pos_PR( $t_i$ )
12.         Incrementer_Horloge( $t_i$ )
13.     Fin de Pour
```

La première fonction que le modèle de simulation lance est le processus de traitement des commandes de pièces qui sont passées par le technicien de maintenance (Algorithme 2).

Le rôle de cette fonction est de répondre à la demande des techniciens de maintenance en envoyant des pièces, quand elles sont disponibles. Le processus récupère la valeur de la demande pour le jour t_i , puis crée des commandes pour chaque demande des techniciens de maintenance. Pour envoyer les pièces dans notre modèle de simulation, nous devons préciser l'origine et la destination pour que le modèle puisse estimer le temps nécessaire pour le

transport. Dans le cas de l’algorithme 2 où on livre les pièces aux techniciens, l’origine est un centre de distribution (WH) et la destination est le site du client (CSite).

Algorithme 2– Fonction : Traitement de l’expédition des commandes des techniciens

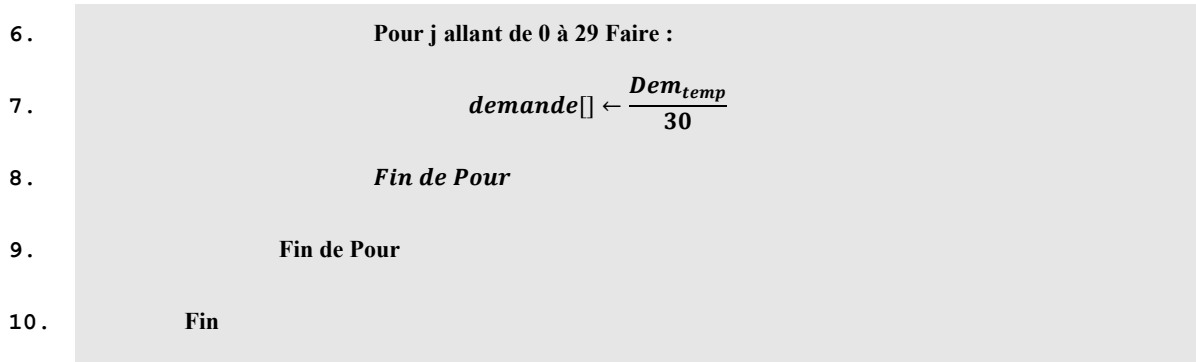
```
1.      Fonction Expedier_PO_CSite( t )
2.
3.      demande[] ← getDemande(t)
4.
5.      Pour chaque d dans demande [] Faire :
6.
7.          Pour i allant de 1 à taille(d) Faire :
8.
9.              PO_i = créerPo(type = `FE`,orig = `WH`,dest = `CSite`)
10.
11.             ChangerStatusPO (PO_i,`inTransit`)
12.
13.          Fin de Pour
14.
15.      Fin de Pour
16.
17.      Fin
```

L’algorithme 3 est responsable de l’actualisation des prévisions de la demande. Ceci est nécessaire dans notre contexte, car nous nous basons sur des prévisions déjà faites pour la demande des pièces de rechange, qui sont actualisées chaque jour. Notre modèle doit être capable de prendre les nouvelles valeurs en considération lors de son exécution et mettre à jours ces résultats.

La demande sur laquelle le modèle se base est mensuelle, puis elle est distribuée par jour, d’une façon uniforme.

Algorithme 3 – Fonction : Actualisation des prévisions de la demande

```
1.      Fonction Actualiser_Prev_Demande()
2.
3.      Dem_Actualisee[] = getDemandeActualise()
4.
5.      demande[] ← []
6.
7.      Pour i allant de 0 à 11 Faire :
8.
9.          Dem_temp ← Dem_Actualisee[i]
```



L’algorithme 4 représente la simulation des décisions d’approvisionnement pour une référence donnée pour un instant t_i donné.

Pour commencer, le modèle récupère d’abord les délais d’approvisionnement pour l’achat (LTNB) et pour la réparation (LTRP). Ensuite, le modèle doit se projeter sur le maximum des deux délais, vu que l’inventaire doit être suffisant pour satisfaire la demande à partir du jour courant et pendant la période du délai maximum d’approvisionnement.

Pour calculer le stock projeté de l’inventaire GOH, le modèle calcule la consommation pour la période $max(LTNB, LTRP)$ qui est représentée par la variable LT_{conso} et puis vérifie si le stock projeté est inférieur au stock de sécurité.

Si c’est le cas, l’algorithme passe à l’étape de choix de critère de décision. Il existe deux critères principaux pour choisir entre la réparation ou l’achat. Ces critères sont le délai d’approvisionnement et le prix de réparation.

Pour choisir sur quoi se baser, le modèle vérifie l’existence de commandes en attente. Ceci est la première priorité, il faut répondre à ces commandes qui sont déjà en attente et pour lesquelles les clients n’ont pas trouvé de pièces disponibles, surtout pour des pièces de rechange utilisées sur des systèmes médicaux. S’il y a des commandes en attente, le système priorise les délais d’approvisionnement comme critère et choisi la solution la plus rapide. Sinon, le système priorise le coût de la solution et prend la moins chère.

Ensuite, le modèle calcule les quantités à passer en commande :

- Pour la réparation en se basant sur le taux moyen d’échec de réparation.
- Pour l’achat en se basant sur la quantité minimale de commande ainsi que la taille de lot imposée par le fournisseur.

Algorithme 4 – Fonction : Simulation des décisions d’approvisionnement

```

1.      Fonction Decisions_Approvisionnement(p, t)
2.
3.       $n \leftarrow \text{Int} \left( \frac{\max(p.LTNB, p.LTRP)}{30} \right)$ 
4.
5.       $p \leftarrow \max(p.LTNB, p.LTRP) \% 30$ 
6.
7.       $LT_{conso} \leftarrow \sum_{i=1}^{n-1} \sum_{j=1}^{30} (p.demande_{(i,j)}) + \sum_{j=1}^p (p.demande_{(n,j)})$ 
8.
9.       $p.GOH_{projete} \leftarrow p.GOH_{actuel} - LT_{conso}$ 
10.
11.      Si  $p.GOH_{projete} < p.SS$  Faire :
12.
13.           $besoin_{base} = p.SS - p.GOH_{projete}$ 
14.
15.          Si  $p.BO > 0$  Faire:
16.
17.               $Besoin_{eoq} = \left( E_{sup} \left( \frac{\max(besoin_{base}, p.MinOQ)}{p.EOQ} \right) + 1 \right) \times p.EOQ$ 
18.
19.              Si  $p.LTRP < p.LTNB$  Faire:
20.
21.                   $BesoinBOH = \frac{besoin_{base}}{1 - p.RWR}$ 
22.
23.                   $Qt_{RP} = \min(p.BO, BesoinBOH)$ 
24.
25.                   $BesoinNB = besoin_{base} - (1 - RWR)Qt_{RP}$ 
26.
27.                   $Qt_{NB} = \left( E_{sup} \left( \frac{\max(BesoinNB, p.MinOQ)}{p.EOQ} \right) + 1 \right) \times p.EOQ$ 
28.
29.                   $Creer\_POs(\text{type} = \text{'RP'}, \text{orig} = \text{'WH'}, \text{dest} = \text{'RC'}, \text{qt} = Qt_{RP})$ 
30.
31.                   $Creer\_POs(\text{type} = \text{'NB'}, \text{orig} = \text{'NBSup'}, \text{dest} = \text{'WH'}, \text{qt} = Qt_{NB})$ 
32.
33.                  Sinon Faire :
34.
35.                       $Creer\_POs(\text{type} = \text{'NB'}, \text{orig} = \text{'NBSup'}, \text{dest} = \text{'WH'}, \text{qt} = besoin_{eoq})$ 
36.
37.                  Fin de Si
38.
39.                  Sinon Faire :
40.
41.                      Si  $p.RPC < p.NBP$  Faire:

```



```

31. 
$$BesoinBOH = \frac{besoin_{base}}{1 - p.RWR}$$

32. 
$$Qt_{RP} = \min(p.BOH, BesoinBOH)$$

33. 
$$BesoinNB = besoin_{base} - (1 - RWR)Qt_{RP}$$

34. 
$$Qt_{NB} = \left( E_{sup} \left( \frac{\max(BesoinNB, p.MinOQ)}{p.EOQ} \right) + 1 \right) \times p.EOQ$$

35.
36. 
$$Creer\_POs(type = `RP`, orig = `WH`, dest = `RC`, qt = Qt_{RP})$$

37. 
$$Creer\_POs(type = `NB`, orig = `NBSup`, dest = `WH`, qt = Qt_{NB})$$

Sinon Faire :

$$Creer\_POs(type = `NB`, orig = `NBSup`, dest = `WH`, qt = besoin_{eoq})$$

Fin de Si
Fin de Si
Fin

```

L'algorithme 5 représente la fonction de traitement des pièces en transport. Comme expliqué précédemment, les pièces en transit sont représentées par une liste contenant toutes les commandes en route.

A chaque itération, le modèle parcourt toutes les pièces et vérifie si la date actuelle est égale à la date estimée d'arrivée de la pièce. Si ce n'est pas le cas, le modèle continue le parcours, sinon il « livre » la pièce à la nouvelle liste de destination et supprime la commande de la liste en transit.

Algorithme 5 – Fonction : Traitement du transport des pièces de rechange

```

1. Fonction Traiter_Pos_Intransit(t)
2.  $pos \leftarrow getIntransitPos()$ 
3. Pour chaque po dans pos Faire:
4.  $Si po.date_{arrivee} == t$  Faire:
5.  $Si po.part.status == `Good`$  Faire:
6.  $Si po.destination == `WH`$  Faire:

```

```

7.          po.part.GOH ← po.part.GOH + 1
8.          fermerPo(po)
9.          Sinon Si po.destination == `CSite` Faire:
10.         success ← (math.rand.uniform > po.part.RR)
11.         Si success == True Faire:
12.         fermerPo(po)
13.         Sinon Faire:
14.         ExpédierPo(orig = `CSite`, Dest
15.                = `WH`)
16.         Fin de Si
17.         Fin de Si
18.         Sinon Faire:
19.         ChangerStatusPO(Status = `InRepair`)
20.         Fin Si
21.         Fin Si
          Fin

```

La dernière fonction dans notre modèle est représentée par l’algorithme 6, qui décrit la logique de traitement des commandes NB (achat en neuf).

Pour traiter les commandes, l’algorithme parcourt les objets de la liste d’attente, puis pour chacune vérifie si la date de livraison est atteinte. Si c’est le cas, le modèle « expédie » les pièces en les ajoutant à la liste des commandes en route vers les entrepôts de stockage. Ceci en définissant l’origine de la pièce et sa destination.

Algorithme 6 – Fonction : Simulation des commandes en cours, d’achat en neuf

```
1.      Fonction Traiter_Pos_NB(t)
2.          pos ← getPendingNBPos()
3.          Pour chaque po dans pos Faire:
4.              Si po.fin_LTNB == t Faire :
5.                  ExpedierPo(orig = `NBSup`, Dest = `WH`)
6.              Sinon
7.                  Continuer
8.              Fin de Si
9.          Fin de Pour
10.     Fin
```

VI Le modèle de résultats de la simulation de la CL-PR-BF

Dans cette partie, nous présentons l’implémentation du modèle de simulation présenté dans la section précédente, les technologies utilisées, ainsi que les résultats du modèle et sa comparaison avec l’historique des opérations de GEHC.

1 Implémentation du modèle de simulation

Afin d’implémenter le modèle de simulation, nous avons utilisé un ensemble de technologies et de bibliothèques simplifiant ainsi la tâche de programmation, surtout les tâches de traitement (jointures, séparation, sélection...) des données de grand volume, ainsi que la gestion du parallélisme dans le modèle. Les principales technologies que nous avons employées sont les suivantes :



Python :

Développé à l’origine par Guido van Rossum en 1989, Python est un langage de programmation de haut niveau, supportant la programmation structurée, fonctionnelle et Orientée Objet. Python est un langage open source, maintenu par une équipe de contributeurs partout dans le monde.



Ray

Développé par une équipe de chercheurs de l'Université de Californie à Berkeley, Ray a comme objectif principal le parallélisme d'un certain nombre de bibliothèques d'apprentissage automatique d'une façon distribuée. Ray ne se limite pas aux tâches d'apprentissage, il permet également de paralléliser toutes les tâches Python afin d'accélérer leur performance.



Pandas

Pandas est une bibliothèque écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques et de séries temporelles.



Jupyter Notebooks

Jupyter est un outil web permettant d'avoir une interface intuitive pour programmer dans plus de 40 langages de programmation, dont Python, Julia, Ruby, R, ou encore Scala.

Jupyter est principalement utilisé pour créer des « notebooks » contenant à la fois des programmes exécutables dans le navigateur web, ainsi que des figures, graphes et du texte, ce qui est très pratique pour faire des analyses de données.



NumPy

NumPy est une extension du langage de programmation Python, destinée à manipuler des matrices ou tableaux multidimensionnels ainsi que des fonctions mathématiques opérant sur ces tableaux.

Le modèle que nous avons implémenté est intégré au système d'informations de GEHC. A chaque lancement, l'outil se connecte automatiquement à un datawarehouse que nous avons mis en place pour cet objectif, contenant la base de données complètes des pièces de rechange

de l'entreprises, toutes leurs informations, ainsi que l'historique des opérations qui y sont liées (Figure IV.20).

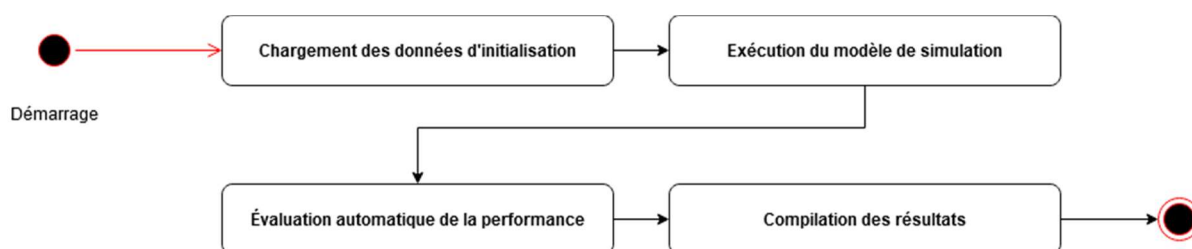


Figure IV.20: Etapes d'exécution du modèle de simulation

Le résultat est une application exécutable permettant la simulation des opérations sur le catalogue complet des pièces de rechange de GEHC. Par défaut, l'application simule les opérations sur 6 mois et génère le résultat dans des fichiers CSV qui peuvent être traités et visualisés à l'aide de logiciels tels que Microsoft Excel ou Google Sheets.

En plus des rapports finaux de la simulation, l'application offre la possibilité de visualiser l'évolution des variables clés pour une référence donnée, tels que les niveaux d'inventaire, le nombre de pièce en réparation et les commandes en cours (Figure IV.21).

```

- ELAPSED TIME: 0years, 0 months, 8days
- PART: 2355552
- REPAIR LENGTH : 15.0
- REPAIR COST : 340.47
- PROCUREMENT LENGTH : 10.0
- PART COST : 626.45
#####
#                                     PART STATE                                     #
#-----#-----#-----#-----#-----#-----#-----#-----#-----#-----#
#   GOH   |   BOH   |   DIT   |   RIT   |   RET RATE |   WSH RATE   |
#-----#-----#-----#-----#-----#-----#-----#-----#
# 5770    |   177    |    0     |    7     |    13      |    1         |
#-----#-----#-----#-----#-----#-----#-----#
#   36M_DEMND |   24M_DEMND |   12M_DEMND |   CUR_MTH_FRCST |   CUR_MTH_DEM   |
#-----#-----#-----#-----#-----#-----#-----#
# 3284    |   2641    |   2700    |   172.31 |    48      |
#-----#-----#-----#-----#-----#-----#
#                                     IN_REPAIR |                                     ON_ORDR   |
#-----#-----#-----#-----#-----#-----#
# 4       |           |           |           |           |           |
#####
- W_TO_RC TRANSPORT LOAD: [.....] - [00%]
- RC_TO_W TRANSPORT LOAD: [.....] - [00%]
- REPAIR CENTER LOAD:    [=====.....] - [12%]
  
```

Figure IV.21: L'outil de simulation en cours d'exécution. Le modèle simule les opérations pour toutes les pièces existantes dans la base de données, et offre la possibilité de visualiser l'évolution des différentes variables pour une pièce en temps réel.

2 Résultats du modèle de simulation sur les pièces de rechange de GEHC

Afin de tester et analyser les différents scénarios du modèle de simulation, sur les différentes pièces de GEHC, nous divisons le test en trois catégories de pièces :

- **Catégorie 1** : Les pièces pour lesquelles l'achat du neuf (NB) est la seule solution
- **Catégorie 2** : Les pièces de rechange pour lesquelles la réparation est un meilleur choix, et pour lesquelles on dispose d'un inventaire BOH (pièces défectueuse) suffisant pour satisfaire la demande pendant 6 mois.
- **Catégorie 3** : Les pièces de rechange pour lesquelles la réparation est optimale, cependant, le stock BOH n'est pas suffisant pour satisfaire la demande. Pour ce type de pièces, le modèle doit combiner la réparation avec l'achat du neuf d'une façon optimale afin de respecter les contraintes imposées (Safety Stock, EOQ, MinOQ...)

2.1 Catégorie I : Les pièces avec l'achat seul (NB) comme solution optimale

Pour cette catégorie, nous avons filtré un ensemble de pièces avec les critères nécessaires pour que le modèle prenne la décision d'achat seul. Ces critères sont le délai d'approvisionnement de l'achat qui est plus rapide que la réparation, ou/et le prix d'achat qui est moins cher ou bien que les pièces pour lesquelles nous ne disposons pas d'inventaire défectueux à réparer.

La figure IV.22 montre le résultat du modèle sur trois pièces différentes. Pour les deux premières pièces, le modèle calcule le stock projeté en fonction du niveau d'inventaire initial et de la prévision de la demande des techniciens. Ensuite le modèle n'a pas fait de commandes jusqu'à ce que le stock projeté descende au niveau du stock de sécurité.

Pour la dernière pièce, nous testons le modèle sur une pièce avec un stock projeté au-dessus du stock de sécurité dès le début de la simulation. Dans ce cas, le modèle passe dès la première itération des commandes pour satisfaire la demande. Nous remarquons que les niveaux d'inventaire s'ajustent par la suite et ne descendent pas sous le niveau de sécurité.

Nous remarquons que le niveau de stock GOH minimal n'est pas parfaitement aligné avec le stock de sécurité, ceci est lié au fait que le modèle doit respecter la taille de lot des commandes, et avec la variation aléatoire de la consommation des pièces.



Figure IV.22: Résultat du modèle de simulation sur la catégorie I des pièces de rechange (NB)

2.2 Catégorie II : Les pièces avec la réparation seule (RP) comme solution optimale

Pour cette catégorie, nous avons pris un échantillon de pièces pour lesquelles :

- La réparation est une solution optimale, en se basant sur la différence du prix entre la réparation et l'achat des pièces neuves, ainsi que sur le délai d'approvisionnement.
- Le stock défectueux BOH est suffisant pour être réparé et répondre à la demande.



Figure IV.23: Résultat du modèle de simulation sur la catégorie II des pièces de rechange (RP)

Pour cette catégorie de pièces, au lancement du simulateur, le modèle avance dans le temps jusqu'à la détection de besoin d'approvisionnement, ensuite envoi des pièces BOH aux centres de réparation. Ceci se voit dans la chute du niveau du stock des pièces défectueuses représentée par la courbe orange.

Le niveau du stock GOH continue à baisser en fonction de la consommation. Ensuite, après le déroulement de la durée d'approvisionnement en réparation (LTRP), les pièces sont réparées et expédiées aux entrepôts, ce qui fait lever le niveau du stock GOH et illustré par les pics de la courbe bleue.

Un point à noter est que théoriquement, en absence d'achat de pièces neuves, le niveau des pièces défectueuses baisse dans le temps, dû au taux d'échecs de réparation pour chaque référence. Quand la réparation échoue, la pièce est éliminée et n'est plus utilisée. Ceci se voit bien dans le résultat de la simulation au travers de la tendance à la baisse de la courbe orange (Figure IV.24).

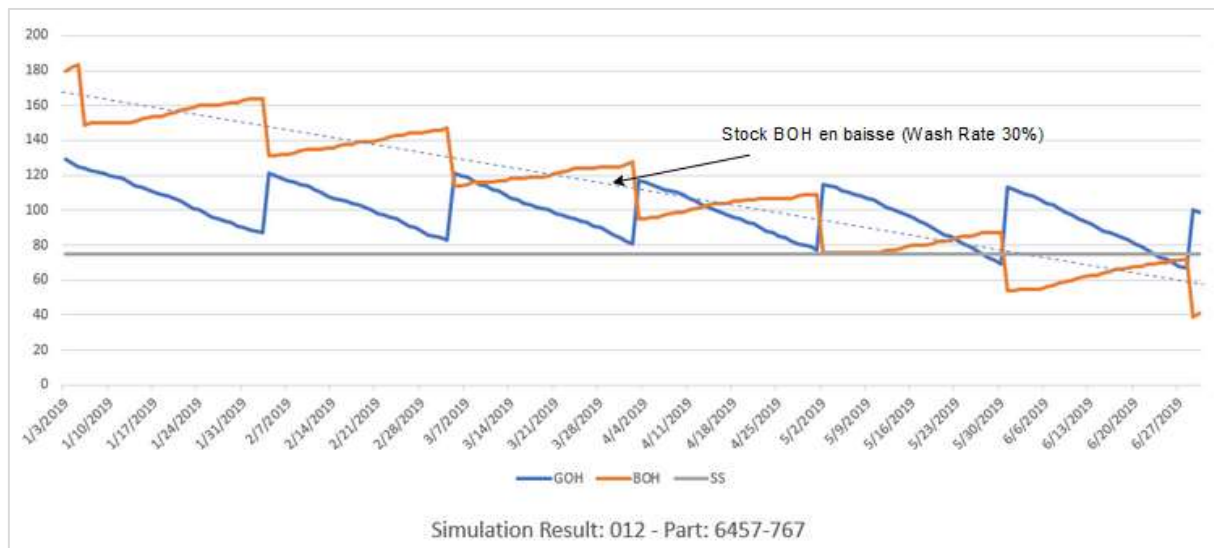


Figure IV.24: Niveau du BOH en baisse dû au taux d'échec des réparations.

2.3 Catégorie III : Les pièces nécessitant une combinaison entre la réparation (RP) et l'achat de pièces neuves (NB)

Pour la dernière catégorie des références, la solution optimale est la réparation qui est moins chère et plus rapide. Cependant, cet échantillon de référence ne dispose pas d'un niveau d'inventaire BOH suffisant pour satisfaire la totalité de la demande. Pour compenser le manque du BOH, le modèle doit passer des commandes d'achat NB.

Les commandes de réparation et d'achat du neuf n'ont pas les mêmes délais d'approvisionnement. Le modèle doit calculer les besoins, le stock projeté et passer les deux types de commandes aux bons moments pour ne pas descendre au-dessous du stock de sécurité

et en même temps ne pas avoir un surstock, tout en respectant les tailles de lots des commandes fournisseurs NB des pièces et la taille minimale de commandes.

La figure IV.25 montre les décisions que le modèle de simulation a prises. Pour la première pièce, le stock projeté à la première itération est au-dessous du stock de sécurité. Le modèle a directement expédié le stock complet BOH en réparation même s'il n'est pas suffisant, puis a compensé par une commande en neuf. Le délai d'approvisionnement LTNB est supérieur au LTRP ce qui fait que le niveau d'inventaire GOH est descendu à zéro, en attendant l'arrivée des commandes de pièces neuves, en générant des backorders.



Figure IV.25: Résultat du modèle de simulation sur des pièces avec une combinaison RP, NB comme solution adéquate

Comme pour la section précédente avec la réparation seule, le niveau d'inventaire BOH baisse en fonction du taux d'échec de réparation de chaque référence de pièces. Cependant, dans ce cas, l'inventaire est compensé par la passation des commandes en achat de pièces neuves.

3 Performance du modèle pour la prévision de la charge de réparation

Dans cette section, nous allons présenter le test de performance du modèle que nous avons implémenté sur les données de GEHC. Nous allons commencer par le processus de test du modèle, puis nous allons présenter les résultats et les analyser.

3.1 Processus de test de la performance du modèle

L'objectif du modèle de simulation est de calculer l'estimation des volumes de pièces de rechange à réparer qui seront reçues dans les centres de réparation. Afin d'évaluer leur performance, nous allons suivre la méthodologie suivante :

- Sur un échantillon de cinquante références de pièces de rechange, nous allons initialiser le modèle avec des données historiques (Par exemple, le 1^{er} Mars 2019).
- Puis, nous allons lancer le modèle pour simuler 6 mois d'opérations avec plusieurs combinaisons de ces paramètres :
 - Variation de la demande : Faible, moyenne ou haute
 - Durées de transport :
 - Durées fixes pour chaque origine et destination
 - Durées variables en fonction de l'historique de la référence et la destination (Moyenne, P95).
 - Délais de réparation :
 - En utilisant le RLT fixe
 - En utilisant les données historique (Moyenne, P95).
 - Délais d'approvisionnement :
 - En utilisant le NBLT fixe
 - En utilisant les données historique (Moyenne, P95).
- Ensuite, nous allons comparer le résultat de la simulation avec les vraies données historiques de GEHC et évaluer sa performance en se basant sur les métriques : MSE, RMSE et R2.

Note : Le P95 centile représente la valeur maximale des 95% points inférieurs d'un échantillon. Cette valeur est généralement utilisée pour éviter les données aberrantes.

3.2 Métriques d'évaluation de la performance des modèles

L'évaluation de la performance des modèles de prévision consiste à évaluer l'écart entre la valeur estimée par le modèle pour une observation et la valeur réelle associée à cette dernière.

Nous avons décidé d'utiliser quatre métriques couramment utilisées dans les études de performance de modèles de régression (Savic *et al.* (1991), Prairie *et al.* (1996)). Nous allons utiliser les mêmes métriques pour évaluer tous les modèles de notre travail.

3.2.1 Erreur quadratique moyenne (ou Mean Squared Error en anglais : MSE)

La MSE ou erreur quadratique moyenne est l'une des mesures les plus utilisées pour calculer la précision pour les modèles de régression. Il s'agit simplement de la moyenne de la différence au carré entre la valeur cible et la valeur prédite par le modèle de régression.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

3.2.2 Racine carrée de l'erreur quadratique moyenne (ou RMSE)

La RMSE est aussi une métrique classique pour les tâches de régression. Elle est définie par la racine carrée de la différence quadratique moyenne entre la valeur cible et la valeur prédite par le modèle. Elle est préférable dans certains cas à la MSE car, plus sensible aux écarts des erreurs, elle entraînera une pénalisation plus importante aux erreurs importantes. Cela implique que la RMSE est utile lorsque des erreurs importantes ne sont pas souhaitées.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2}$$

3.2.3 L'erreur R²

Le coefficient du R² nous aide à comparer notre modèle actuel avec une ligne de base constante et nous indique à quel point notre modèle est meilleur. La ligne de base constante est choisie en prenant la moyenne des données et en traçant une ligne à la moyenne. R² est un score sans échelle qui implique que peu importe que les valeurs des erreurs, le R² sera toujours inférieur ou égal à 1. Une valeur proche de 1 signifie une bonne adéquation au modèle.

$$R^2 = 1 - \frac{MSE(\text{modele})}{MSE(\text{reference})} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

3.3 Résultats

Nous commençons l'analyse par la comparaison entre l'historique des commandes de réparation du 1^{er} Mars 2019 au 31 Aout 2019, avec le résultat du modèle de simulation sur la même durée.

Le tableau IV.1 présente les différentes combinaisons des configurations des différentes parties du modèle de simulation, avec la performance de chacune basée sur l'application du modèle sur les vraies données historiques de GEHC.

Pour nos analyses de résultats, sur des pièces individuelles ou le catalogue complet, nous allons utiliser la combinaison avec le meilleur score du tableau IV.1.

DEM_RAND_VAR	TRANSPORT_CONFIG	REPAIR_CONFIG	NB_CONFIG	MSE	RMSE	R2
LOW	FIXED	FIXED_RLT	FIXED_NBLT	1296.00	36.00	0.25%
LOW	FIXED	FIXED_RLT	VAR_AVG	1936.00	44.00	2.91%
LOW	FIXED	FIXED_RLT	P95	1369.00	37.00	0.69%
LOW	FIXED	VAR_AVG	FIXED_NBLT	841.00	29.00	1.00%
LOW	FIXED	VAR_AVG	VAR_AVG	1225.00	35.00	3.47%
LOW	FIXED	VAR_AVG	P95	225.00	15.00	21.39%
LOW	FIXED	P95	FIXED_NBLT	1600.00	40.00	3.54%
LOW	FIXED	P95	VAR_AVG	1521.00	39.00	1.68%
LOW	FIXED	P95	P95	196.00	14.00	18.17%
LOW	VAR_AVG	FIXED_RLT	FIXED_NBLT	900.00	30.00	4.22%
LOW	VAR_AVG	FIXED_RLT	VAR_AVG	1600.00	40.00	1.94%
LOW	VAR_AVG	FIXED_RLT	P95	144.00	12.00	23.15%
LOW	VAR_AVG	VAR_AVG	FIXED_NBLT	256.00	16.00	7.51%
LOW	VAR_AVG	VAR_AVG	VAR_AVG	324.00	18.00	14.09%
LOW	VAR_AVG	VAR_AVG	P95	900.00	30.00	1.05%
LOW	VAR_AVG	P95	FIXED_NBLT	1849.00	43.00	3.01%
LOW	VAR_AVG	P95	VAR_AVG	289.00	17.00	15.92%
LOW	VAR_AVG	P95	P95	289.00	17.00	4.17%
LOW	P95	FIXED_RLT	FIXED_NBLT	225.00	15.00	4.38%
LOW	P95	FIXED_RLT	VAR_AVG	2025.00	45.00	2.45%
LOW	P95	FIXED_RLT	P95	1849.00	43.00	0.21%
LOW	P95	VAR_AVG	FIXED_NBLT	900.00	30.00	1.46%
LOW	P95	VAR_AVG	VAR_AVG	961.00	31.00	1.05%
LOW	P95	VAR_AVG	P95	1296.00	36.00	2.08%
LOW	P95	P95	FIXED_NBLT	225.00	15.00	22.64%
LOW	P95	P95	VAR_AVG	1936.00	44.00	1.58%
LOW	P95	P95	P95	1444.00	38.00	0.81%
MEDIUM	FIXED	FIXED_RLT	FIXED_NBLT	1296.00	36.00	1.54%
MEDIUM	FIXED	FIXED_RLT	VAR_AVG	1600.00	40.00	0.19%

MEDIUM	FIXED	FIXED_RLT	P95	1849.00	43.00	2.91%
MEDIUM	FIXED	VAR_AVG	FIXED_NBLT	361.00	19.00	0.40%
MEDIUM	FIXED	VAR_AVG	VAR_AVG	324.00	18.00	5.70%
MEDIUM	FIXED	VAR_AVG	P95	729.00	27.00	0.87%
MEDIUM	FIXED	P95	FIXED_NBLT	144.00	12.00	23.15%
MEDIUM	FIXED	P95	VAR_AVG	900.00	30.00	4.02%
MEDIUM	FIXED	P95	P95	1849.00	43.00	1.79%
MEDIUM	VAR_AVG	FIXED_RLT	FIXED_NBLT	841.00	29.00	3.11%
MEDIUM	VAR_AVG	FIXED_RLT	VAR_AVG	169.00	13.00	11.86%
MEDIUM	VAR_AVG	FIXED_RLT	P95	289.00	17.00	5.58%
MEDIUM	VAR_AVG	VAR_AVG	FIXED_NBLT	400.00	20.00	9.54%
MEDIUM	VAR_AVG	VAR_AVG	VAR_AVG	1225.00	35.00	1.40%
MEDIUM	VAR_AVG	VAR_AVG	P95	1156.00	34.00	3.79%
MEDIUM	VAR_AVG	P95	FIXED_NBLT	441.00	21.00	11.20%
MEDIUM	VAR_AVG	P95	VAR_AVG	361.00	19.00	7.81%
MEDIUM	VAR_AVG	P95	P95	2025.00	45.00	2.50%
MEDIUM	P95	FIXED_RLT	FIXED_NBLT	529.00	23.00	6.74%
MEDIUM	P95	FIXED_RLT	VAR_AVG	1936.00	44.00	2.73%
MEDIUM	P95	FIXED_RLT	P95	1764.00	42.00	2.79%
MEDIUM	P95	VAR_AVG	FIXED_NBLT	400.00	20.00	6.57%
MEDIUM	P95	VAR_AVG	VAR_AVG	1156.00	34.00	3.14%
MEDIUM	P95	VAR_AVG	P95	1024.00	32.00	0.94%
MEDIUM	P95	P95	FIXED_NBLT	1936.00	44.00	2.66%
MEDIUM	P95	P95	VAR_AVG	144.00	12.00	2.09%
MEDIUM	P95	P95	P95	1156.00	34.00	2.27%
HIGH	FIXED	FIXED_RLT	FIXED_NBLT	144.00	12.00	7.69%
HIGH	FIXED	FIXED_RLT	VAR_AVG	256.00	16.00	10.64%
HIGH	FIXED	FIXED_RLT	P95	729.00	27.00	6.45%
HIGH	FIXED	VAR_AVG	FIXED_NBLT	1225.00	35.00	4.57%
HIGH	FIXED	VAR_AVG	VAR_AVG	2025.00	45.00	3.00%
HIGH	FIXED	VAR_AVG	P95	441.00	21.00	0.07%
HIGH	FIXED	P95	FIXED_NBLT	729.00	27.00	7.32%
HIGH	FIXED	P95	VAR_AVG	2025.00	45.00	0.75%
HIGH	FIXED	P95	P95	1296.00	36.00	2.69%
HIGH	VAR_AVG	FIXED_RLT	FIXED_NBLT	961.00	31.00	0.34%
HIGH	VAR_AVG	FIXED_RLT	VAR_AVG	900.00	30.00	4.07%
HIGH	VAR_AVG	FIXED_RLT	P95	256.00	16.00	0.71%
HIGH	VAR_AVG	VAR_AVG	FIXED_NBLT	144.00	12.00	22.20%
HIGH	VAR_AVG	VAR_AVG	VAR_AVG	576.00	24.00	3.53%
HIGH	VAR_AVG	VAR_AVG	P95	576.00	24.00	0.62%
HIGH	VAR_AVG	P95	FIXED_NBLT	1225.00	35.00	3.87%
HIGH	VAR_AVG	P95	VAR_AVG	576.00	24.00	8.02%
HIGH	VAR_AVG	P95	P95	1764.00	42.00	0.85%
HIGH	P95	FIXED_RLT	FIXED_NBLT	1024.00	32.00	2.75%
HIGH	P95	FIXED_RLT	VAR_AVG	400.00	20.00	9.84%
HIGH	P95	FIXED_RLT	P95	169.00	13.00	2.71%

HIGH	P95	VAR_AVG	FIXED_NBLT	484.00	22.00	3.38%
HIGH	P95	VAR_AVG	VAR_AVG	196.00	14.00	11.43%
HIGH	P95	VAR_AVG	P95	1849.00	43.00	0.50%
HIGH	P95	P95	FIXED_NBLT	676.00	26.00	3.72%
HIGH	P95	P95	VAR_AVG	576.00	24.00	0.75%
HIGH	P95	P95	P95	324.00	18.00	2.86%

Tableau IV.1: Présentation des différentes combinaisons des différentes parties du modèle de simulation avec leur performance (MSE, RMSE et R2).

3.3.1 Performance sur des pièces individuelles

Nous présentons d'abord les résultats sur des pièces individuelles. Nous avons choisi des références représentatives de la généralité des résultats, ensuite nous allons passer aux résultats sur la globalité des pièces de rechange.

Sur la figure IV.26 (A), la courbe en bleue représente les vraies quantités créées pour chaque jour de notre période de test. Les volumes générés par le module de simulation sont présentés par la courbe noire.

Nous remarquons qu'un écart non négligeable est présent entre les deux courbes (le graphe B de la courbe IV.26) : Les commandes créées par le modèle de simulation sont des quantités plus importantes, mais sont moins fréquentes que les commandes présentes dans l'historique.

Une deuxième différence entre les deux courbes est le fait que le modèle de simulation passe les commandes d'une façon quasi périodique. Ceci est dû à la nature de l'algorithme du modèle DRP qui se projette sur le délai d'approvisionnement, et puisque la demande est quasi constante dans le modèle de simulation, le résultat donne un cycle périodique de passation de commandes de réparation, au contraire des vraies données où on voit que les commandes ne sont pas passées de façons ponctuelles mais graduellement.

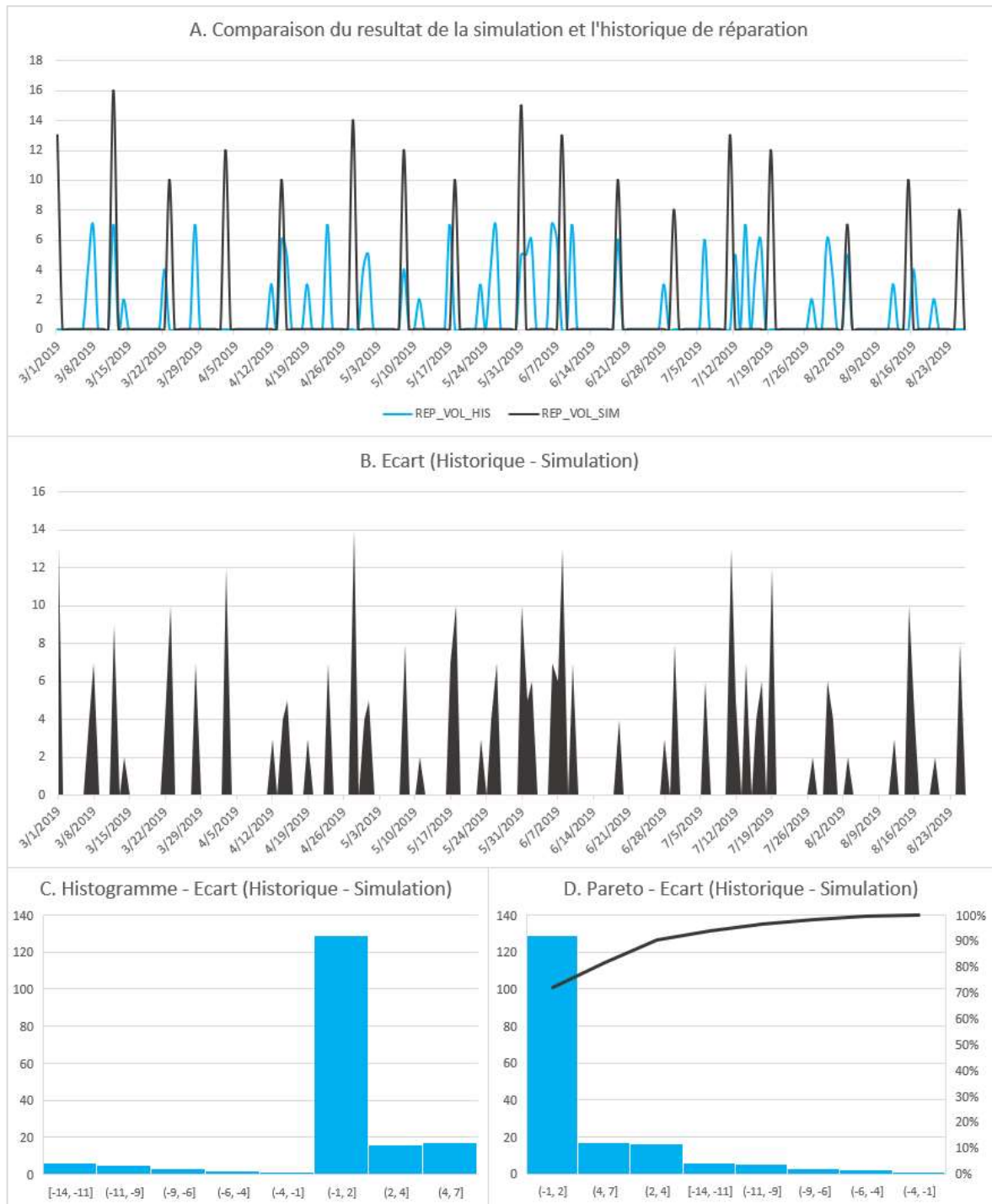


Figure IV.26: Comparaison du résultat de la simulation avec l'historique des opérations de réparation de GEHC, pour une référence de pièces.

Théoriquement, si la prévision de la demande des pièces dans le modèle de simulation est constante, le modèle doit générer des commandes avec des quantités constantes. Cependant, pour notre cas, nous avons implémenté une variation aléatoire (avec 3 niveaux ajustables : faible, moyenne et haute variation), ce qui fait que la consommation des pièces n'est pas constante et que le modèle s'ajuste pour la compenser.

Malgré l'écart entre le modèle de simulation et la réalité, la différence entre les quantités par commande reste globalement faible. Les graphes (C) et (D) de la figure IV.26 présentent un histogramme et un diagramme Pareto de la différence entre les vrais volumes de réparation et le résultat de la simulation. Nous voyons que, dans 70% des cas, la différence entre les deux valeurs se situe entre -1 et 2, ce qui est relativement faible. Cependant, la prévision des volumes de réparation se fait sur plusieurs mois et une différence faible peut se cumuler et générer un grand écart sur une grande période, par exemple de 6 mois.

Pour voir la performance du modèle sur une unité de temps où les différences se cumulent, nous présentons la figure IV.27 présentant une comparaison entre le résultat de la simulation et l'historique, par semaine (A) et par mois (B).

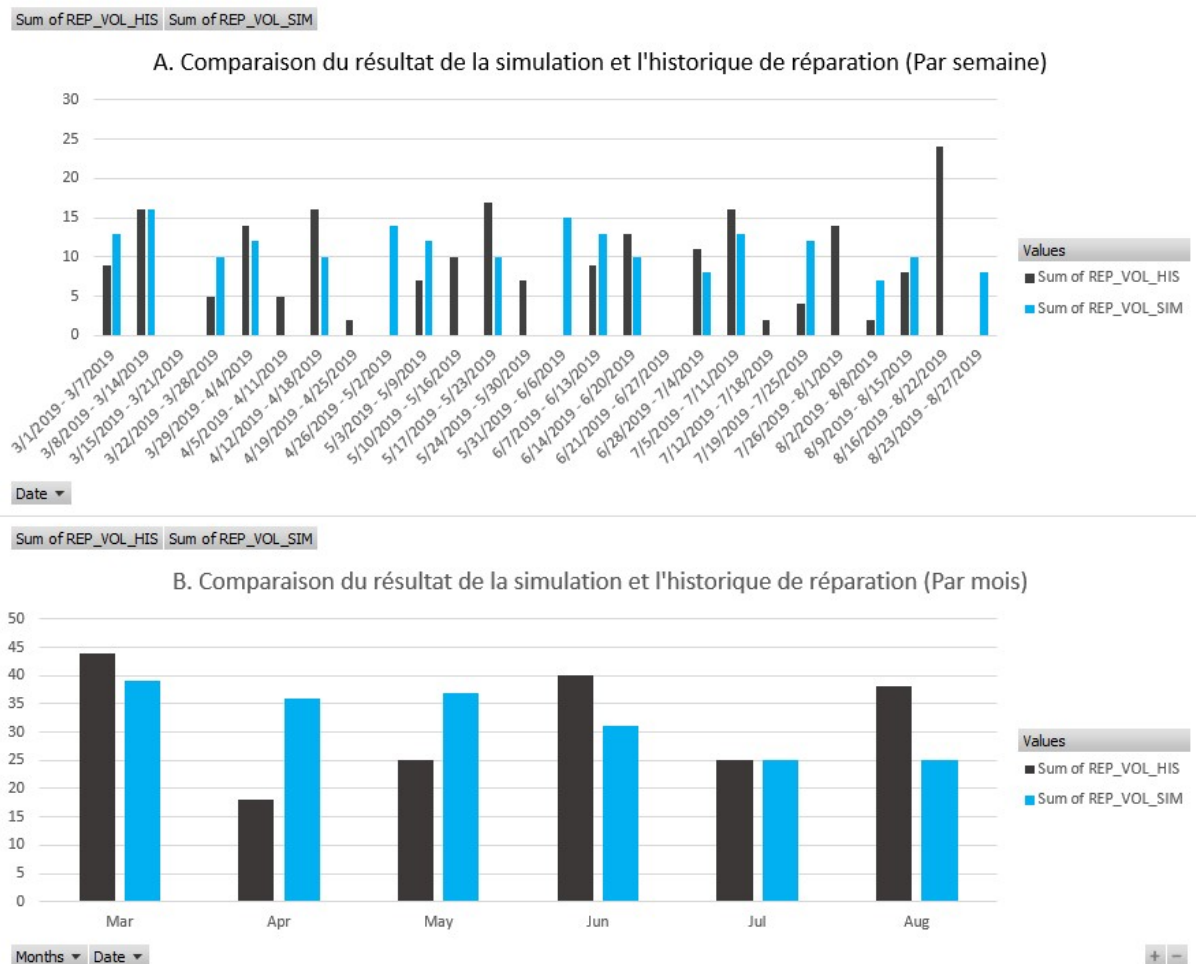


Figure IV.27: Les volumes de réparation créés par le modèle de simulation en comparaison avec l'historique, par semaine (A) et par mois (B).

La figure IV.28 présente le volume cumulé créé par le modèle de simulation ainsi que celui présent dans l'historique, effectué par les planners à GEHC. Nous remarquons que l'écart devient plus important avec le temps et représente plus de 40% de différence.

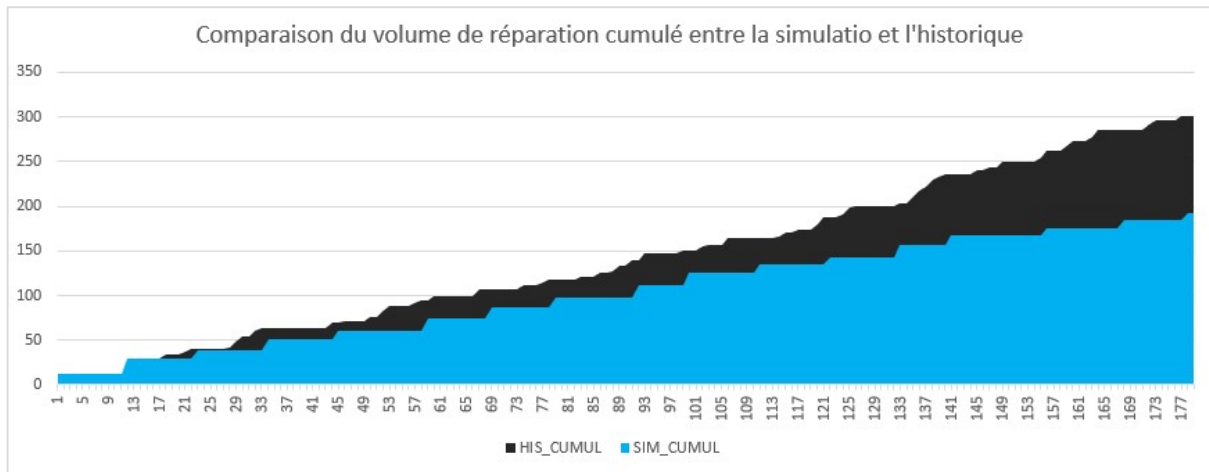


Figure IV.28: Comparaison du volume de réparation cumulé entre la simulation et l'historique

3.3.2 Performance sur l'ensemble des pièces du catalogue GEHC

Comme pour sur les pièces individuelles, nous commençons par la comparaison de l'historique des opérations de réparation de GEHC avec le résultat du modèle de simulation sur la période du 1^{er} Mars 2019 au 31 Aout 2019.

La première chose que nous constatons est la quasi-stabilité des prévisions fournies par le modèle de simulation, représentées par la courbe noire (Figure IV.29). Celles-ci ne sont pas représentatives des vraies décisions faites par les responsables de planification. Les volumes prévus par le modèle de simulation ont une tendance constante vu que le modèle se base sur un algorithme déterministe. Nous remarquons un volume important au début du graphe en comparaison avec le reste : ceci est parce qu'au lancement, le modèle calcule les besoins initiaux et passe les commandes pour les références pour lesquelles le stock projeté est inférieur au stock de sécurité.

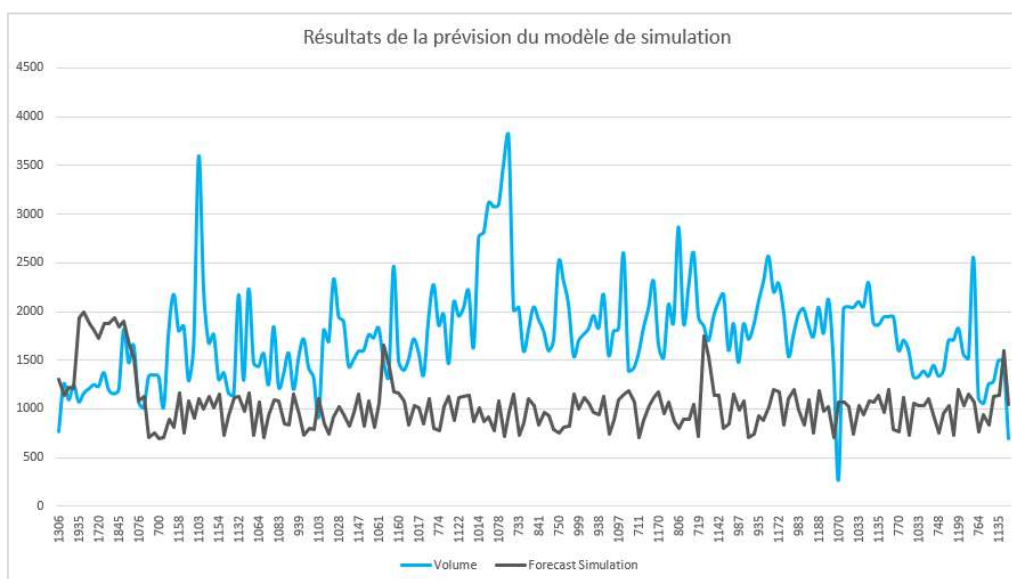


Figure IV.29: Comparaison du résultat de la simulation avec l'historique des opérations de réparation de GEHC

En plus de la stabilité du modèle de simulation, ce dernier ne permet pas de prévoir la tendance ou les pics des demandes de réparation, ce qui est très important pour les centres de réparation car elle permet aux responsables de préparer l'équipe de réception, l'espace et le personnel en cas d'augmentation brusque de la charge.

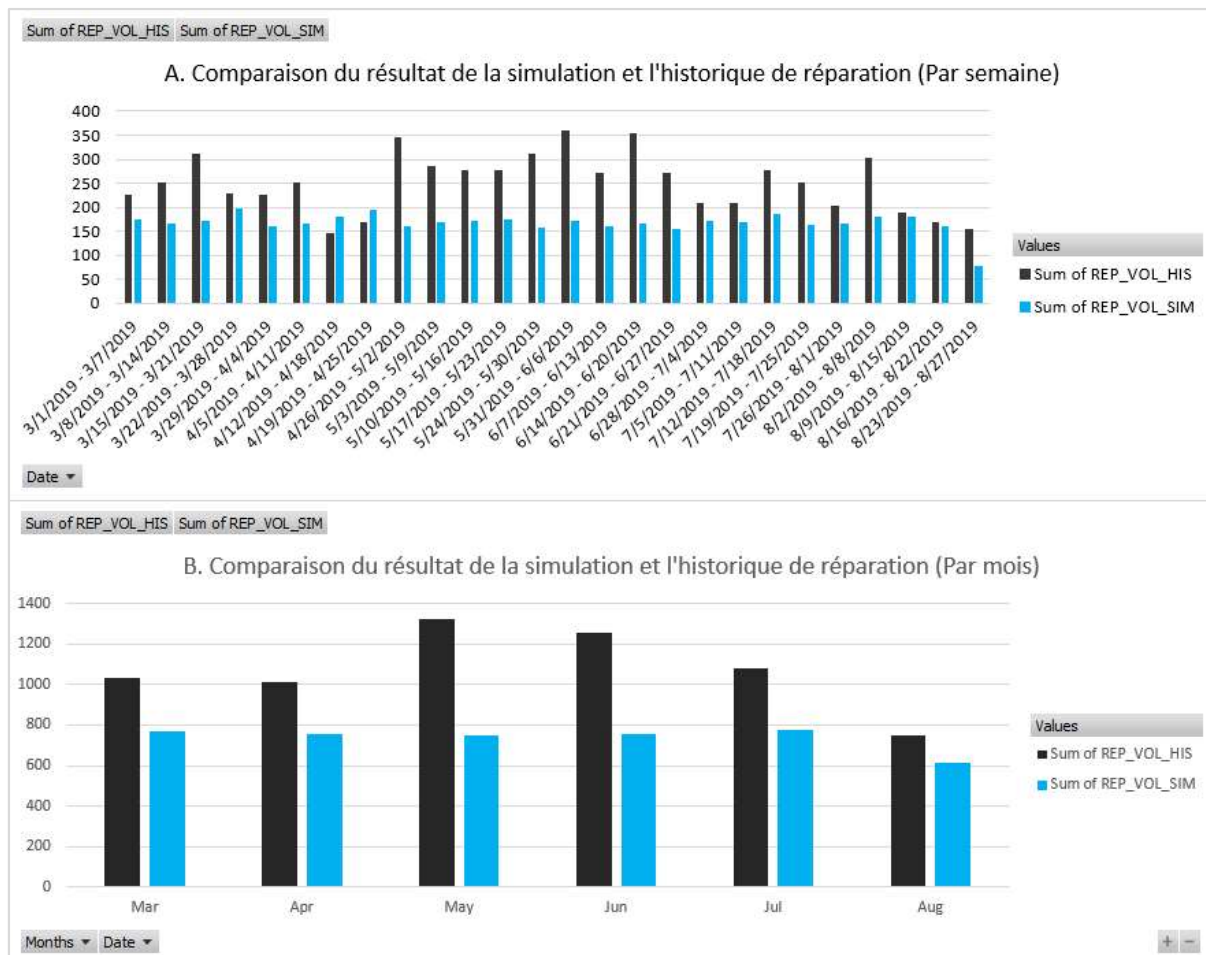


Figure IV.30: Visualisation des écarts entre la vraie demande des pièces de rechange de GEHC et le résultat de la simulation

Comme pour le résultat sur une pièce individuelle, la figure IV.30 montre l'écart avec des pas de prévision de semaines et de mois. En se basant sur cette dernière, on peut voir que le modèle n'est pas capable de prévoir les tendances (augmentation au milieu de notre horizon), et les pics qui sont présents dans la réalité et qui représentent la plus grande difficulté pour les centres de réparation.

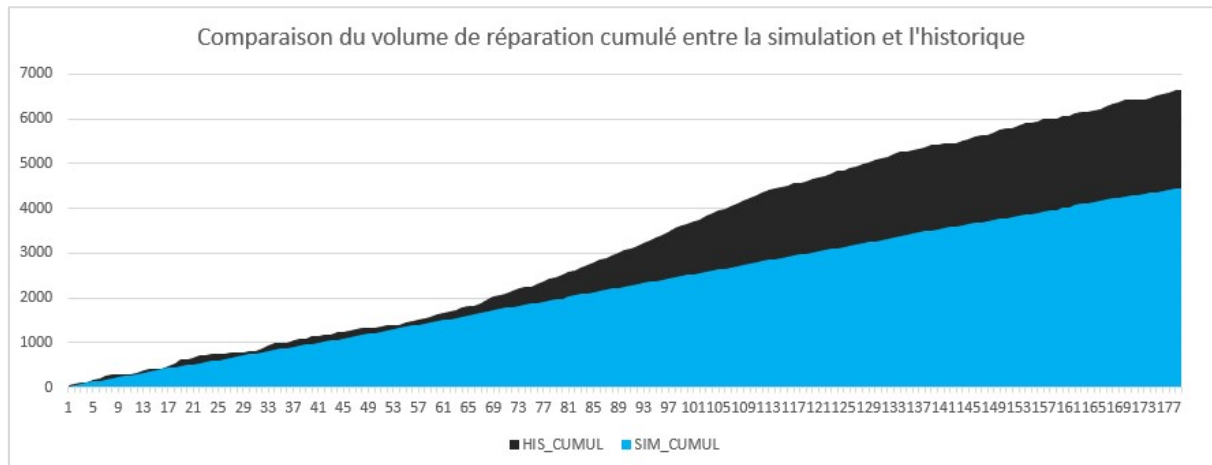


Figure IV.31: Visualisation de l'écart cumulé entre la vraie demande des pièces de rechange de GEHC et le résultat de la simulation

L'écart avec les vraies données devient plus important avec le temps et atteint plus de 2000 pièces que le modèle a prévu de ne pas envoyer en réparation mais qui le seront dans la réalité (Figure IV.31). Le modèle a tendance à sous-estimer l'envoi en réparation.

En analysant les données, nous trouvons plusieurs cas où l'algorithme théorique implique un achat de pièces neuves, mais, en réalité, ce sont des décisions de réparations qui ont été prises, même si elles sont plus longues.

Une des causes principales qui expliquent ces écarts entre ce que le modèle prévoit et la réalité est les prix qui sont très élevés d'une grande partie des pièces du catalogue de l'entreprise : pour cette catégorie de pièces, les responsables ne prennent pas de décisions systématiques suivant l'algorithme théorique mais traitent cas par cas. Par exemple, au lieu d'acheter une pièce neuve pour laquelle nous pouvons avoir une réparation à un dixième du coût, les planners prennent la décision d'attendre un peu plus quand il n'y a pas d'urgence, pour avoir une réparation au lieu de l'acheter beaucoup plus cher.

Le modèle de simulation ne permet pas un traitement des références d'une façon individuelle, l'algorithme s'applique d'une façon systématique quoi que ce soit la référence, ses caractéristiques ou les conditions du moment de la commande.

VII Conclusion

Dans ce chapitre, nous avons présenté une modélisation de l'algorithme de planification pour les pièces de rechange dans une chaîne logistique de pièces de rechange en boucle fermée. Ce modèle intègre les différentes solutions présentées dans ce contexte, qui sont la réparation et l'achat et intègre les différentes règles de gestion ainsi que les contraintes de notre contexte.

Ensuite, nous avons présenté l'implémentation du modèle de simulation et son application sur les vraies données du service de réparation de pièces de rechange de GEHC. Après l'analyse de la performance du modèle de simulation, nous constatons un ensemble de limites :

- Vu que le modèle se base sur des prévisions qu'il répartit d'une façon uniforme sur la période de simulation, le résultat de la simulation est quasi constant et ne représente pas la réalité des opérations.
- Le modèle de simulation ne permet pas de prévoir les pics de la demande et/ou l'augmentation ou la baisse de la charge de réparation, ce qui représente le besoin le plus important dans notre contexte.
- L'intégration de nouvelles contraintes pour améliorer la performance est assez complexe. En plus, il est très difficile ou impossible d'intégrer des règles permettant la prévision des pics ou les tendances de la charge dans les centres de réparation, vu qu'il faut se baser sur un grand nombre de variables, ce qui va générer un algorithme de grande taille qui est difficilement maintenable, et qui nécessitera des mises à jour manuelles à chaque changement dans les règles de gestions.

Quelques exemples de ces règles de gestion sont :

- **Les sources du BOH** : Dans notre modèle, la seule source est les pièces GOH qui se transforme en BOH après que les techniciens de maintenance les remplacent. En réalité, d'autres sources existent telles que les pièces récupérées des anciens systèmes que l'entreprise récupère des clients lors de mise à niveau de leurs systèmes. La donnée du nombre de systèmes récupérés dans le futur n'est pas disponible et son intégration est assez complexe. Ceci est une des raisons pour lesquelles le modèle génère moins de réparations que ce qu'on voit dans l'historique, car le modèle ne trouve pas de stock BOH pour effectuer les réparations, et compense pas du NB.
- **La création des commandes** : En réalité, et en se basant sur les données, les approvisionnements ne sont pas faits d'une façon ponctuelle comme le modèle théorique, parfois les planners passent une partie des commandes et pas tout le besoin théorique, en fonction du prix de la pièce et de la situation. Pour les pièces chères par exemple, les responsables d'approvisionnement tendent à être plus « prudents » et ne passent des commandes que lorsqu'ils s'assurent que le besoin existe.
- **Les différences entre les clients et les pays destinations** : Le modèle que nous avons implémenté ne traite pas les différences entre les clients, et considère que tous les clients acceptent les pièces neuves et réparées. Cependant, en réalité, ceci n'est pas vrai et dépend du client et du pays destination des pièces de

rechange, il existe des pays qui n'acceptent que les pièces neuves et non les pièces réparées. Pour les pays de ce type, le service de planification des pièces de rechange passe des commandes d'achat même si théoriquement une réparation fait plus de sens. L'intégration de cette contrainte est très complexe surtout que la liste des clients et des pays acceptant des pièces réparées évolue dans le temps.

- **Les réallocations de l'inventaire entre les différentes régions du monde :** Une des hypothèses de simplification que nous avons faites lors de l'implémentation est les différentes transactions logistiques. Dans notre modèle, nous considérons uniquement les transactions entre : (i) les entrepôts et les centres de réparation, (ii) les clients et les entrepôts et (iii) les fournisseurs et les entrepôts de stockage. En réalité d'autres types de transactions existent, telles que les ajustements inter-entrepôts et inter-pôles, qui sont également très difficiles ou impossibles à intégrer d'une façon précise dans notre modèle.

A cause de ces différentes limites, le modèle n'arrive pas à simuler d'une façon précise les décisions d'approvisionnement avec de la réparation, et vu la complexité élevée du travail nécessaire pour améliorer le modèle et intégrer toutes les contraintes et les combinaisons de décisions possibles, ainsi que le travail manuel obligatoire lors de n'importe quel changement dans la logique, l'intégration des règles de gestion dans l'algorithme n'est pas envisageable.

Amélioration de la performance du modèle

Pour prévoir des décisions complexes comme la décision d'approvisionnement qui se base sur un grand nombre de variables ; et au lieu de décrire l'algorithme de décision manuellement, un ensemble d'approches plus « intelligentes » sont présentes dans la littérature, qui sont les algorithmes d'apprentissage. Ce type d'approches permet de ne pas décrire analytiquement des algorithmes complexes, mais plutôt se baser sur des quantités importantes de données historique (expérience) pour répondre à des nouvelles situations.

Dans le chapitre suivant, nous allons présenter une implémentation d'un modèle d'apprentissage des décisions d'approvisionnement, notamment les approvisionnements par réparation. Nous allons voir une revue de littérature sur les modèles d'apprentissage, présenter notre méthodologie d'implémentation et enfin présenter les résultats du modèle sur les données historiques du service de réparation des pièces de rechange de GEHC.

Chapitre V : Modélisation et implémentation d'un model Machine Learning pour la prévision de la charge dans les centres de réparation dans une CL-PR-BF

I Introduction

Dans une chaîne logistique complexe et de grande échelle, le flux informationnel représente un élément important dans la stratégie de l'entreprise. Les nouvelles technologies permettent l'accès à un nombre massif de données à un coût économiquement acceptable. Ceci motive les entreprises à exploiter ces dernières pour améliorer leur niveau de service et de performance (Ellram *et al.* (2004), Van Jaarsveld *et al.* (2011), Baryannis *et al.* (2019))

L'exploitation des données massives inclut généralement des algorithmes et des modèles se partageant entre les sciences informatiques, mathématiques et statistiques. Le Machine Learning est une approche reposant sur de l'apprentissage statistique introduite dans la littérature dans les années 1950s, avec notamment des publications comme celle de Alan Turing « Computing machinery and intelligence-AM Turing » (Turing *et al.* (1950)). Cependant, la technologie ne permettait pas l'implémentation de ces modèles qui sont extrêmement demandeurs de ressources informatiques. De nos jours, la technologie permet l'implémentation de tels algorithmes et ouvre des pistes de recherche et de projets d'exploitation de données massives pour nombre d'industriels.

L'objectif de ce chapitre est de présenter l'apport de modèles d'apprentissage dans le contexte de la chaîne logistique de pièces de rechange pour la prévision de la charge dans les centres de réparation.

Le chapitre est décomposé de la manière suivante : (1) la présentation générale de modèles d'apprentissage, (2) la méthodologie d'implémentation des modèles Machine Learning et (3) l'évaluation de la performance sur des données réelles.

Cette étude est basée sur des données réelles opérationnelles provenant de la chaîne logistique après-vente de l'entreprise General Electric Healthcare.

II Notions de base en Machine Learning

1 Définition

Le Machine Learning ou ML est une catégorie de modèles d'intelligence artificielle apportant une capacité d'apprentissage aux ordinateurs sans nécessiter l'intervention humaine. Ce type de modèles a pour objectif de permettre l'implémentation de solutions informatiques capables de réagir à des nouvelles situations, autrement dit rendre les ordinateurs plus « intelligents » en accumulant de « l'expérience » des données qui leurs sont soumises.

Praveena *et al.* (2017) définit les méthodes de Machine Learning comme une sous-catégorie des domaines d'Intelligence Artificielle, où l'on cherche à apprendre des connaissances aux ordinateurs sans les programmer en intégralité, permettant ainsi de s'adapter à toute nouvelle situation tout en accumulant de la connaissance au fur et à mesure de la collecte de données.

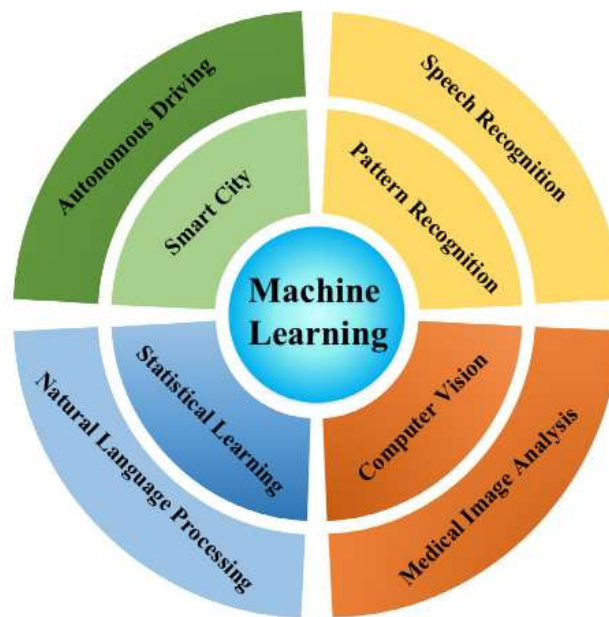


Figure V.1: Les applications principales de Machine Learning (Gao *et al.* (2020))

Le Machine Learning est actuellement l'un des domaines techniques en croissance rapide, se situant à l'intersection de la recherche informatique, les statistiques, l'intelligence artificielle et de la science des données (Baryannis *et al.* (2019)).

L'usage et l'enrichissement des approches d'apprentissage automatique attirent l'intérêt de nombreux chercheurs et la curiosité des entreprises. Il y trouve en effet un grand nombre d'applications réelles et variées, telles que de l'optimisation de la recherche sur le Web, le

filtrage de contenu sur les réseaux sociaux, les voitures autonomes et leur sécurité, et même les jeux vidéo et le divertissement.

2 L'apprentissage supervisé et non supervisé

La Figure V.2 présente les principaux types de modèles Machine Learning, qui sont l'apprentissage supervisé, l'apprentissage non-supervisé et d'autres types de modèles tels que l'apprentissage profond et les réseaux de neurones. Les différents types de modèles sont différents en termes de structure et de mode de fonctionnement.

En pratique, chaque catégorie de modèles est adaptée à un certain type de problèmes, ce qu'on voit dans la littérature (Gao *et al.* (2020)).

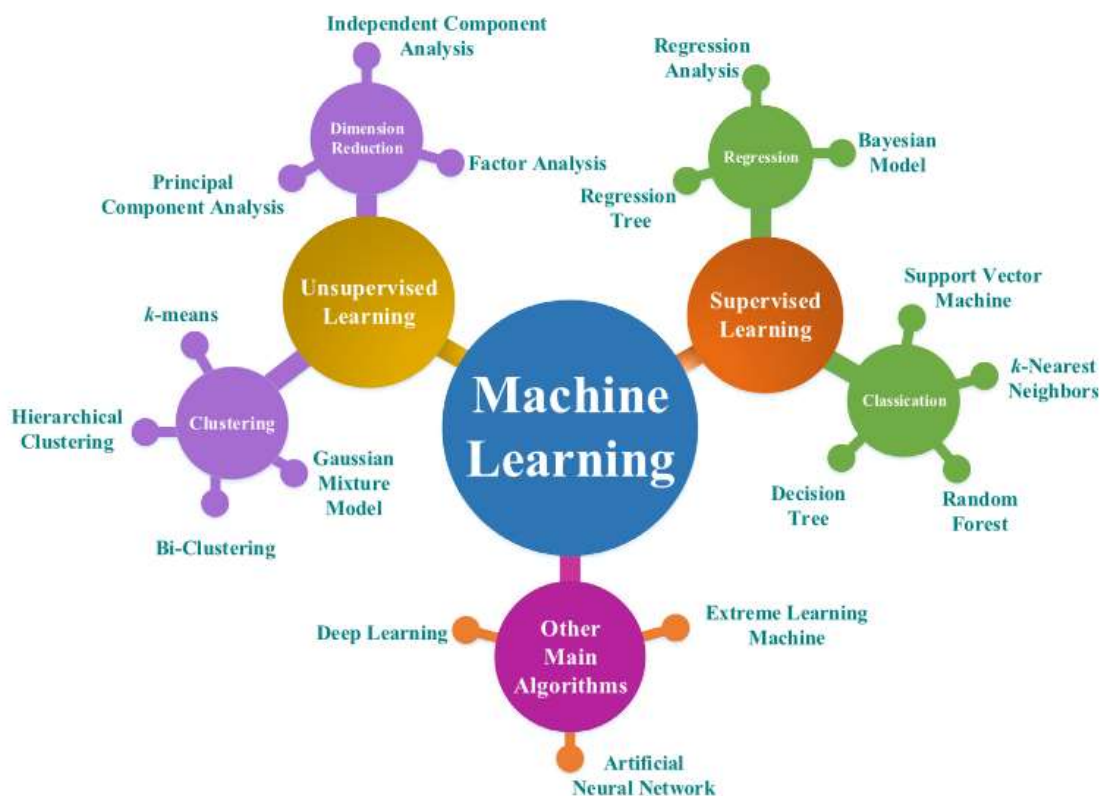


Figure V.2: Les principaux types de modèles Machine Learning (Gao *et al.* (2020))

2.1 L'apprentissage supervisé

Le premier type de modèles est lié à l'apprentissage supervisé. C'est appelé ainsi puisqu'on « supervise » l'apprentissage pendant la phase de calibration des modèles. De plus, les paramètres du modèle peuvent être dynamiquement ajustés en fonction de la problématique et du besoin. L'apprentissage supervisé consiste à estimer la fonction transfert entre les

observations et les données de sortie. Il existe deux catégories de modèles supervisés, comme l'indique la figure V.3 (Praveena *et al.* (2017), Sutton-Charani *et al.* (2014)).

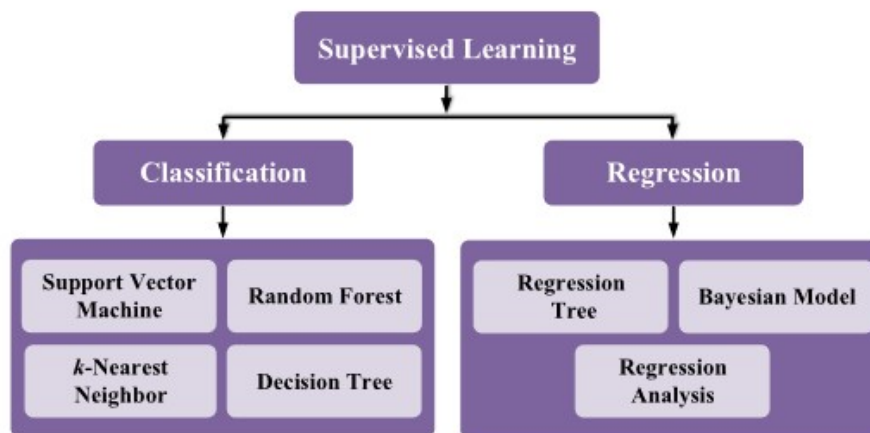


Figure V.3: Les différents modèles d'apprentissage supervisé (Gao *et al.* (2020))

L'ensemble de nos contributions se base sur l'apprentissage supervisé, que nous présenterons dans les sections suivantes.

2.2 L'apprentissage non-supervisé

L'apprentissage non supervisé ne s'appuie que sur l'analyse brute des données avec une catégorisation de ces données réalisées en autonomie par le programme, sans chercher à s'appuyer sur des éléments prédéfinis. Pour ce faire, le système va croiser les informations qui lui sont soumises, de manière à identifier des similitudes entre les données et pouvoir les rassembler en classe de similitudes. Ainsi, en fonction du but recherché, il reviendra à l'opérateur ou au chercheur de les analyser afin d'en déduire les différentes hypothèses. Il est sûr que la qualité des résultats fournis dépend à la fois de la quantité et de la pertinence des données.

La différence entre ces deux principes de fonctionnement réside dans le fait que l'apprentissage supervisé peut être influencé par des paramètres au moment de l'étiquetage des données. Par ailleurs, il faut noter que des travaux existent pour également faire intervenir un mode de fonctionnement mixte qui utilise les deux types d'apprentissage pour arriver à des résultats plus précis.

2.3 Autres méthodes

En plus des méthodes supervisées et non-supervisées, d'autres types de modèles existent, tels que les réseaux de neurones (Chen *et al.* (2017)), le Deep Learning (Goodfellow *et al.* (2016), LeCun *et al.* (2015)) et l'apprentissage extrême (Huang *et al.* (2006)).

3 La classification et la régression

Il existe deux principales catégories de modèles de Machine Learning supervisés, chacune choisie en fonction de la problématique traitée. Ces deux méthodes sont : (i) les modèles de classification et (ii) les modèles de régression.

Dans notre problématique, nous avons employé les deux types de modèles pour des parties différentes de la problématique.

4 La classification

4.1 Définition

Les modèles de Machine Learning pour la classification consistent à catégoriser une nouvelle observation parmi un ensemble de classes possible. La classe est définie a priori et repose sur un étiquetage des observations pour lesquelles on connaît les propriétés. On parle aussi d'observations libellées. Il revient alors d'affecter la nouvelle donnée à une classe existante (Figure V.4).

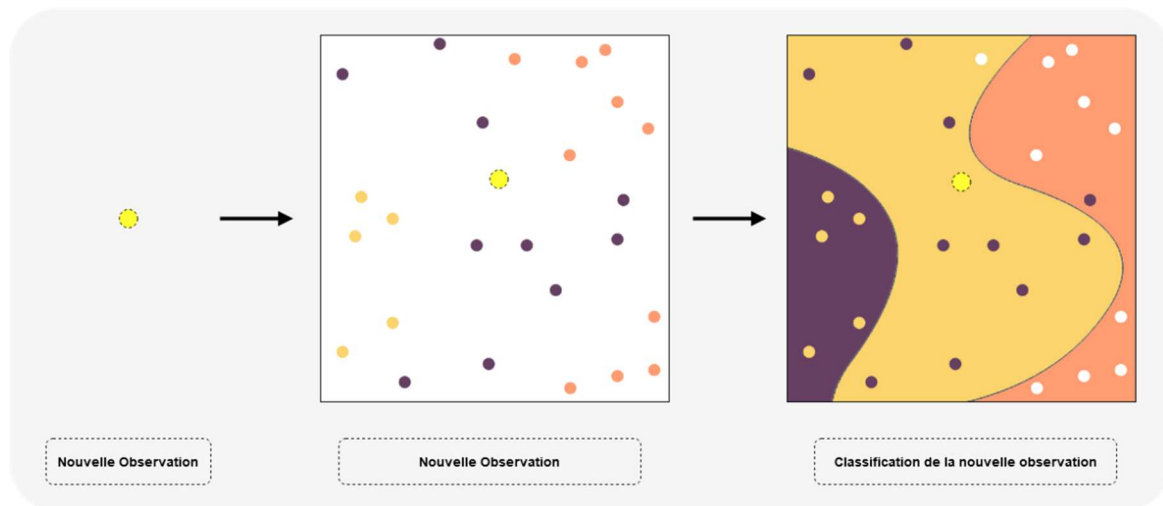


Figure V.4: Mode de fonctionnement des modèles de classification

Plusieurs types de classification existent :

- Le type le plus simple est la classification avec une seule classe ($E = \{C\}$). Ce type permet détecter les anomalies, en détectant les observations n'appartenant pas à la classe normale. Un des exemples d'utilisation de ce type de classification est la

détection des emails spam, où la seule classe existante (la classe « normale ») est la catégorie des emails non-spam.

- Le deuxième type, et qui est le plus connu, est la classification binaire où l'objectif est de catégoriser une nouvelle observation parmi deux classes. L'ensemble des classes disponibles est $E = \{N, P\}$, où N est la classe négative et P est la classe positive. Ce type d'observations à un grand nombre d'applications tels que l'authentification en utilisation la reconnaissance faciale.
- Le dernier type est la classification multi-classes (aussi appelée multi-étiquettes), où le nombre de classes possibles est supérieur à deux ($E = \{1, 2, \dots, k\}$). Une des utilisations classiques est la traduction automatique pour laquelle l'ordinateur doit détecter automatiquement la langue avec laquelle le texte est écrit avant de le traduire.

4.2 Evaluation de performance des modèles de classification

Les observations positives et négatives

L'évaluation de performance d'un système de classification est composée de quatre métriques principales (Figure V.5). On appelle des observations positives quand elles appartiennent à la même classe estimée par le modèle, et négatives dans le cas contraire (Sauvanaud *et al.* (2016)).

Les deux premières métriques évaluent le nombre d'observations positives, ce qui veut dire, celles pour lesquelles le modèle a prédit correctement leur appartenance à la classe concernée (TP : True Positif) ou celles que le modèle a classées comme non appartenant à la classe (TN : True Negative).

Les deux dernières métriques comptent le nombre d'observations dites « négatives » qui sont incorrectement classifiées, soit celles que le modèle a classées comme appartenant à la classe et qui ne le sont pas (FP : False Positive) ou celles qui ont été classées comme non appartenant mais qui le sont (FN False Negative).

Généralement on s'appuie sur le nombre total d'observations positives ou négatives pour évaluer la performance d'un modèle de classification, alors que les autres métriques sont plus utilisées pour l'analyse et l'amélioration de la performance d'un modèle.

	Observation positive	Observation négative	Total
Décision positive	Vrai positif <i>TP</i>	Faux positif <i>FP (I)</i>	<i>P'</i>
Décision négative	Faux négatif <i>FN</i>	Vrai négatif <i>TN (II)</i>	<i>N'</i>
Total	<i>P</i>	<i>N</i>	

Figure V.5: Grandeurs élémentaires pour l'évaluation des performances (Matrice de confusion) (Sauvanaud *et al.* (2016))

Probabilité de bonne classification

Selon Lecomte *et al.* (2013), la probabilité de bonne classification (P_{BC}) mesure la proportion des observations testées qui ont été correctement classées. Cette mesure simple permet d'évaluer très rapidement les performances d'un détecteur sur une base de test équilibré. On l'utilise en particulier lors des étapes de validation croisée afin d'estimer les paramètres d'un modèle. On détermine la probabilité de bonne classification par l'expression suivante :

$$P_{BC} = \frac{TP+TN}{P+N} \quad (\text{Lecomte } et al. (2013))$$

5 La régression

5.1 Définition

Les modèles de régression sont utilisés pour estimer une valeur réelle pour une observation donnée. Selon Roy *et al.* (2018), lorsque l'espace de sortie est une valeur réelle, on parle alors d'un modèle de régression.

Comme pour les modèles de classification, les modèles de régression se basent sur un nombre important d'observations avec leurs valeurs réelles associées. Ces données sont utilisées pour estimer une fonction transfert liant les observations avec leurs résultats.

L'avantage des modèles de régression est leur performance en termes de temps de calcul. Au contraire des modèles de classification pour lesquels on traite toutes les observations disponibles pour estimer une classe, les modèles de régression se calibrent entièrement pendant la phase d'apprentissage, puis, le paramétrage étant fixé, aucun post-traitement ne sera fait pendant la phase d'utilisation

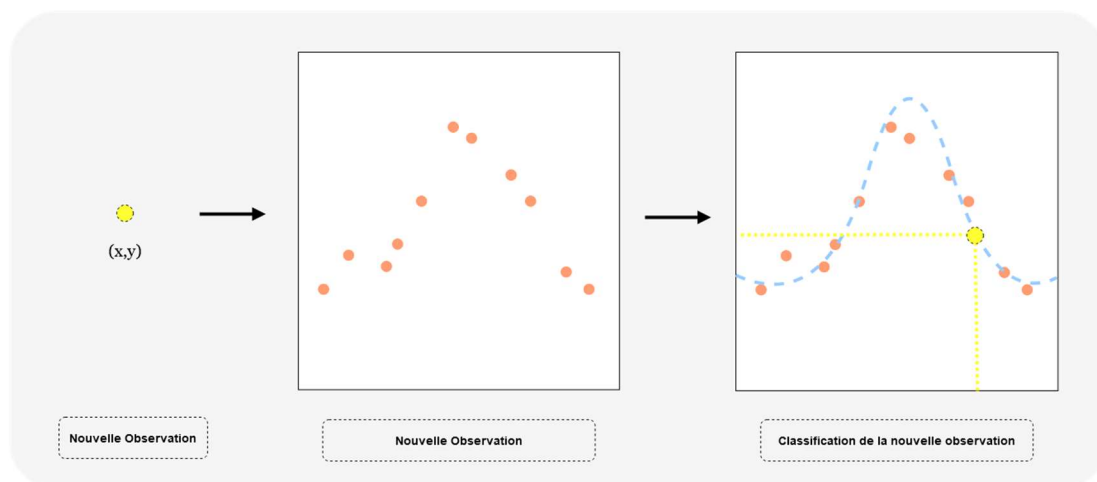


Figure V.6: Principe simplifié des modèles de régression

6 Le sur apprentissage et le sous apprentissage

Il ne suffit pas de mettre en place un modèle pour obtenir des résultats de qualité, surtout en termes de prédictibilité. Une phase préalable de paramétrage est primordiale pour éviter les erreurs dues au :

- Sous-apprentissage, qui est le fait de ne pas ajuster suffisamment le modèle aux données. Ainsi, le modèle apprend « mal » l'historique et ne peut donc pas non plus généraliser le phénomène en question sur de nouvelles données.
- Sur-apprentissage, qui consiste au contraire à trop ajuster le modèle aux données. Le modèle apprend même le bruit et cela a un impact négatif sur la prédiction face à de nouvelles observations dont le bruit sera aléatoirement différent.

Il est important de trouver un compromis « biais-variance » du modèle. Le biais d'un modèle correspond à la différence asymptotique entre la moyenne des prédictions et la moyenne réelle des données d'apprentissage. La variance du modèle quant à elle correspond au bruit présent dans le jeu de données qui modifie radicalement la prédiction d'un jeu de données à l'autre. Le sous-apprentissage contient un biais trop important tandis qu'une variance trop forte est le signe d'un sur-apprentissage. Toutes les méthodes algorithmiques sont concernées par ce compromis. Plus un modèle est complexe, plus il intègre de paramètres, et plus il est flexible donc capable de s'ajuster aux données engendrant ainsi une erreur faible d'ajustement. En revanche, un tel modèle peut s'avérer défaillant lorsqu'il s'agira de prévoir ou de généraliser, c'est-à-dire de s'appliquer à des données qui n'ont pas pris part à sa construction.

Des critères de sélection de modèles tels que l'AIC (Akaike Information Criterion) ou dans un contexte bayésien le BIC (Bayesian Information Criterion) permettent de venir qualifier un modèle en fonction de sa complexité, à savoir le nombre de paramètres. L'étude des performances de ces modèles ne fait pas partie de cette thèse qui a un objectif plus applicatif par rapport à la problématique industrielle. Nous invitons le lecteur à se référer par exemple à Lebarbier *et al.* (2004) pour une première analyse.

III Implémentation d'un modèle Machine Learning pour la prévision de la charge des centres de réparation dans une CL-PR-BF

1 Méthodologie d'implémentation de modèles d'apprentissage dans la littérature

Afin de mettre en place un système d'apprentissage pour répondre à un problème donné, l'application d'un ensemble d'étapes est nécessaire afin de s'assurer du bon fonctionnement du modèle. Cette section introduit une méthodologie d'implémentation de modèles de Machine Learning allant de la collecte de données à l'évaluation au choix du modèle optimal.

La méthodologie que nous présentons est extraite de la littérature. Elle a été déployée sur de nombreuses applications telles que l'extraction et la classification des émotions à partir d'un

texte, Gamal *et al.* (2019), ou encore la prévision de la consommation énergétique, Gallagher *et al.* (2018).

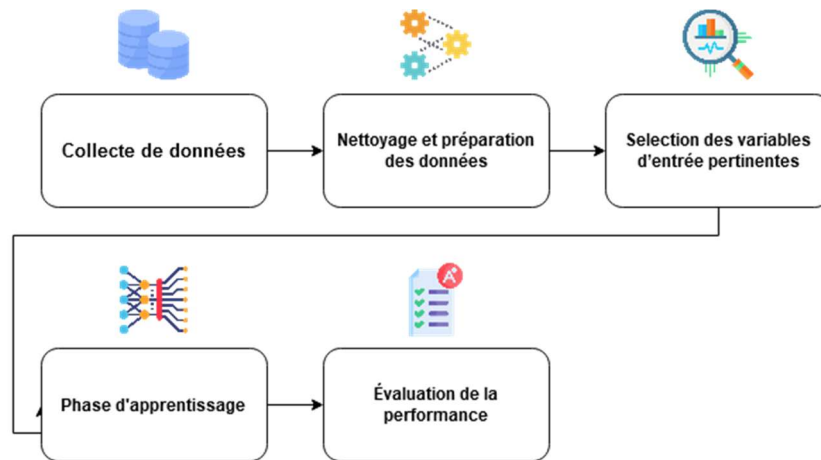


Figure V.7: Méthodologie d'implémentation de modèles d'apprentissage

Comme le montre la figure V.7, à haut niveau, nous proposons de suivre les cinq étapes suivantes :

- Collecte de données que nous allons utiliser pour l'apprentissage du modèle
- La préparation des données en effectuant un « nettoyage » et élimination du « bruit » des données
- Le choix des variables que nous allons employer pour l'apprentissage
- La phase de l'apprentissage
- L'évaluation des performances du modèle.

2 Implémentation du modèle de résolution

La méthodologie de base présentée dans la section précédente offre un cadre général d'implémentation des modèles d'apprentissage. Nous présentons ici (figure V.8) sa déclinaison générale sur le problème spécifique de la gestion des données dans le contexte de la chaîne logistique de pièces de rechange.

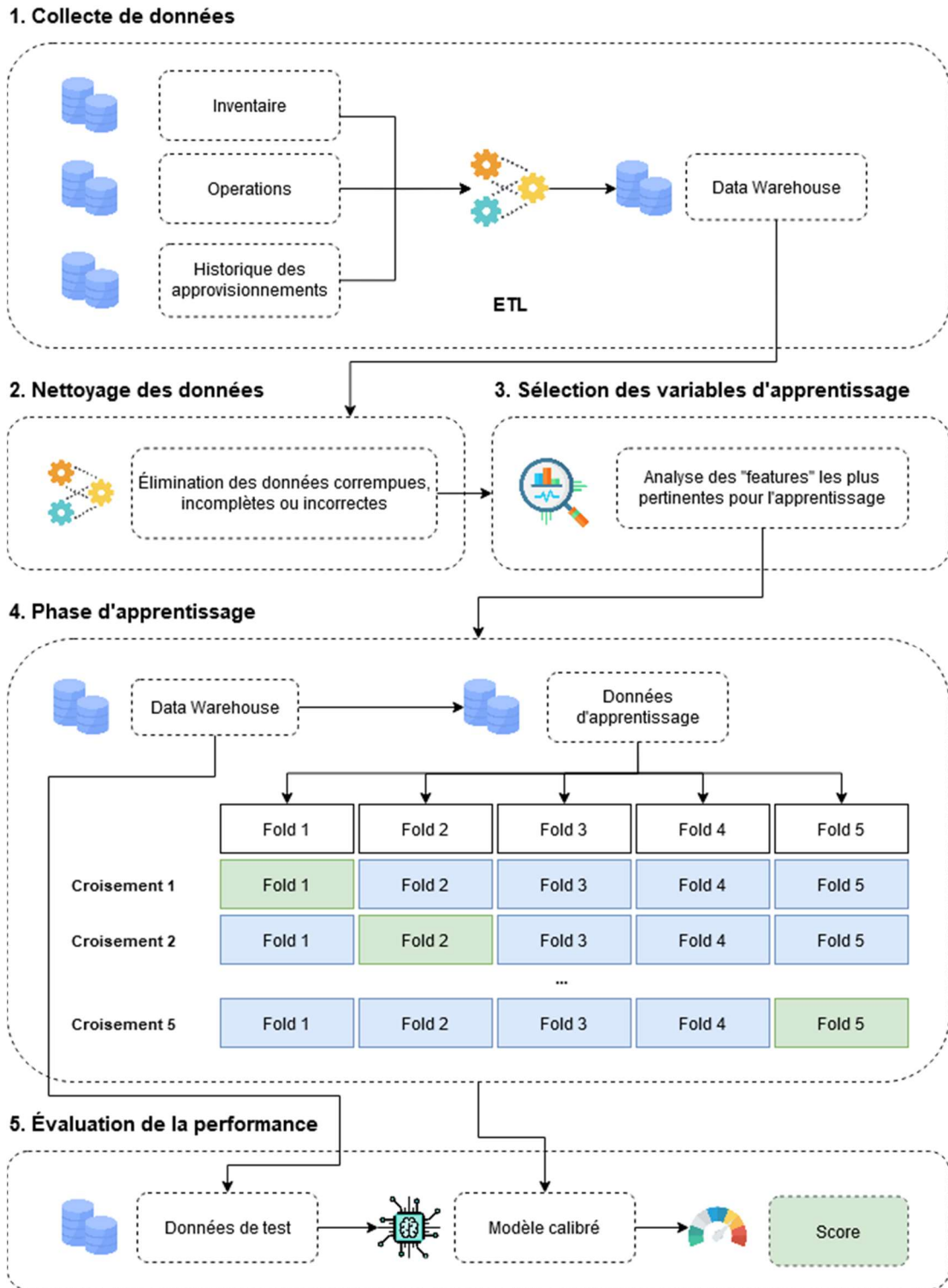


Figure V.8: Application de la méthodologie d'implémentation des modèles d'apprentissage dans notre contexte

2.1 La collecte de données

Comme nous l'avons vu dans les sections précédentes, un nombre important d'observations est nécessaire pour calibrer un modèle Machine Learning. Ces observations doivent contenir les

différentes variables de décision permettant d'estimer la variable de sortie, qui est le volume de réparation. L'un des problèmes que nous avons pu rencontrer dans cette thèse est la collecte des données quantitatives de masse. General Electric Healthcare a une longue expérience dans la gestion de ses systèmes et de ses clients. Cependant, l'acquisition de l'ensemble des données reste délicate au vu de la complexité de l'organisation GEHC, de la multiplicité et la diversité des systèmes d'information et l'accessibilité voire la propriété même de ces données pouvant être chez le client. Pour ces raisons, nous avons préféré une approche supervisée à l'approche non supervisée en définissant des classes et critères de décision sur la base d'avis d'experts.

Ainsi, nous nous sommes appuyés sur une dizaine de responsables de planification pour la définition des critères de décision. Nous avons déduit que les variables les plus utilisées et que nous devons récupérer pour la phase d'apprentissage sont : (i) l'historique de l'état d'inventaire, (ii) l'historique des opérations de réparation (iii) les historiques des décisions d'approvisionnement. Comme le montre la figure V.8, nous allons récupérer ces données de plusieurs parties du système d'information de GEHC.

Nos données proviennent de plusieurs sources avec des différents types de données (APIs, Bases de données relationnelles, Fichiers de données etc..). Après l'extraction des données, un travail de traitement et de transformation a été effectué pour les mettre dans une structure adéquate, afin de les charger dans un « datawarehouse » commun avec une structure relationnelle, permettant de faire le lien entre les différentes entités d'une façon automatique. Ce processus est appelé ETL (Extraction, Transformation, Load) et il est fréquemment utilisé dans la littérature pour le même objectif (Lecomte *et al.* (2013)).

D'après des sessions de travail avec les experts d'approvisionnement, les « planners » se basent sur 5 catégories de données (parmi plusieurs d'autres comme les données financières ou de facturation). Des analyses de type ACP (Analyse de composantes principales) existent pour choisir les données les plus critiques, cependant, dans notre cas, et pour des raisons de complexité et de multitude de données et de systèmes d'informations, nous nous sommes basés sur l'expérience des experts pour choisir les catégories :

- Etat de l'inventaire pour chaque référence, dans les différents pôles de l'entreprise
- Prévision de la demande des pièces de rechange
- Détails de chaque référence (prix neuve, cout de réparation, Age, garantie...)
- Détails des fournisseurs, délais d'approvisionnement pour chaque référence
- Historique des commandes d'approvisionnement

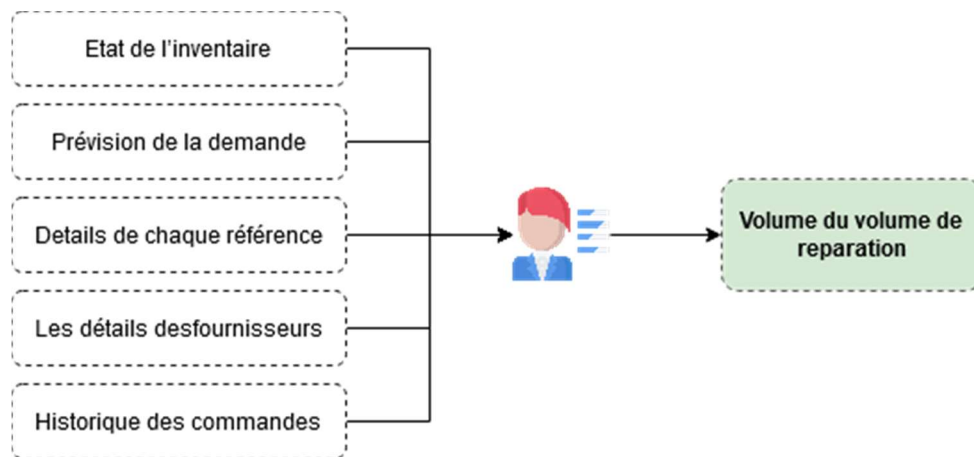


Figure V.9: Les processus de décision d’approvisionnement en se basant sur les données

Chaque « planner » est responsable sur un ensemble de références (généralement entre 100 et 200 références). Chaque jour et pour chacune de ces références, le responsable doit suivre l’état d’inventaire dans les différents entrepôts de l’entreprise, la demande des techniciens de maintenance, l’évolution des prix... afin de décider s’il y a besoin de faire des approvisionnements. Dans le cas de besoin, le planner doit décider quelles solutions sont les plus appropriées, en se basant sur les délais de livraison et sur les coûts (Figure V.9).

L’objectif du modèle de Machine Learning est de prévoir les décisions de planification en intégrant les données de réparation sur la base des mêmes informations disponibles aux planners, à savoir les données factuelles d’entrée et les règles de décision d’approvisionnement.

Ainsi, on appelle ligne de données, un ordre d’opération formé de ~120 données et représentée par une ligne du tableau suivant :

P/N	...	GOH	BOH	NBLT	RLT	NB Price	R Price	Frcst12M	Sol	Vol
___523-R	...	46	154	90	30	\$1389.00	\$798.33	221	NB	15
___318-R	...	17	32	60	20	\$2839.00	\$1411.50	175	NB	53

Tableau V.1: Exemple de lignes des données d’apprentissage

2.2 Le nettoyage de données

Le nettoyage de données est une étape primordiale dans la mise en place d’un modèle d’apprentissage. Ceci est dû au fait que la performance du modèle dépend principalement de la qualité des données d’entrée. Les données d’entrée doivent avoir une bonne et complète couverture du phénomène étudié, qui est les décisions d’approvisionnement dans notre cas.

Dans le cas où d’une faible représentativité des cas contenus dans les données d’apprentissage, le modèle ne pourra simplement pas décider ou prévoir d’une façon fiable la variable de sortie.

Un autre avantage quant à cette phase de nettoyage des données est la simplification des modèles de traitement qui en suivront. En effet, on peut remarquer la bonne performance de modèles de classification et de régression dits simples sur la base de données de « bonne qualité » (Gudivada *et al.* (2017)). Ceci a clairement l'avantage de gains de temps pour les phases d'apprentissage et de traitements de données de modèles complexes.

2.2.1 L'élimination des données non désirées

Chaque échantillon de données peut contenir des observations non désirées, qui peuvent être soit des doublons ou des observations non-représentatives du phénomène étudié.

Dans notre cas, des doublons sont présents puisqu'on combine plusieurs sources de données et c'est crucial de les éliminer, sinon des calculs comme celui des volumes des pièces de rechange (défectives ou en état de fonctionnement) sera en double.

Pour les observations non-représentatives de notre problématique, elles considèrent les cas de décision d'approvisionnement qui sont présents dans l'historique mais qui sont soit des exceptions soit des erreurs. Par exemple, le tableau V-2 présente un ensemble d'observations qui sont considérées des exceptions à éliminer.

P/N	...	GOH	BOH	NBLT	RLT	NB Price	R Price	Frcst12M	Sol	Vol
__523-R	...	46	154	90	30	\$1389.00	\$798.33	221	NB	15
__318-R	...	17	32	60	20	\$2839.00	\$1411.50	175	NB	53
__748-R	...	84	45	90	15	\$2690.00	\$837.65	183	NB	6
__910-R	...	65	115	180	20	\$1720.00	\$967.00	253	NB	11

Tableau V.2: Exemple de données non représentatives de la problématique

Dans les exemples ci-dessus (Tableau V.2), nous retrouvons des observations pour lesquelles les responsables d'approvisionnement ont choisi de passer par l'achat de pièces neuves même si le service dispose de pièces réparables pouvant être remises en état de fonctionnement plus rapidement et à moindre coût.

Après analyse, on peut trouver plusieurs raisons pour ces décisions comme la politique d'import-export de certains pays n'acceptant que des pièces neuves (Notons ici que ceci restant en marge de nos cas, ce critère n'a pas été retenu). Cependant, dans un ensemble de cas, il n'y a pas d'explication et d'après les experts eux-mêmes, ces observations peuvent être affectées à des erreurs dans les données que nous devons éliminer pour éviter d'avoir du bruit dans notre modèle.

De la même façon, nous avons éliminé un ensemble de cas non-représentatifs. Le tableau V.3 montre ainsi le volume des éléments à traiter et les résultats retenus après traitement. Notons que près de 10% des données a été ici considéré comme erroné (ceci peut être considéré comme un faible pourcentage au vu de certaines autres applications).

Echantillon de base	Echantillon après élimination des doublons	Après élimination des exceptions et erreurs
+14 000 000 obs	+12 000 000 obs	+11 000 000 obs

Tableau V.3 : Nombre des observations dans l'échantillon de test avant et après les opérations de nettoyage

2.2.2 Mise en forme des données

La mise en forme des données consiste à effectuer des transformations sur un échantillon de base afin de structurer en couples (observation, résultat). Pour notre cas, notre échantillon de données contient 46 variables qui représentent le « vecteur » observation, donnant un résultat qui est le volume de réparation demandé.

Notre échantillon de données (Tableau V.4) contient, pour chaque référence et pour chaque semaine fiscale, l'état de l'inventaire, les pièces en transit, le volume en réparation, les prévisions de la demande, et toutes les autres informations dont nous disposons, accompagnées par les décisions prises pour chaque semaine. En total, nous avons un échantillon de 11 millions d'observations couvrant chaque semaine fiscale de 2017 à 2020.

Vecteur d'observation										Résultat
YEAR	FW	Reference	...	GOH	BOH	SS	MinOQ	EOQ	...	Vol
2017	01	___523-R	...	46	154
2017	02	___523-R	...	17	32
...	...	___523-R	...	84	45
2020	05	___523-R	...	65	115

Tableau V.4: structure des données d'apprentissage

La visualisation des données et l'analyse des observations est une étape importante avant de passer à l'implémentation et à la calibration du modèle. Par exemple, cette étape permet de détecter des valeurs incorrectement formatées. Ceci est particulièrement pertinent pour les

variables catégoriques, par exemple, pour le type de commandes d'approvisionnement qui est une variable catégorique pouvant signifier un achat du neuf ou bien une réparation.

Si on visualise les valeurs de cette colonne pour une partie de l'échantillon (Figure V.10), on remarque qu'il y a plusieurs catégories qui doivent être regroupées : « repair » veut dire la même chose que « Repair » et « REAIR », et de même pour « Newbuy », « new_buy » et « NEW_BUY ». Ceci est un simple exemple pour montrer l'intérêt de cette étape, en réalité, des dizaines d'autres opérations sont effectuées pour le même objectif.

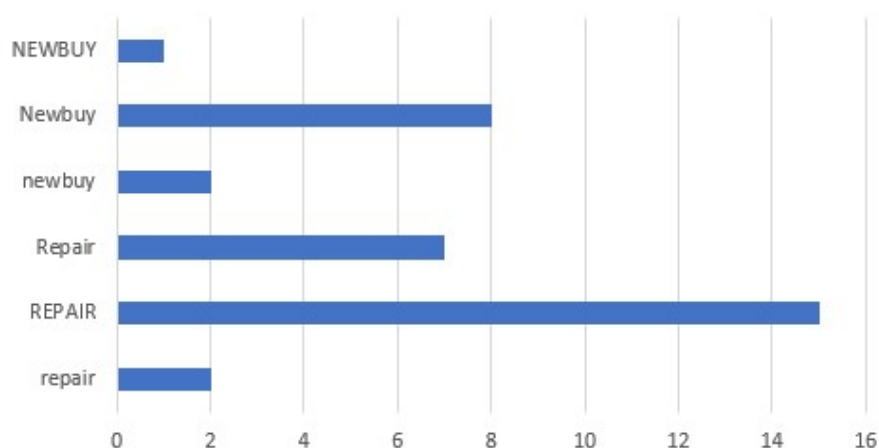


Figure V.10: Volume en fonction de la solution d'approvisionnement avant la mise en forme

Les figures V.10 et V.11 montrent la visualisation d'une des colonnes sur laquelle nous avons effectué ce processus, avant et après la transformation.

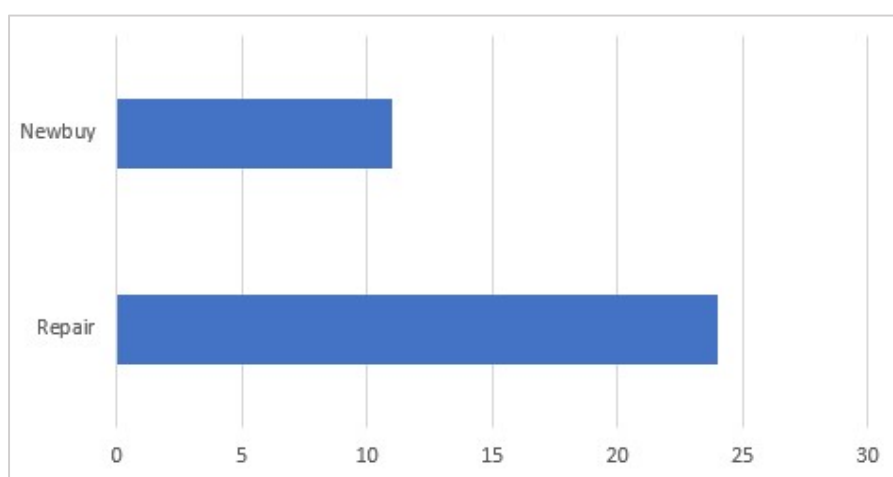


Figure V.11: Volume en fonction de la solution d'approvisionnement après la mise en forme

2.2.3 Le filtrage des données aberrantes

Pendant la préparation des données, nous avons remarqué la présence de données dites aberrantes dans le sens où elles présentent des caractéristiques très éloignées du nuage des

points. Après des sessions de travail avec les experts, ce phénomène s'explique par deux raisons : (i) des erreurs dans les données (ii) des cas isolés qui ne représentent pas les règles de gestion.

Chaque observation peut être aberrante pour un ou plusieurs critères. Pour notre cas, nous allons éliminer chaque observation aberrante pour au moins un critère.

Il existe plusieurs méthodes pour détecter des points aberrants. Nous avons utilisé la « boîte à moustache », qui est largement utilisée en littérature (Le Guen *et al.* (2002)).

Selon Le Guen *et al.* (2002), la boîte à moustache utilise 5 valeurs qui résument des données : le minimum, les 3 quartiles Q1, Q2 (médiane), Q3, et le maximum (Figure V.12).

Selon la même référence (Le Guen *et al.* (2002)), la boîte a pour hauteur la distance interquartile ($Q3-Q1$), et les moustaches sont basées généralement sur 1,5 fois la hauteur de la boîte. Une valeur est atypique si elle dépasse de 1.5 fois l'écart interquartile au-dessous du 1er quartile ou au-dessus du 3ème quartile. La valeur 1.5 est selon TUKEY *et al.* (1977) « une valeur pragmatique (« rule of thumb »), qui a une raison probabiliste ».

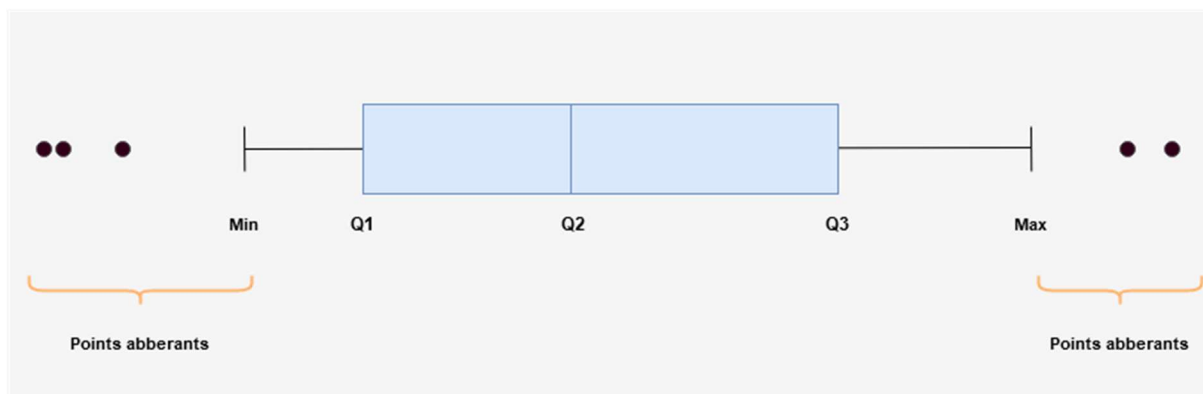


Figure V.12: La « boîte à moustache » et les points aberrants

2.2.4 La gestion des données manquantes

Dans cette partie de la préparation des données, il est impératif de ne pas laisser des données manquantes dans les données d'apprentissage, vu qu'une donnée vide peut être, dans certains cas, interprétée comme une valeur à part entière (0 ou null, dans la majorité des cas). Par ailleurs, précisons que nombre de bibliothèques informatiques d'implémentation des modèles ne supportent pas des valeurs vides et vont générer des erreurs en cas de présence de ces valeurs manquantes.

Deux façons existent pour gérer les valeurs vides :

- **La complétion** : Remplir la variable manquante à partir des données, soit en la copiant directement, ou en la calculant. Dans notre cas, nous nous sommes limités à copier les données de d'autres sources quand c'est possible.

- **L'élimination** : Consiste à supprimer les observations avec des variables manquantes.

2.3 Apprentissage du modèle

2.3.1 Les algorithmes utilisés pour la prévision de la charge dans les centres de réparation

- **La régression polynomiale (Modèle linéaire généralisé)**

Selon Chesneau *et al.* (2017), un modèle de régression polynomiale est un modèle non-linéaire, visant à exprimer une variable quantitative Y à partir de p variables X_1, X_2, \dots, X_p , tel que :

$$Y = f(X, \beta_0, \beta_1, \dots, \beta_p) + \varepsilon$$

Où $f(X, \beta_0, \beta_1, \dots, \beta_p)$ étant un polynôme en X_0, X_2, \dots, X_p de coefficients $\beta_0, \beta_1, \dots, \beta_p$.

$$f(X, \beta_0, \dots, \beta_p) = \beta_0 + \beta_1 X + \dots + \beta_p X^p$$

Le modèle a un seul paramètre à définir lors de la phase d'apprentissage, qui est le degré p du polynôme. La calibration du modèle consiste à trouver les paramètres β_i minimisant l'écart ε entre les prévisions du modèle et les vraies données. Le problème est alors d'estimer f par un polynôme en choisissant le "bon" degré p parmi un ensemble d'entiers possible $\{1, \dots, N\}$ où N est le degré maximal qu'on s'autorise.

Avec un degré p élevé, le modèle devient plus « flexible », capable de s'adapter plus aux tendances dans les données. Cependant, un polynôme avec un degré trop élevé va être susceptible de sur-apprentissage et donc ne pas suivre la tendance générale des données, mais aussi les bruits. En même temps, un degré p trop faible peut engendrer un sous-apprentissage (Figure V.13).

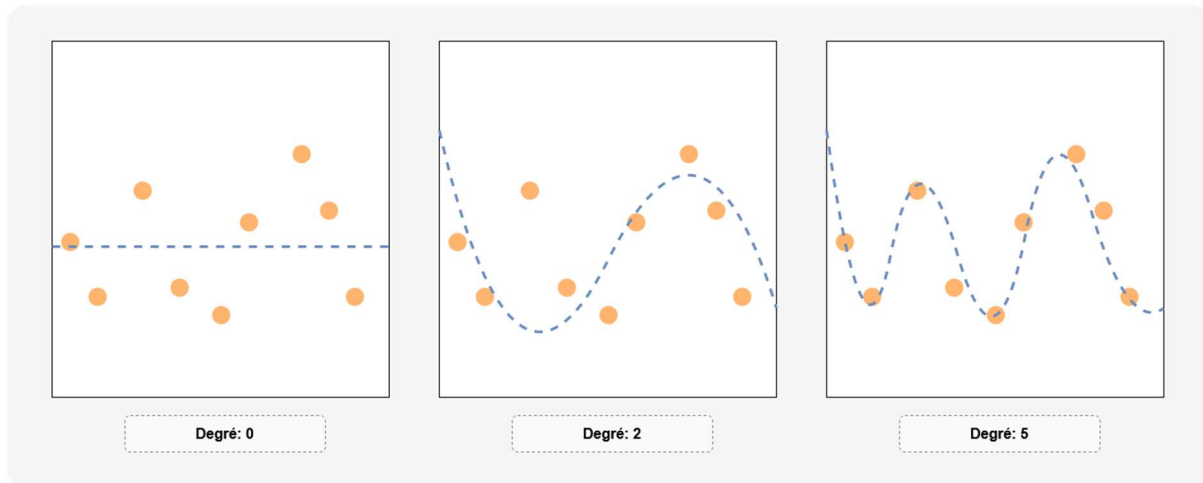


Figure V.13: Résultat d'une régression polynomiale avec plusieurs degrés : 1, 2 et 5.

Il est recommandé de commencer l'apprentissage avec un degré faible, puis de remonter d'une façon progressive afin de voir la tendance la performance en fonction de la valeur du degré du polynôme. Ceci est important puisque la complexité augmente d'une façon considérable en augmentant le degré du polynôme, et la phase de résolution prend un temps considérable quand le degré du polynôme est élevé.

- **La régression Ridge**

La régression Ridge est basée sur la régression linéaire multiple, en y ajoutant une contrainte sur les coefficients lors de la modélisation pour maîtriser l'amplitude de leurs valeurs. Ce type de méthodes est appelé : méthodes de contraction de coefficients. Pour une régression linéaire multiple, la modélisation prédictive d'une variable cible quantitative est donnée par l'équation suivante.

$$y = a_0 + a_1x_1 + \dots + a_nx_n + \varepsilon$$

Disposant d'un échantillon d'observations de taille n , nous cherchons les paramètres a_0, a_1, \dots, a_p qui minimisent l'erreur entre les observations et les estimations, par exemple l'erreur quadratique moyenne (MSE : mean squared error).

$$\min_{a_1, a_2, \dots, a_p} \sum_{i=1}^n \left(y_i - \sum_{j=1}^p a_j x_j \right)^2$$

Cependant, cette méthodologie peut générer des coefficients « aberrants » créant du bruit dans la prévision. En introduisant des contraintes sur les coefficients, on maîtrise l'amplitude de

leurs valeurs. La méthode Ridge propose une pénalisation avec un paramètre λ sur les coefficients (formule ci-dessous). Ce paramètre est ajustable pendant la phase d'apprentissage.

$$\min_{a_1, a_2, \dots, a_p} \sum_1^n \left(y_i - \sum_{j=1}^p a_j x_j \right)^2 + \lambda \sum_{j=1}^p a_j^2$$

○ La régression LASSO

La méthode LASSO (acronyme pour « Least Absolute Shrinkage and Selection Operator ») est une méthode introduite par Tibshirani *et al.* (1996). Comme pour la méthode Ridge, la méthode LASSO est une méthode de contraction de coefficients qui se base sur une approche de pénalisation. Cependant, LASSO propose une pénalisation de la norme L1 (valeur absolue) des coefficients.

Cette différence entre les deux méthodes fait que certains coefficients peuvent avoir une valeur nulle (zéro), au contraire de Ridge où les coefficients peuvent être négatifs (voir formule ci-dessous).

$$\min_{a_1, a_2, \dots, a_p} \sum_1^n \left(y_i - \sum_{j=1}^p a_j x_j \right)^2 + \lambda \sum_{j=1}^p |a_j|$$

2.4 Evaluation de la performance des modèles de régression

Pour générer nos prévisions, nous devons tester les différents modèles que nous avons présentés, puis choisir celui qui offre la meilleure précision. La calibration et l'application de chacun des modèles vont être par référence. Ceci va nécessiter un temps important de calibration vu le grand nombre de références. Cependant, ceci est nécessaire dû aux différences importantes entre les diverses références de pièces de rechange :

- **L'âge de la référence** : joue un rôle important dans les décisions de gestion d'inventaire, pour une pièce relativement ancienne le taux de réparation baisse et doit être pris en compte. Les décisions de réparation de deux références avec des niveaux d'inventaires identiques et une demande identique peuvent varier en fonction de l'âge des pièces.
- **La valeur des pièces** : Les prix des pièces de rechange peuvent varier d'un montant de 100 dollars à un montant de 80.000 dollars américains ou plus. En fonction de la valeur de chaque référence, les décisions d'approvisionnement peuvent varier. Comme vu dans le chapitre précédent, un ensemble de cas sont présents dans les données où les

planners prennent plus de temps pour passer des commandes, pour des pièces de haute valeur.

- **La taille de la base installée pour chaque référence** : dépend du cycle de vie du produit sur lequel la référence est employée. Le cycle de vie du produit et son impact sur la demande de pièces de rechange sont illustrés dans la figure V.15. Dans la phase initiale, la taille de la base installée augmente avec l'augmentation des ventes de produits. La demande de pièces de rechange est plutôt faible, car les produits sont encore relativement « jeunes ». Dans la phase de maturité où les ventes de produits reculent progressivement et la base installée atteint sa taille maximale, la demande de pièces détachées augmente. Dans la phase de fin de vie (End Of Life ou EOL), les ventes de produits diminuent et éventuellement s'arrêtent. La demande de pièces de rechange peut continuer d'augmenter, avant de diminuer progressivement à mesure que davantage de produits atteignent leur fin d'utilisation (Kim *et al.* (2017)). D'après ce chiffre, on peut voir que la demande de pièces de rechange suit la demande du produit, mais avec un certain retard.

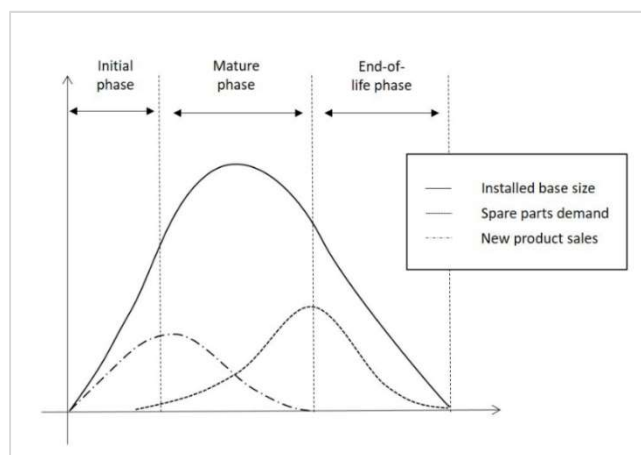


Figure V.14: Cycle de vie du produit, évolution de la base installée et demande de pièces de rechange (Dekker *et al.* (2013))

- **Les régions où les références sont demandées le plus** : ce critère joue un rôle important dans le choix des solutions d'approvisionnement, vu qu'un ensemble de pays impose un certain nombre de règles concernant les opérations d'import / export des pièces de rechange. Par exemple, un nombre de pays n'accepte pas l'importation des pièces de rechange réparées, mais uniquement des pièces neuves. Pour ce cas, même si on dispose d'un stock BOH qu'on peut réparer plus rapidement et moins cher, nous allons choisir d'acheter des pièces neuves.

Pour ces différentes raisons, nous choisissons d'appliquer à chaque référence, un modèle qui lui est spécifiquement calibré.

L'algorithme 7 présente la logique que nous avons appliquée pour calibrer les modèles. On commence par l'extraction de l'ensemble des références des pièces ainsi que l'extraction des modèles que nous avons implémentés.

Ensuite nous parcourons toutes les pièces et, pour chacune, nous calibrons le modèle sur les données de celle-ci, puis trions le résultat de tous les modèles en fonction de la métrique R2, et enfin choisir le meilleur en termes de « score » le plus élevé et l'attribuer à la référence courante.

Algorithme 7 : Calibration des modèles sur les différentes références de pièces

```
1.      Fonction Calibrer_modeles()
2.          modeles[] ← getListeModeles()
3.          pieces[] ← getParts()
4.          Pour chaque m dans modeles Faire:
5.              temp[] ← []
6.              Pour chaque p dans pieces Faire :
7.                  temp[] ← m.calibrer(p)
8.              Fin de Pour
9.              temp[] ← temp.calibrer(critere = 'R2', ordre = 'Asc')
10.             m.modele ← temp[0]
11.          Fin de Pour
12.      Fin
```

Algorithme 7 : Fonction de calibration des modèles implémentés sur le catalogue des pièces de rechange de GEHC

L'algorithme 7 commence par récupérer la liste des pièces disponibles, puis, pour chacune, on calibre chacun modèle et on choisit celui qui donne la meilleure précision. Le « meilleur » modèle est affecté à la pièce et utilisé pour générer ses prévisions. Ceci permet d'utiliser le modèle approprié pour chaque pièce.

2.4.1 Les données de test : Les 500 références les plus demandées de GEHC

Pour calculer la performance des différents modèles présentés dans la section précédente, nous choisissons un échantillon de 500 références de pièces de GE Healthcare. La raison de ce choix est que ces 500 références représentent 80% des commandes passées depuis janvier 2017 (Figure V.16) mais qui représentent uniquement ~25% du nombre total des références. Cette restriction permet de réduire énormément le temps de calcul, tout en gardant une bonne couverture des commandes des pièces de rechange.

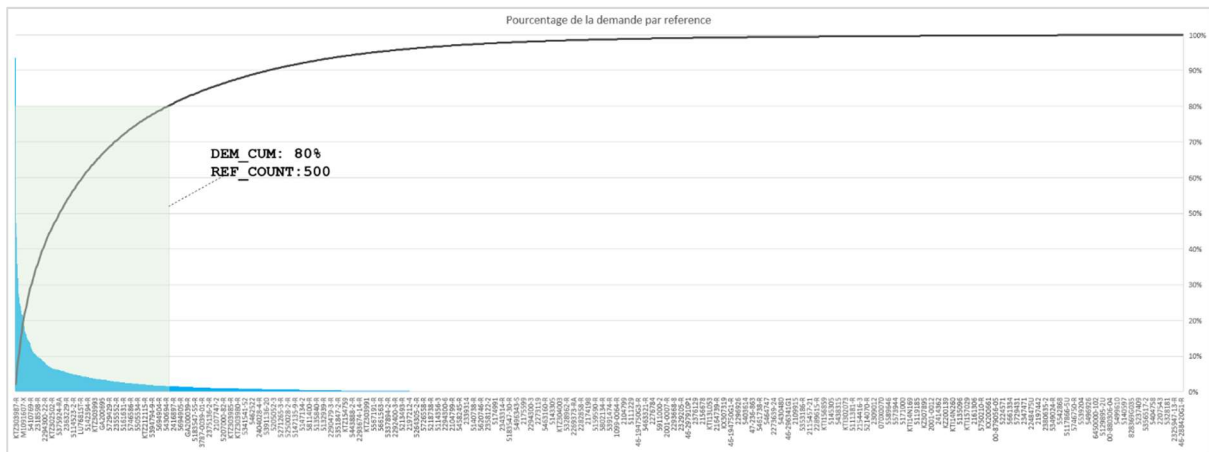


Figure V.15: Présentation des 500 références représentant 80% de la demande globale des pièces de rechange de GEHC

2.4.2 Temps de calibration des modèles

Le temps nécessaire pour la calibration d'un système de prévision est un élément important pour le choix des modèles qui y sont utilisés. La durée de calibration de modèles d'apprentissage peut varier de quelques secondes à plusieurs jours voir plus, en fonction du type du modèle, le nombre de ses paramètres et de la quantité de données à prendre en considération.

La figure V.16 présente les durées nécessaires pour la calibration de chacun des modèles sur l'échantillon des 500 références de pièces de rechange (en minutes). Le modèle de régression polynomiale est celui qui prend en moyenne le plus de temps, suivi du modèle Lasso et puis la régression Ridge.

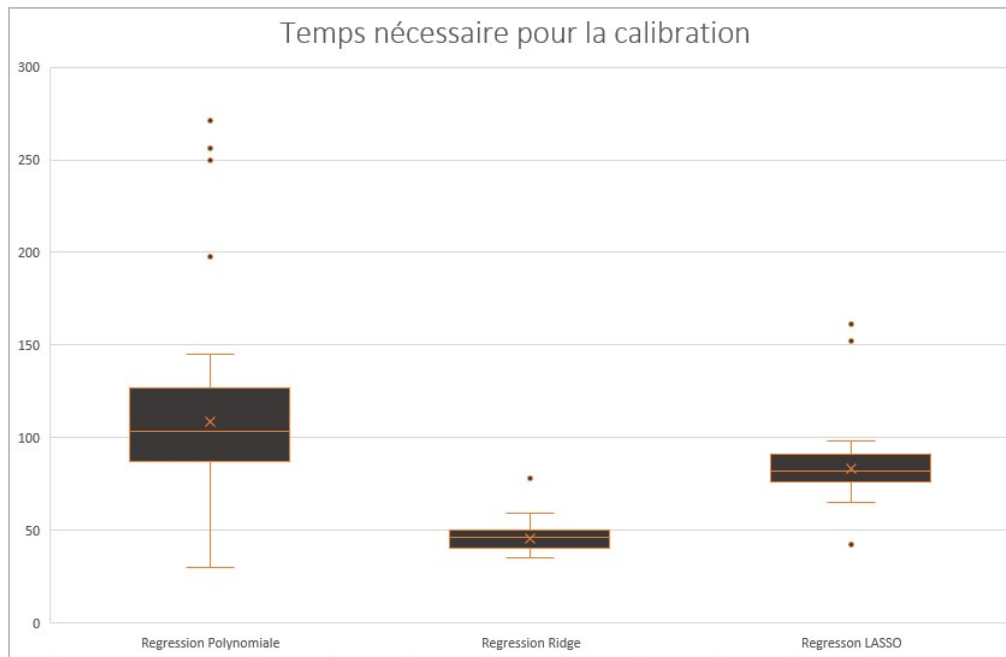


Figure V.16: Distribution des temps de chargement de chaque modèle sur l'échantillon des 500 références

2.4.3 Evaluation de la performance des modèles

Nous commençons d'abord par les résultats spécifiques pour chaque modèle. Nous allons présenter la performance de chacun des algorithmes quand on les applique séparément à toutes les références de pièces.

2.4.3.a Le modèle de régression polynomiale

La figure V.17 présente la prévision de la globalité des commandes de réparation générées par le modèle de régression polynomiale, quand il est appliqué à toutes les références.

Au contraire du modèle de simulation où le volume total est quasi-constant, la régression polynomiale génère des volumes variables qui augmentent ou diminuent en fonction de sa calibration. Dans nos résultats, le modèle ne suit pas la tendance des données réelles et s'éloigne des vraies données avec le temps.

Notons que, pour une meilleure représentation de notre problématique, nous avons volontairement choisi des données contenant des pics de charge et une chute à la fin de notre horizon. Le modèle n'a pas permis de détecter ces changements brusques avec toutes les configurations possibles.

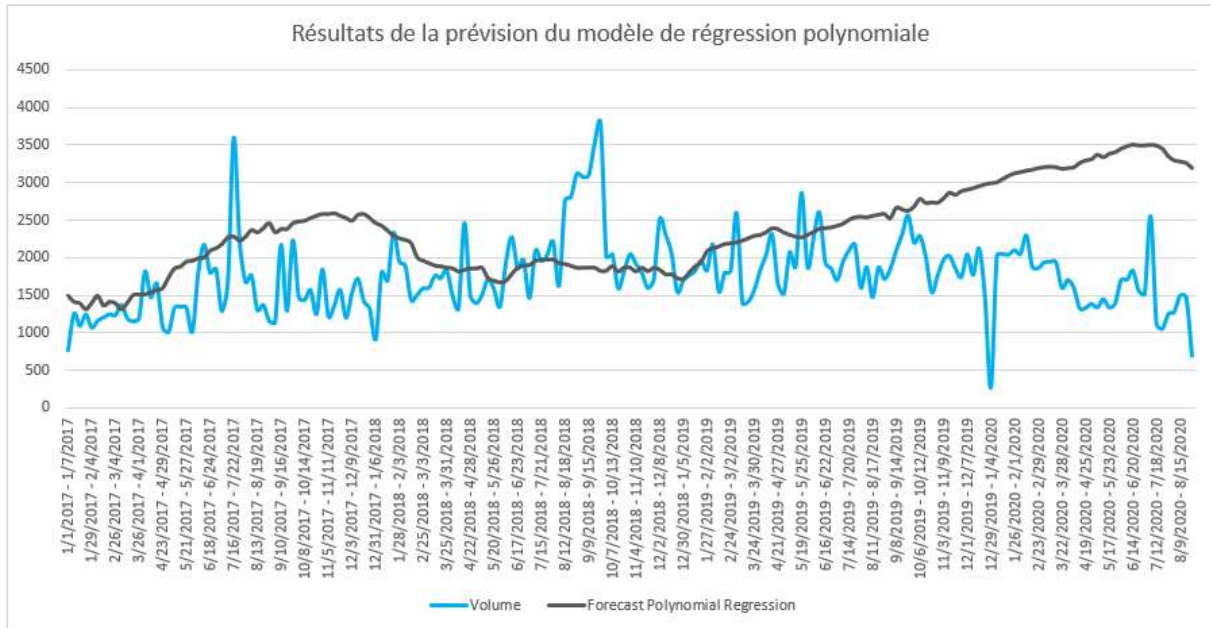


Figure V.17: Résultats de la prévision du modèle de régression polynomiale, applique sur les données de GEHC.

La valeur absolue de l'écart entre la prévision du modèle et les données historiques est représentée par la figure V.18. Dans 50% des cas, la différence entre le nombre des commandes que le modèle prévoit et celui présent dans les données est inférieure à 400 pièces (Figure V.21). Ce résultat est meilleur que le modèle de simulation vu dans le chapitre précédent. Cependant, il ne représente pas les tendances de la vraie charge de réparation (Figure V.18).

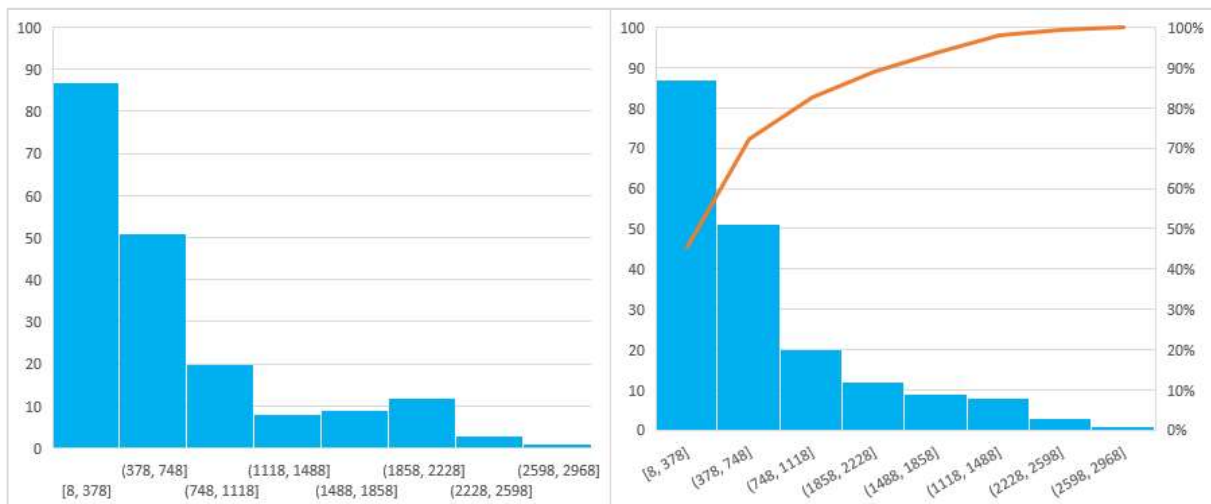


Figure V.18: La distribution de l'erreur de prévision du modèle de régression polynomiale.

2.4.3.b Le modèle de régression Ridge

La figure V.19 présente la prévision générée par le modèle de régression Ridge. Comme pour la régression polynomiale, le modèle ne suit pas la tendance générale des données et ne permet pas de détecter les pics de la demande quand il est appliqué à toutes les références.

En comparaison avec la régression polynomiale, la régression Ridge donne des résultats différents et une prévision de tendance différente. Cependant, en termes de précision, les deux modèles ont une précision assez semblable.

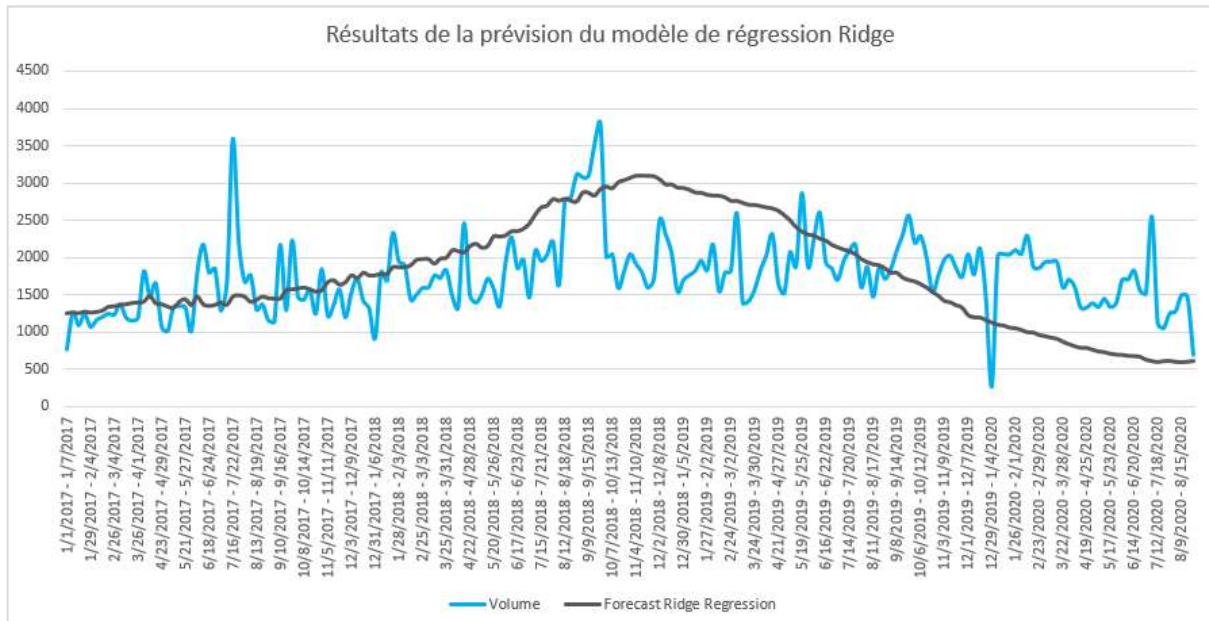


Figure V.19: Résultats de la prévision du modèle de régression ridge, applique sur les données de GEHC

Dans plus de 50% des observations (Figure V.20), la prévision générée par le modèle a un écart de plus de 500 pièces par semaine. Même si le modèle a, en globalité, une précision proche de la régression polynomiale, l'histogramme des erreurs montre que pour un ensemble de pièces, l'erreur est beaucoup plus importante que celles du modèle polynomial.

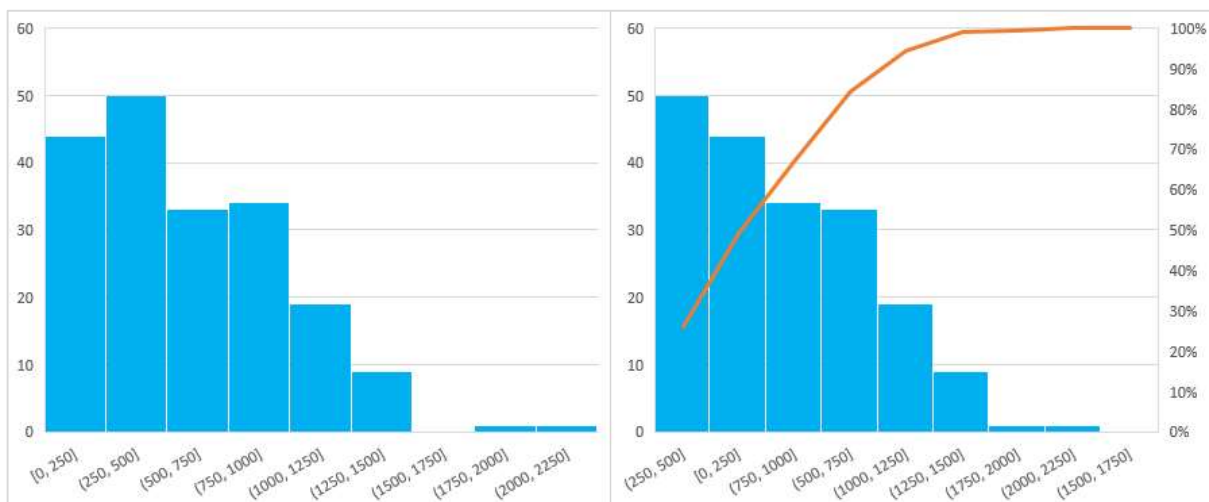


Figure V.20: La distribution de l'erreur de prévision du modèle de régression Ridge.

2.4.3.c Le modèle de régression Lasso

La figure V.21 visualise la même prévision mais avec la régression Lasso. Un écart clair entre la prévision et l'historique est présent. Le modèle est le dernier en classement en termes de précision par rapport aux modèles déjà présentés.

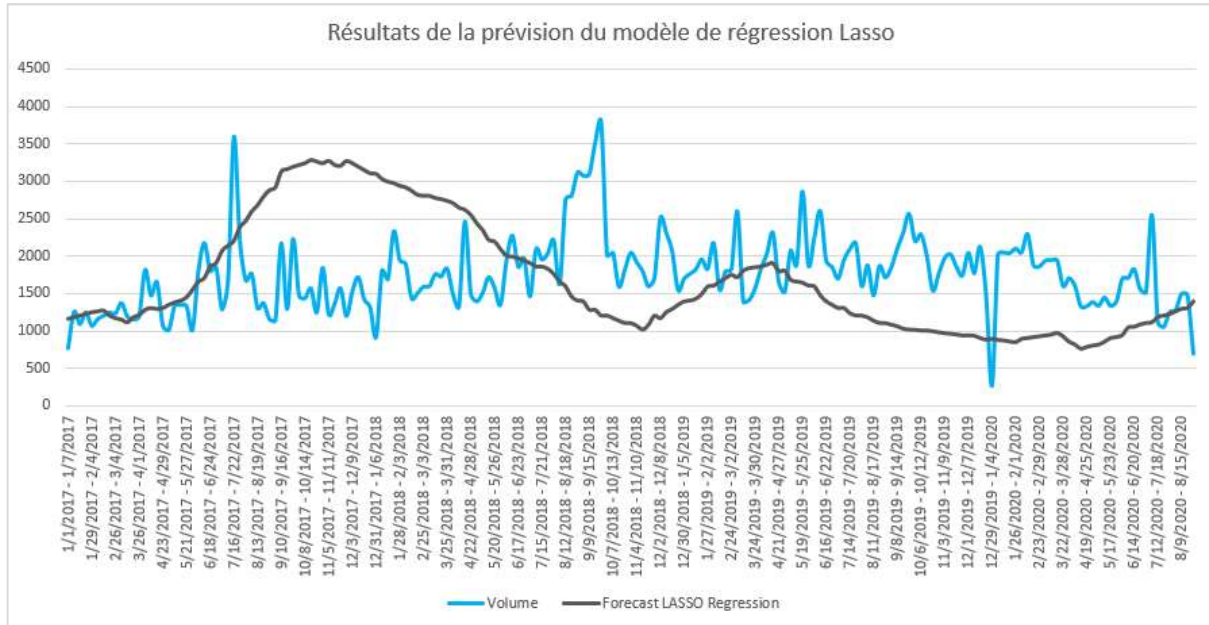


Figure V.21: Résultats de la prévision du modèle de régression Lasso, applique sur les données de GEHC

De la même façon que pour les modèles précédents, dans la figure V.22, nous montrons la distribution de l'erreur que l'algorithme génère. En comparaison avec les deux modèles précédents, ici nous avons l'erreur la plus importante.

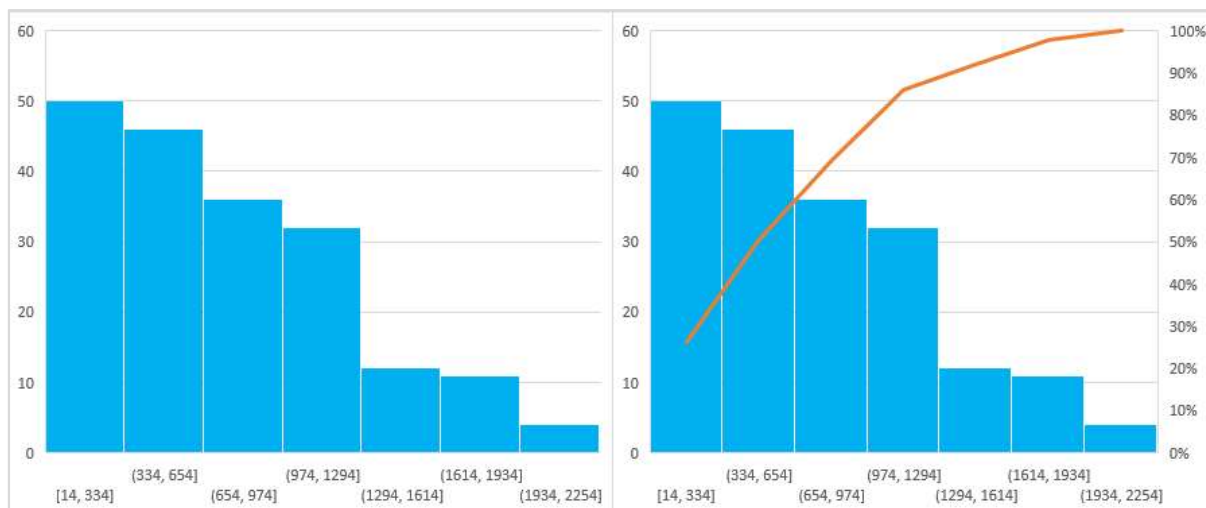


Figure V.22: La distribution de l'erreur de prévision du modèle de régression Lasso.

2.4.4 Résultats avec l'affectation des modèles par référence

Dans cette section nous allons donner une vue globale de la performance de chacun de nos modèles, mais chacun appliqué par référence. Ceci veut dire que chaque modèle a été uniquement appliqué aux pièces pour lesquelles il a offert la meilleure performance.

Le tableau V.5 présente le nombre et le pourcentage des références de pièces pour lesquelles chaque modèle offre la meilleure performance.

Modèle	Nombre de références pour lesquelles le modèle est meilleur	Pourcentage du l'échantillon
Régression Polynomiale	227	45.40%
Régression Ridge	125	25.00%
Régressons Lasso	148	29.60%

Tableau V.5 : Nombre de références pour lesquelles chaque modèle a été choisi comme le meilleur.

Sur notre échantillon contenant un ensemble des 500 références des pièces les plus commandées, le classement commence par le modèle de régression polynomiale qui a offert la meilleure performance sur un ensemble de 227 références, représentant 45.4% des références.

Suivi par le modèle de régression Lasso avec 148 références (29.6%) et enfin la régression Ridge avec 125 références (25%).

Le tableau V.6 présente la performance des modèles, en prenant la précision maximale, moyenne et minimale. Dans le calcul de la performance, nous prenons uniquement les différentes références auxquelles ces modèles ont été affectés, et sur lesquelles ils ont été utilisés.

Modèle	Score des modèles	
Régression Polynomiale	Min	0.001
	Avg	0.516
	Max	0.705
Régression Ridge	Min	0.001
	Avg	0.380
	Max	0.481
Régressons Lasso	Min	0.001
	Avg	0.397
	Max	0.651

Tableau V.6: Le minimum, maximum et la moyenne des scores de chaque modèle de prévision

La performance que le tableau V.6 présente est le score calculé sur l'historique. La meilleure performance moyenne a été réalisée par le modèle de régression polynomiale avec un score de 0.516 sur un ensemble de 227 références, suivie par le modèle régression Lasso avec 0.397 sur 148 références et ensuite la régression Ridge avec un score de 0.380 sur un ensemble de 125 références.

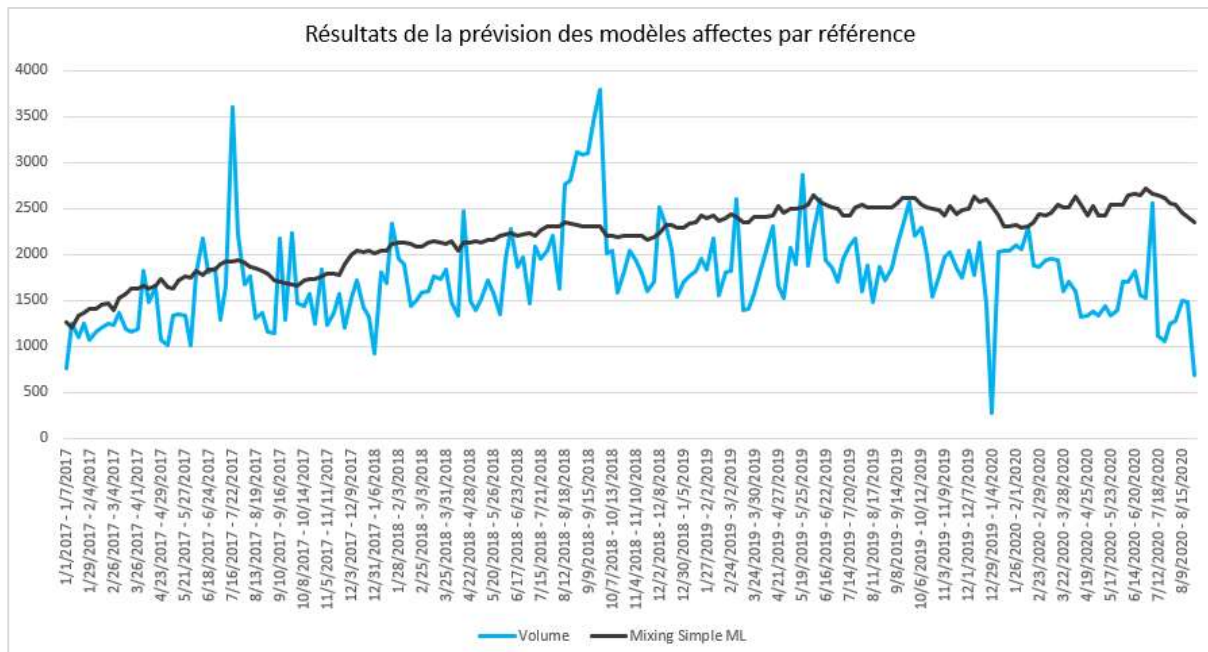


Figure V.23: Résultats de la prévision du modèle Machine Learning appliqué sur les données de GEHC.

La figure V.23 visualise le résultat de la prévision générée par le modèle de Machine Learning lorsqu'on applique chaque modèle aux références sur lesquelles il performe le mieux. Au contraire de la première stratégie où on applique chaque modèle séparément, ce modèle ML permet de prévoir la tendance générale du volume de réparation qui, dans notre cas, augmente d'une façon progressive au milieu de l'horizon (entre le mois de Mai et le mois de Juin) et décroît à la fin.

L'application des modèles par référence apporte une grande amélioration à la performance du modèle, avec un score de 0.601. Cependant, le modèle ne permet toujours pas de détecter les changements brusques dans la charge.

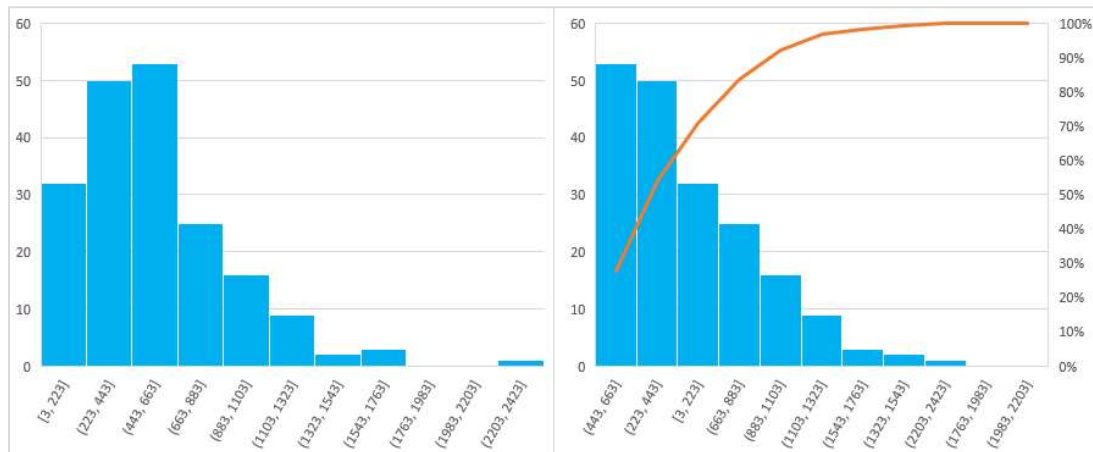


Figure V.24: Distribution de l'écart entre la prévision et l'historique

Les graphes de la figure V.24 présente la distribution de l'écart entre les vrais volumes de réparation et les prévisions du modèle. L'écart que cette approche fournie est beaucoup plus faible en comparaison des méthodes précédentes. L'application par référence a permis une réduction de +50% de l'erreur et a amélioré sa fiabilité. Cependant, pour notre contexte et pour la nature de l'activité, la détection des changements brusques de la charge reste la problématique la plus importante.

IV Conclusion

Dans ce chapitre, nous avons présenté l'implémentation d'un modèle Machine Learning pour l'objectif de prévision de la charge dans les centres de réparation. Rappelons que la finalité de notre travail est l'analyse des améliorations potentielles d'approches de type Machine Learning et sa mise en place en contexte opérationnel et n'intègre pas le développement de nouvelles techniques de classification ou d'inférence. Nous avons ainsi présenté une méthodologie d'implémentation des modèles d'apprentissage empruntée à la littérature que nous avons déclinée pour l'appliquer sur un système de prévision particulier.

Nous avons ensuite testé ce système sur les données réelles de l'historique des opérations de GEHC et puis présenté et analysé les résultats. Par rapport à la simulation seule, les modèles de Machine Learning offrent des résultats prometteurs en améliorant sensiblement la précision de prédiction. Cependant, quand les modèles sont appliqués d'une façon systématique à toutes les références, ils donnent des résultats avec une précision très faible.

Ensuite nous avons introduit un algorithme testant et appliquant chaque modèle sur les références pour lequel il offre la meilleure performance. Avec cette approche, les modèles permettent d'avoir une meilleure précision et aussi de détecter la tendance générale des données, au contraire des approches précédentes.

Avantages de l'approche

Les modèles Machine Learning offrent deux avantages principaux par rapport aux modèles classiques :

- ***La prévision des tendances de la charge :***

Comme vu dans le chapitre 4, les prévisions avec la simulation suivent une logique systématique et génère des valeurs prédictibles. Ceci fait que les volumes cumulée sont stables et quasi constants, au contraire de la réalité où les volumes sont rarement constants.

Les modèles Machine Learning que nous avons utilisés permettent de prévoir des volumes variables et permettent de détecter l'augmentation ou la baisse des volumes de réparation avec une meilleure précision que celle de la simulation.

- ***La prise en considération des nouvelles données :***

Les modèles de prévisions classiques se basent uniquement sur les données précédentes et sur l'historique (dépendant de l'algorithme) pour décider la valeur du prochain point dans les prévisions.

Pour les modèles Machine Learning, on se base plutôt sur les données derrière chaque observation et non les valeurs de ces observations elles-mêmes. Ceci permet de détecter les « patterns » et les liens entre les données d'entrée et la valeur de leurs observations. Grâce à cette différence fondamentale entre les deux approches, les modèles Machine Learning permettent de prévoir d'une façon plutôt fiable la tendance générale des données.

Quand on les applique à toutes les références, les modèles de Machine Learning offrent une meilleure précision de 0.601 par rapport aux modèles appliqués d'une façon individuelle : PR, RR et LR avec respectivement 0.516, 0.380 et 0.397 comme précision.

Cependant, et malgré les avantages des modèles ML, la précision de ce dernier n'est pas suffisante. Les pics de la charge (ou chute) restent indétectables par notre approche ce qui ne répond pas à notre problématique générale et au besoin des centres de réparation.

Limites de l'approche

Malgré les avantages du modèle d'apprentissage par rapport aux modèles classiques, on peut dégager les limites suivantes.

Le modèle ML permet de prévoir la tendance générale de la charge de réparation mais ne permet pas de détecter les détails de cette demande, comme les pics ou les chutes brusques. Ceci peut être considéré comme du « sous-apprentissage ». Comme on l'a présenté au début

de ce chapitre, il signifie que le modèle n'est pas assez détaillé pour détecter les changements sur une durée réduite de temps mais uniquement sur un grand intervalle.

Dans certains cas, le sous-apprentissage peut être résolu via la recalibration du modèle et son paramétrage. Cependant, dans notre cas, il n'est pas dû aux paramètres (toutes les combinaisons possibles ont été testées) du modèle mais bien à sa nature. Pour arriver à cette conclusion, nous avons fait des tests sur des échantillons de taille réduite générés de façon manuelle contenant des pics et des chutes de la demande.

Après une analyse des erreurs que le modèle génère, nous remarquons que, dans plus de 70% des observations, le modèle prévoit une solution (réparation ou achat) différente de la réalité, mais un volume avec une « bonne » précision. Ce qui montre que la limite du modèle n'est pas l'estimation du volume lui-même mais dans le choix des décisions (réparation ou achat).

Pistes d'améliorations de la performance du modèle

Après investigation des problèmes de performance de notre modèle ML, notre conclusion est que les limites de performance sont principalement liées à la capacité de la prévision des solutions d'approvisionnement (réparation ou achat) et non à la prévision du volume lui-même. Le volume de réparation que nous essayons de prévoir est en fait une combinaison de trois étapes de décision, connectées pour générer la donnée finale, à savoir le nombre de pièces à réparer. Ces étapes sont les suivantes :

- ***La décision d'approvisionnement*** : qui signifie qu'à ce moment-là, il y a un besoin et qu'il faut passer des commandes d'approvisionnement.
- ***Le choix de la solution d'approvisionnement*** : où les experts choisissent entre l'achat des pièces ou la réparation de l'inventaire défectueux pour répondre à la demande.
- ***Le calcul des volumes à commander*** : où les responsables calculent la quantité d'approvisionnement en prenant en compte le besoin, la taille de lots, la quantité économiques et d'autres contraintes qui peuvent être imposées par les fournisseurs.

Pour remédier à cette limite, dans le chapitre suivant, nous allons introduire une subdivision de notre problème en 3 sous-problèmes, puis, nous allons modéliser et implémenter un modèle de type Machine Learning Hybride pour la prévision des trois étapes nécessaires pour l'estimation de la charge de réparation. Nous allons présenter la structure du modèle, le processus d'implémentation et ensuite évaluer sa performance par rapport aux modèles que nous avons testés précédemment.

Chapitre VI : Implémentation d'un modèle Machine Learning Hybride pour la prévision de la charge dans les centres de réparation dans une CL-PR-BF

I Introduction

Dans la section précédente, nous avons vu l'application de modèles d'apprentissage pour la prévision de la charge dans les centres de réparation de pièces de rechange de GEHC. Ces modèles donnent des résultats qui sont mieux que la simulation, cependant, ils ne permettent pas de prévoir les pics et les chutes de la demande.

Dans ce chapitre nous proposons un modèle Machine Learning Hybride pour la prévision de la charge de réparation. Un modèle Machine Learning Hybride est un modèle combinant plusieurs modèles d'apprentissage différents pour répondre à une problématique donnée.

Selon Tsai *et al.* (2010), les modèles Machine Learning Hybride utilisent deux ou plusieurs modèles d'apprentissage adaptés à des parties de la problématique afin de capturer les caractères discrets ou continus du problème d'optimisation.

II Les modèles Machine Learning Hybrides dans la littérature

Un nombre important d'applications de modèles Machine Learning Hybride est présent dans la littérature. Ce type de modèles a été appliqué dans nombre de domaines tels la médecine, l'aéronautique, la bourse et l'énergie.

Banihashemi, Saeed *et al.* (2016) implémente un modèle hybride combinant un modèle de régression et un modèle de classification pour la prévision et la classification de la consommation énergétique. Le travail combine un modèle de réseau de neurones (Artificial Neural Network) couvrant les paramètres continus, avec un arbre de décision (Decision Tree) couvrant les paramètres discrets. La performance du modèle final a été évaluée avec MSE (Mean Squared Error) et une fonction objective hybride. (Figure VI.1)

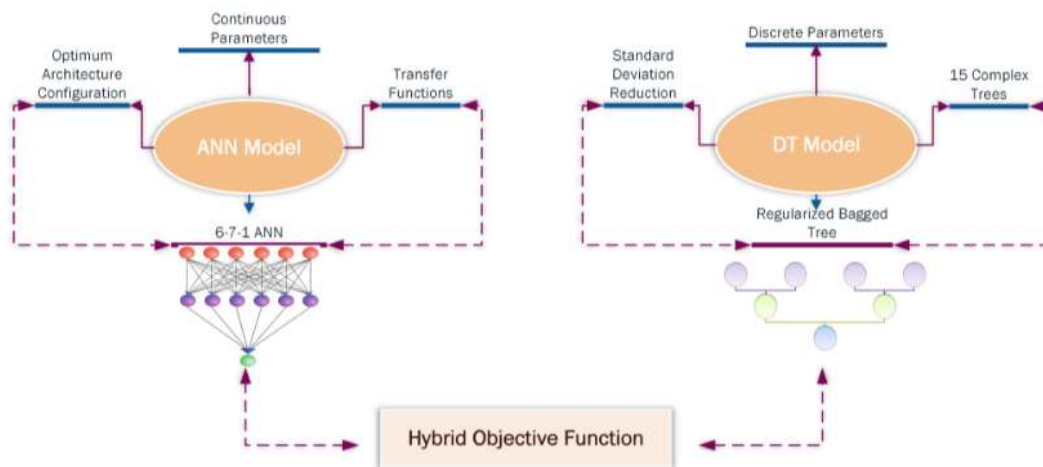


Figure VI.1: Structure du modèle hybride (Banihashemi, Saeed *et al.* (2016))

Nguyen *et al.* (2020) propose une combinaison entre un modèle LSSVM (Least Squares Support Vector Machines) avec un PSO (Particle Swarm Optimization). En suivant la même approche que la référence précédente, le modèle utilise une fonction d'évaluation hybride évaluant les deux modèles en même temps afin de trouver les paramètres avec lesquels la combinaison des deux modèles est la plus performante (Figure VI.2).

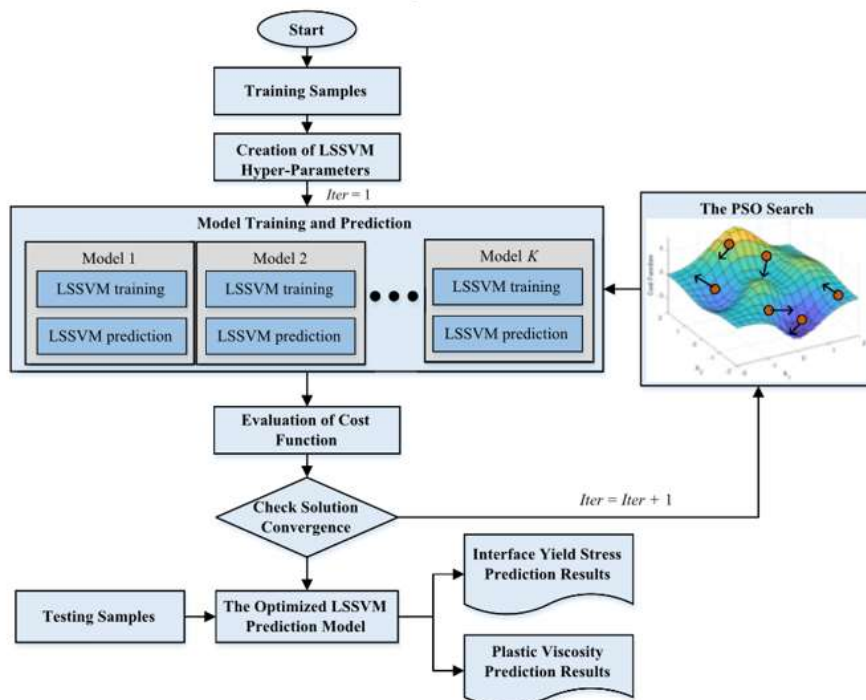


Figure VI.2: Structure du modèle hybride proposé par Nguyen *et al.* (2020)

Après l'étude d'autres travaux tels que les publications de Gao *et al.* (2020), Lecomte *et al.* (2013) et Olivares *et al.* (2016), nous remarquons que les implémentations suivent toutes la même démarche permettant de manipuler des paramètres discrets ainsi que des paramètres continus.

Notre implémentation suit la même méthodologie. Dans la section suivante, nous allons voir la structure ainsi que l'implémentation d'un modèle hybride combinant trois modèles, deux de classification et un de régression pour la décomposition du processus de prévision de la charge de réparation dans une CL-PR-BF.

III Implémentation du modèle Machine Learning Hybride pour la prévision de la charge de réparation dans une CL-PR-BF

1 La structure du modèle Hybride

Dans le chapitre précédent, nous avons implémenté un modèle pour la prévision de la charge dans les centres de réparation dans une CL-PR-BF. Basé sur une méthode issue du Machine Learning, il considère le processus dans sa globalité. Or, en spécifiant d'avantage ce dernier, on se rend compte que le processus décisionnel sous-jacent peut se décomposer en plusieurs sous-étapes laissant ainsi apparaître une perspective prometteuse quant à l'utilisation d'un modèle de Machine Learning Hybride.

Aussi nous proposons de nous appuyer sur cette décomposition afin de proposer une méthode de Machine Learning Hybride illustrée dans la figure VI.3.

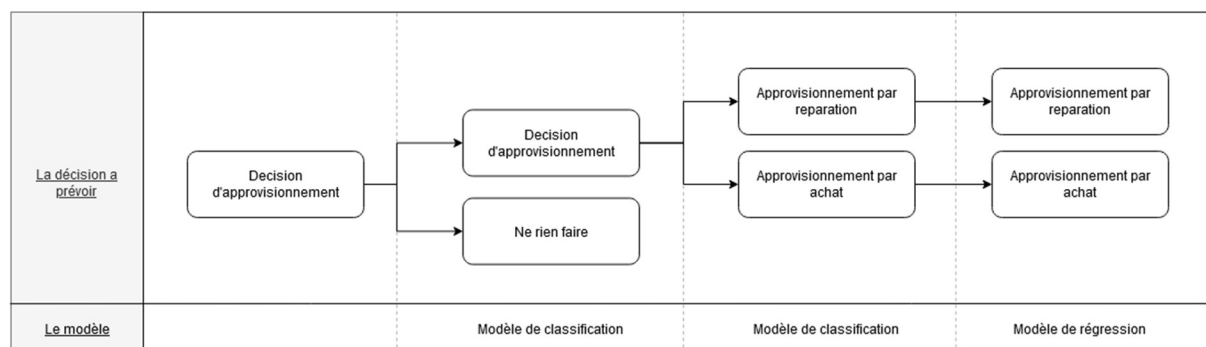


Figure VI.3: Subdivision du processus de décision en plusieurs sous-décisions

La première étape dans le processus de décision est un choix entre « faire un approvisionnement » ou « ne rien faire », cette décision est prise par les experts chaque jour, pour chaque référence, c'est un choix entre deux « classes » de décisions. Les modèles adaptés pour prévoir la « classe » d'une nouvelle observation sont les modèles de classification.

De même pour la deuxième décision, c'est un choix entre deux classes et un modèle de classification est le plus adapté pour la prévoir.

Pour la troisième partie du processus de décision, les planners sont amenés à estimer la quantité de pièces à commander ou à envoyer en réparation. Contrairement aux décisions précédentes, c'est une estimation d'une quantité. Ainsi, un modèle de régression parait le plus adapté.

2 Modèles additionnels pour la classification

Les modèles que nous avons présentés jusqu'à maintenant sont des modèles de régression. Dans le chapitre précédent, la prévision que nous avons effectuée s'est limitée à de la régression pure, et nous avons utilisé les modèles suivants : la Régression Polynomiale, la Régression Ridge et la Régression Lasso.

En plus de ces modèles, nous avons besoin de modèles de « classification » pour les deux premières décisions du processus de planification. Pour faire ceci, nous allons utiliser les modèles de classification les plus utilisés dans la littérature, qui sont KNN (K Nearest Neighbors), SVM (Support Vector Machine), DT (Decision Trees) et RN (Random Forests).

2.1 KNN

L'algorithme KNN (k Nearest Neighbors) est un modèle Machine Learning de classification. KNN est une adaptation du modèle NN (Nearest Neighbors) qui est un modèle plus général qui est utilisé pour d'autres problèmes que la classification.

L'objectif du modèle KNN est de catégoriser un nouveau point (observation) en se basant sur un « vote » de majorité de ses k plus proches voisins. Pour faire ceci, le modèle se base sur deux éléments : le profil de l'observation et une métrique de distance.

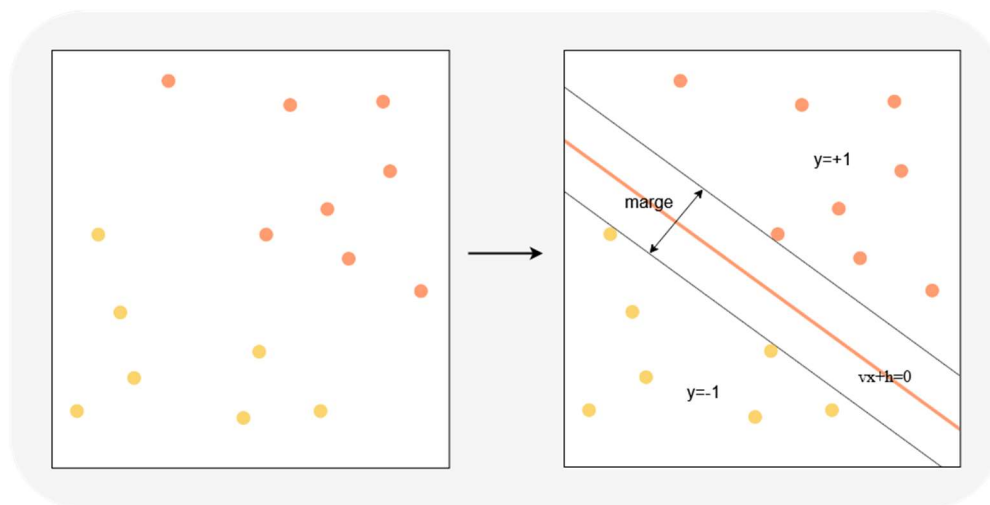
Le paramètre sur lequel le modèle s'appuie pour faire une classification est le nombre de voisins k à prendre en considération. La figure VI.4 présente le résultat d'un exemple simplifié d'une classification KNN en prenant en considération 1 à 6 voisins.



Figure VI.4: Exemple d'une classification « multi-classe » avec une modèle KNN 2 dimensions, en prenant en considération k voisins, avec k allant de 1 à 6 (distance Euclidienne)

2.2 SVM (Support Vector Machines)

Les machines à vecteurs supports (ou Support Vector Machines) ont été introduites au début des années 90. Elles sont des méthodes largement utilisées en théorie de l'apprentissage



statistique.

Figure VI.5: Exemple d'une classification avec un modèle SVM binaire

Les modèles SMV appartiennent à une famille de modèle appelée « Les séparateurs à vastes marges ». Le mode de fonctionnement de ce type de modèles est assez intuitif et consiste à chercher la frontière linéaire ou non linéaire permettant d'avoir la marge maximale entre deux ou plusieurs classes d'observations (Figure VI.5). Le cas le plus simple d'application de modèles SVM est celui où les données d'entraînement viennent uniquement de deux classes différentes, on parle alors de classification binaire.

2.3 Les arbres de décision

Selon Sutton-Charani *et al.* (2014) les arbres de décision sont «un modèle de classification/prédiction simple d'utilisation et qui permet une interprétation très utile aux experts. Aujourd'hui très largement utilisé dans différents domaines, il représente un classifieur standard, facilement modulable selon les applications ».

La première apparition de travaux proposant des arbres de décision se retrouve dans l'article de Belson *et al.* (1956). C'est dans les travaux de Quinlan *et al.* (1986) que le concept d'arbre de décision a été utilisé dans le cadre d'une problématique de classification, ouvrant ainsi la voie à son utilisation dans les algorithmes d'Intelligence Artificielle et notamment de Machine Learning (Sutton-Charani *et al.* (2014)).

Un arbre de décision est une structure composée de l'ensemble des éléments suivants :

- **Une racine :** qui représente le nœud d'origine duquel n'importe quelle estimation ou prévision commence. Ce nœud a deux ou plusieurs branches sortantes depuis lesquelles on choisit la variable d'entrée.
- **Des nœuds internes :** (ou simplement nœuds) qui ont des flèches entrantes ainsi que des flèches sortantes. Un nœud est une étape intermédiaire de décision, permettant d'affiner le choix précédent dans l'objectif d'arriver à une feuille qui est la destination finale.
- **Des branches :** qui est un lien entre une racine, ou un nœud vers un nœud ou une feuille. Une racine contient toujours une condition pour qu'elle soit « activée ». Cette condition est vérifiée sur la variable ou le vecteur d'entrée.

La nature des conditions de branches ainsi que la valeur de variables peuvent être numériques ou sous forme de classes.

- **Les feuilles :** sont simplement des nœuds finaux. C'est la valeur de la feuille finale que nous cherchons, et qui représente le résultat de prévision de l'arbre.

La figure VI.6 présente la structure basique d'un arbre de décision :

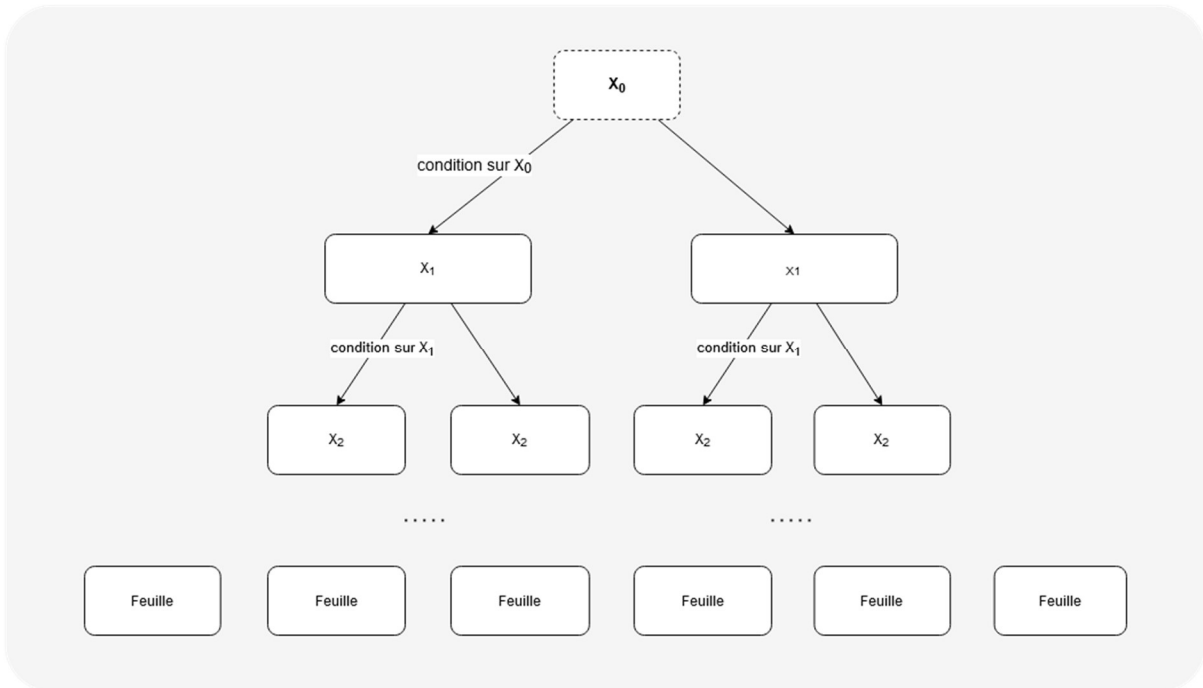


Figure VI.6 : Structure basique d'un arbre de décision

2.4 Les forêts aléatoires

Malgré la simplicité de compréhension, de mise en place et d'utilisation des arbres de décisions ainsi que leur capacité d'offrir une bonne précision sur un ensemble de problématiques, parfois ce type de modèles montre des limites en termes de performances, limites liées à leur nature. Les arbres de décisions sont très dépendants des données d'apprentissage et la structure de l'arbre final dépend principalement de l'arrangement de ces données.

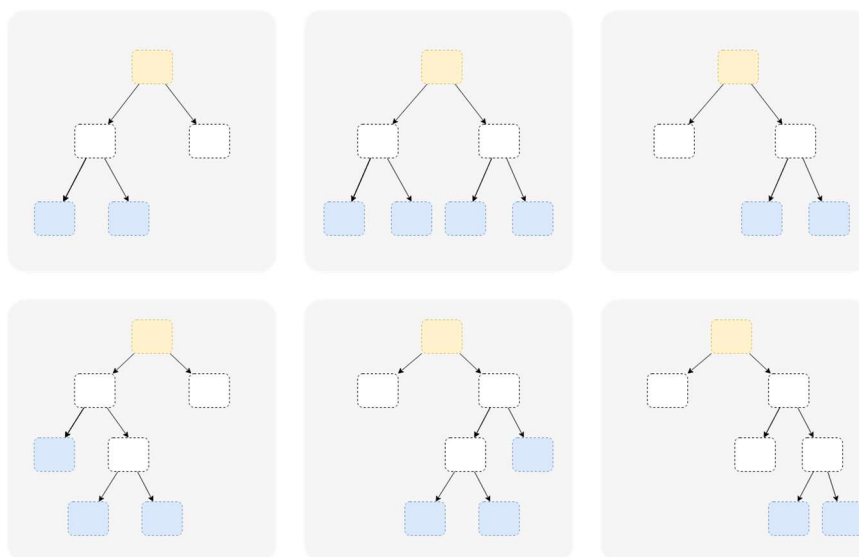


Figure VI.7 : Illustration d'une forêt aléatoire composée de multiples configurations d'un arbre de décision

Les forêts aléatoires ont été introduites la première fois par Breiman *et al.* (2001). A cause des limites qu'on peut rencontrer avec un arbre de décision qui sont liées à la structure des données d'apprentissage. Une forêt aléatoire est un ensemble d'arbres construits de façons aléatoires, avec une variation d'un grand nombre de configurations comme le montre la figure VI.7.

3 Processus d'implémentation

Partant du processus global, nous l'avons décomposé en trois sous processus décisionnels pour lesquels nous allons mettre en place des méthodes de Machine Learning afin de modéliser leurs comportements. La méthodologie suivie se déroule selon le schéma suivant :

Pour chaque phase et pour chaque méthode qui lui sont dédiées, nous commençons par une étape de calibration de la méthode, en utilisant les données de l'historique. Nous mesurons la performance de cette étape grâce à deux indicateurs de performance : le taux de classification correcte pour les modèles de classification et la précision de la régression (présentée auparavant) pour les modèles de régression. Ensuite, nous testons la méthode ainsi paramétrée sur des données de test et nous en mesurons la performance par l'indicateur R2.

Une fois ces phases de calibration et de tests effectuées, nous testons l'ensemble des combinaisons de chainage possibles afin de sélectionner celle offrant la meilleure performance aux yeux de l'indicateur global (R2).

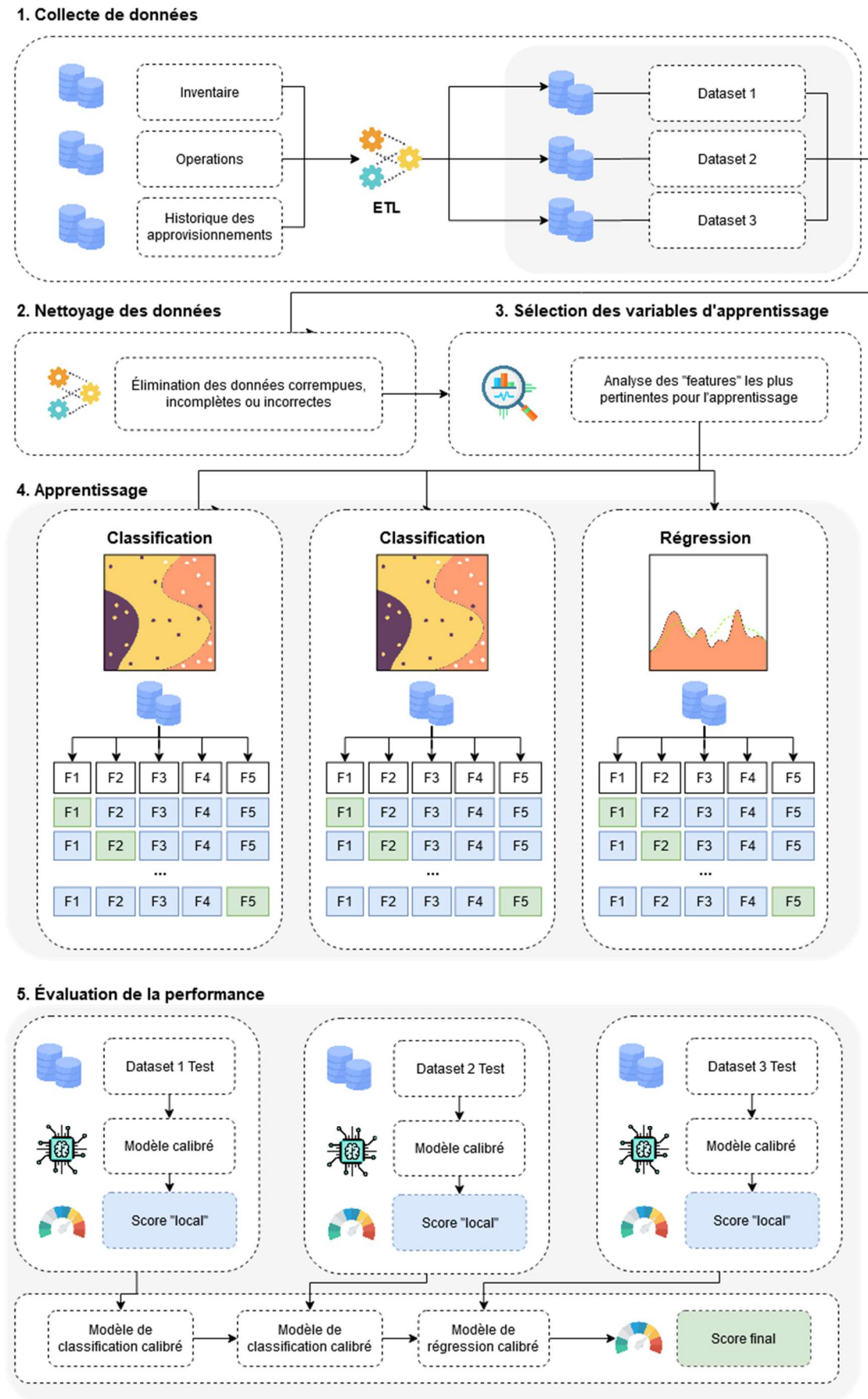


Figure VI.8 : Processus d'implémentation du modèle Machine Learning Hybride

IV Evaluation de la performance du modèle hybride

Afin de tester la performance du modèle hybride, nous avons mis en place trois métriques principales : le score sur les données de calibration, le score sur les données de test et le score final.

Pour les sous-modèles composant le modèle hybride, le score de calibration signifie la performance du modèle sur les données sur lesquelles il a été calibré. Ce score ne permet pas vraiment d'évaluer la performance du modèle, puisque ces données font partie de sa base de « connaissance ». Cependant, il constitue une référence importante pour évaluer la capacité de « généralisation » du modèle.

Avoir l'indicateur local qui permet d'évaluer la capacité de généralisation du modèle est important. Ceci représente comment le modèle performe sur les nouvelles de données de test, par rapport à sa performance de calibration.

Le tableau VI.1 présente les différents scores pour chaque sous modèle ainsi que pour le modèle complet. La meilleure combinaison de notre modèle a un score de 0.704, qui est plus précis qu'un modèle Machine Learning simple ou par rapport à l'utilisation de modèles de simulation.

La combinaison des modèles			La performance sur les données de calibration			La performance sur les données de test			Score final
C1	C2	R	P1	P2	P3	P1	P2	P3	P123
kNN	kNN	RR	0.513	0.804	0.317	0.405	0.775	0.208	0.119
kNN	kNN	PR	0.513	0.804	0.613	0.405	0.775	0.581	0.452
kNN	kNN	Lasso	0.513	0.804	0.401	0.405	0.775	0.332	0.291
kNN	SVM	RR	0.513	0.833	0.317	0.405	0.744	0.208	0.18
kNN	SVM	PR	0.513	0.833	0.513	0.405	0.744	0.581	0.391
kNN	SVM	Lasso	0.513	0.833	0.401	0.405	0.744	0.332	0.103
kNN	DT	RR	0.513	0.602	0.317	0.405	0.594	0.208	0.082
kNN	DT	PR	0.513	0.602	0.513	0.405	0.594	0.581	0.311
kNN	DT	Lasso	0.513	0.602	0.401	0.405	0.594	0.332	0.172
kNN	RF	RR	0.513	0.782	0.317	0.405	0.61	0.208	0.14
kNN	RF	PR	0.513	0.782	0.513	0.405	0.61	0.581	0.332
kNN	RF	Lasso	0.513	0.782	0.401	0.405	0.61	0.332	0.461
SVM	kNN	RR	0.903	0.804	0.317	0.849	0.775	0.208	0.303
SVM	kNN	PR	0.903	0.804	0.613	0.849	0.775	0.581	0.704
SVM	kNN	Lasso	0.903	0.804	0.401	0.849	0.775	0.332	0.461
SVM	SVM	RR	0.903	0.833	0.317	0.849	0.744	0.208	0.271
SVM	SVM	PR	0.903	0.833	0.513	0.849	0.744	0.581	0.304
SVM	SVM	Lasso	0.903	0.833	0.401	0.849	0.744	0.332	0.293
SVM	DT	RR	0.903	0.602	0.317	0.849	0.594	0.208	0.102
SVM	DT	PR	0.903	0.602	0.513	0.849	0.594	0.581	0.171
SVM	DT	Lasso	0.903	0.602	0.401	0.849	0.594	0.332	0.044

SVM	RF	RR	0.903	0.782	0.317	0.849	0.61	0.208	0.241
SVM	RF	PR	0.903	0.782	0.513	0.849	0.61	0.581	0.376
SVM	RF	Lasso	0.903	0.782	0.401	0.849	0.61	0.332	0.228
DT	kNN	RR	0.483	0.804	0.317	0.303	0.775	0.208	0.381
DT	kNN	PR	0.483	0.804	0.613	0.303	0.775	0.581	0.477
DT	kNN	Lasso	0.483	0.804	0.401	0.303	0.775	0.332	0.291
DT	SVM	RR	0.483	0.833	0.317	0.303	0.744	0.208	0.412
DT	SVM	PR	0.483	0.833	0.513	0.303	0.744	0.581	0.571
DT	SVM	Lasso	0.483	0.833	0.401	0.303	0.744	0.332	0.221
DT	DT	RR	0.483	0.602	0.317	0.303	0.594	0.208	0.359
DT	DT	PR	0.483	0.602	0.513	0.303	0.594	0.581	0.41
DT	DT	Lasso	0.483	0.602	0.401	0.303	0.594	0.332	0.204
DT	RF	RR	0.483	0.782	0.317	0.303	0.61	0.208	0.418
DT	RF	PR	0.483	0.782	0.513	0.303	0.61	0.581	0.533
DT	RF	Lasso	0.483	0.782	0.401	0.303	0.61	0.332	0.401
RF	kNN	RR	0.719	0.804	0.317	0.681	0.775	0.208	0.538
RF	kNN	PR	0.719	0.804	0.613	0.681	0.775	0.581	0.602
RF	kNN	Lasso	0.719	0.804	0.401	0.681	0.775	0.332	0.421
RF	SVM	RR	0.719	0.833	0.317	0.681	0.744	0.208	0.388
RF	SVM	PR	0.719	0.833	0.513	0.681	0.744	0.581	0.59
RF	SVM	Lasso	0.719	0.833	0.401	0.681	0.744	0.332	0.484
RF	DT	RR	0.719	0.602	0.317	0.681	0.594	0.208	0.3
RF	DT	PR	0.719	0.602	0.513	0.681	0.594	0.581	0.476
RF	DT	Lasso	0.719	0.602	0.401	0.681	0.594	0.332	0.379
RF	RF	RR	0.719	0.782	0.317	0.681	0.61	0.208	0.303
RF	RF	PR	0.719	0.782	0.513	0.681	0.61	0.581	0.491
RF	RF	Lasso	0.719	0.782	0.401	0.681	0.61	0.332	0.399

Tableau VI.1 : Performance du modèle Machine Learning Hybride, avec les scores de chaque sous-modèle pour la phase de calibration et de test

La figure VI.9 présente le résultat de la prévision du modèle Machine Learning Hybride avec la courbe noire, en comparaison avec l'historique. Il est clair que la performance du modèle avec la subdivision en sous problèmes apporte une amélioration claire par rapport à l'utilisation d'un modèle simple d'apprentissage.

En plus de la performance finale du modèle de 0.704 qui reste la plus forte dans notre étude, le modèle a pu prévoir les pics et la chute de la demande du mois d'avril 2020 (Figure VI.9). La différence la plus importante avec les modèles précédents est que le modèle hybride permet la prévision des pics de la demandes ainsi que la chute de cette dernière d'une façon plutôt fiable Ceci répond à la problématique la plus importante de notre projet qui était la détection des pics de la charge dans les centres de réparation.

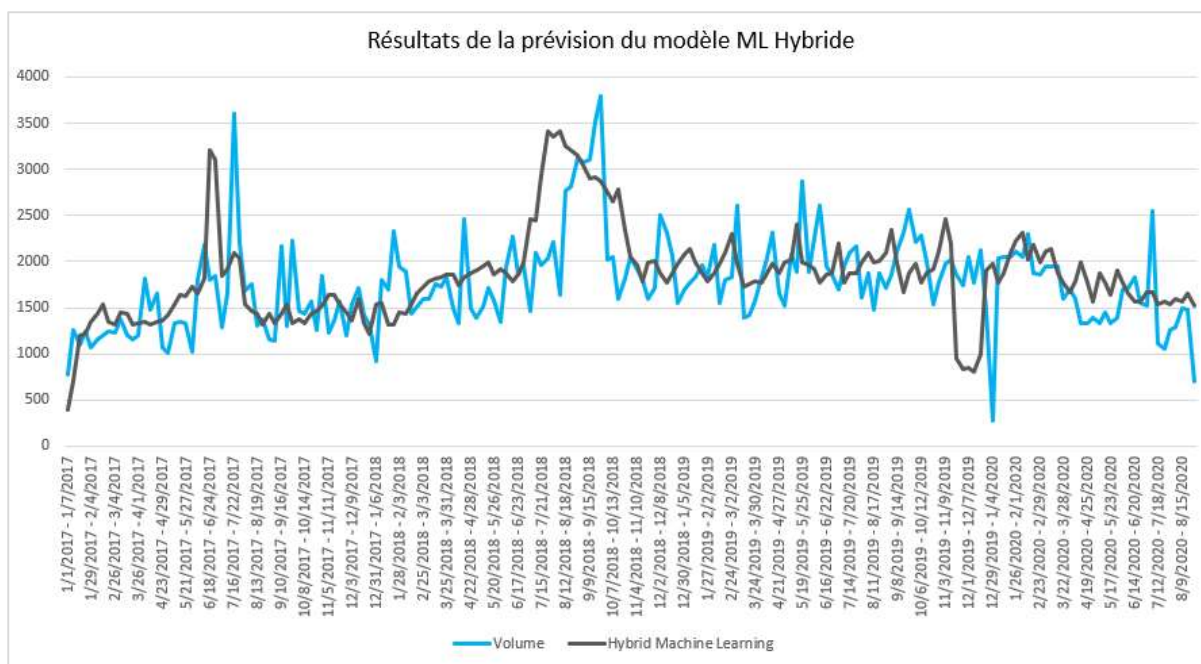


Figure VI.9 : Résultat de la prévision du modèle Machine Learning Hybride par rapport à l'historique

La figure VI.10 présente une visualisation de l'écart entre la prévision et la réalité. En comparaison avec les modèles présentés précédemment, ce modèle offre la meilleure performance et l'écart le plus petit.

Dans la figure VI.10, des points aberrants dans les écarts sont présents avec des valeurs très élevées par rapports aux modèles précédents. Ces points représentent l'écart entre les pics prévus puisqu'ils ont été estimés légèrement en avance. Ce qui est considéré un écart important dans les calculs, mais en réalité, ce résultat ne représente pas un problème car les informations les plus importantes sont la prévision du pic, son amplitude et la période dans laquelle il apparait avec une tolérance d'erreur de plus ou moins quelques semaines (un maximum de ~3 semaines).

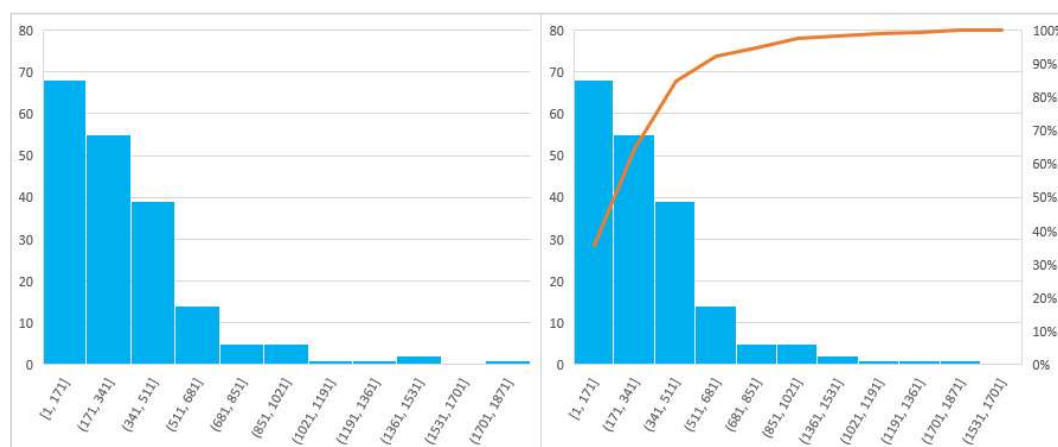


Figure VI.10 : Distribution de la valeur absolue de l'écart entre l'historique et les prévisions du modèle Machine Learning Hybride

V Conclusion

Dans les différents chapitres de ce rapport, nous avons testé l'utilisation de la simulation seule (Chapitre 4), puis les modèles Machine Learning appliqués indépendamment (Chapitre 5), et enfin dans ce chapitre, nous avons mis en place un modèle Machine Learning Hybride pour remédier aux limites des modèles précédents.

Pour avoir également une référence des modèles statistiques classiques de prévision, nous avons appliqué les modèles SMA (Simple Moving Average) et SES (Simple Exponential Smoothing) sur nos données de test. Nous avons ensuite calculé leur précision et nous avons présenté les résultats finaux des scores dans le tableau VI.2 avec les modèles Machine Learning.

Nous avons appliqué tous les modèles sur un ensemble de 500 pièces les plus demandées à GEHC, puis, pour chaque référence, nous choisissons le modèle qui offre la meilleure précision (en calculant la valeur de R2). Le tableau VI.2 présente le récapitulatif des résultats du test :

- Pour chaque modèle, le nombre des références pour lesquelles il a été choisi comme le meilleur.
- Le score moyen que chaque modèle offre sur les différentes références pour lesquelles il a été choisi comme le plus fiable.

Catégorie	Modèle	Meilleur pour x refs	Score Moy	
Approches Classiques	SMA	18	3.60%	0.2336912
	SES	10	2.00%	0.0945333
Machine Learning	Régression Polynomiale	61	12.20%	0.5413877
	Régression Ridge	28	5.60%	0.1833402
	Régression Lasso	22	4.40%	0.1071830
Machine Learning Hybride	[Classificateur, Classificateur, Modèle de régression]	361	72.20%	0.7690038

Tableau VI.2 : Récapitulatif des résultats des modèles classiques, des modèles Machine Learning et des modèles Machine Learning Hybride

Le modèle hybride a offert la meilleure précision pour 361 références d'un échantillon de 500 pièces, ce qui représente un pourcentage de 72.2%, suivi par 12.2% du modèle de régression polynomiale et ensuite le reste des modèles avec des performances semblables.

En se basant sur ce test, il est clair que le meilleur modèle dépend de la nature de chaque référence. Cependant, pour la plupart des pièces, le modèle hybride offre largement la meilleure performance.

Ceci ne signifie pas que nous allons uniquement utiliser le modèle Machine Learning Hybride. En effet, l'implémentation de ce système de calibration, les tests de performance et la

génération de prévision nous a permis d'avoir un mécanisme capable d'appliquer automatiquement le meilleur modèle pour chaque référence, que ça soit un modèle Machine Learning ou un modèle statistique classique.

Chapitre VII : Conclusion, application industrielle et perspectives

1 Conclusion

La prévision de la charge est un besoin crucial pour les centres de réparation dans une chaîne logistique de pièces de rechange. La précision de cette dernière est importante et joue un rôle direct sur le temps de livraison des pièces de rechange et sur la satisfaction des clients.

L'objectif de ce projet de recherche était de proposer un ensemble d'outils méthodologiques, des modèles et des techniques pour répondre à la problématique d'estimation de la charge de réparation dans une CL-PR-BF. Nos travaux se sont articulés autour de trois contributions principales qui nous ont permis de répondre à notre problématique.

En mettant une place un modèle **de simulation de notre chaîne logistique**, nous avons pu implémenter le modèle de gestion DRP et nous l'avons appliqué sur les données historiques pour constater une grande différence entre la logique théorique du modèle et celle qui est effectivement appliquée par les responsables d'approvisionnement.

Pour prévoir les décisions d'approvisionnement qui ne sont pas reflétées par un modèle DRP classique, nous avons appliqué des **modèles d'apprentissage (Machine Learning)** qui se basent sur les données historiques (expérience) pour apprendre et estimer les décisions des responsables d'approvisionnement. Le résultat de notre implémentation a montré une amélioration de la performance par rapport à la simulation. Cependant, le modèle a des limites dans la prévision des pics ou des chutes brusques dans le volume de charge appliqué au centre de réparation.

Après investigation, nous avons conclu que la limitation du modèle n'est pas causée par l'erreur dans l'estimation des volumes mais par une erreur d'appréciation de la décision prise par le responsable d'approvisionnement. En effet, les décisions d'approvisionnement sont soumises aux stratégies d'inventaire et de niveau de service. Ces stratégies peuvent changer drastiquement avec le temps.

Pour remédier à cette limitation, nous avons introduit un modèle de **Machine Learning Hybride**, composée de plusieurs sous modèles, chacun responsable de la prévision d'une partie de la décision finale : (i) la décision de passer une commande, (ii) la décision du type

d'approvisionnement (achat de neuf ou réparation de pièces) et (iii) la quantité de pièces commandées.

Pour prévoir les trois éléments décisionnels précités, nous avons utilisé deux modèles de classification (i et ii) et un modèle de régression (iii). Ensuite, nous avons appliqué ces modèles sur les données de GEHC.

Nos résultats montrent que le modèle hybride offre une réelle amélioration de la prévision par rapport au modèle de Machine Learning classique. En effet, cette approche permet d'offrir la meilleure performance pour 72% des références produits= étudiées, pour 22% des références la meilleure performance est produite par les modèles Machine Learning standards et finalement les méthodes classiques offrent le meilleur résultat pour 6% des références.

Finalement, nous avons implémenté un algorithme qui utilise les données historiques pour choisir la meilleure option entre la prévision par Machine Learning Hybride, Machine Learning standard ou par les approches classiques.

2 Application industrielle des travaux chez GEHC

Au cours de ce projet de thèse, nous avons appliqué nos modèles sur les données opérationnelles de GEHC. Pour le faire, j'ai développé un outil permettant de collecter les données de plusieurs bases de données de l'entreprise, afin de les traiter et les mettre en forme pour les modèles de prévision.

J'ai également implémenté les différents modèles de prévision dans cet outil avec un algorithme capable de choisir le meilleur modèle pour chaque référence de pièces de rechange.

Cet outil a évolué avec le temps pour devenir une plateforme d'aide à la décision qui contient les différentes métriques des centres de réparation. Cette dernière a été développée avec une architecture en modules interconnectés, ce qui permet d'ajouter à chaque fois de nouveaux éléments s'il y a besoin.

Voici quelques interfaces de l'outil :

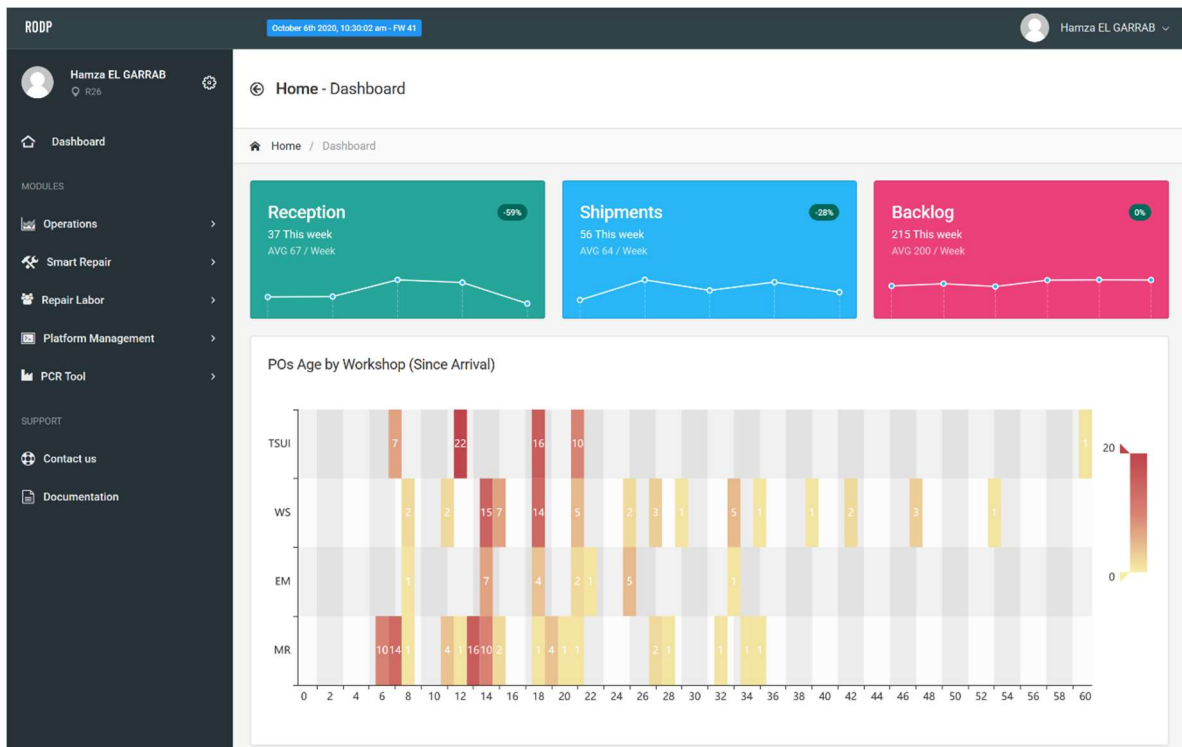


Figure VII.1 : Dashboard d'accueil de la plateforme

Les figures VII.1 et VII.2 illustrent le premier module de la plateforme qui est responsable du calcul des différentes métriques de performances, ainsi que de la visualisation de l'état de l'activité en temps réel.

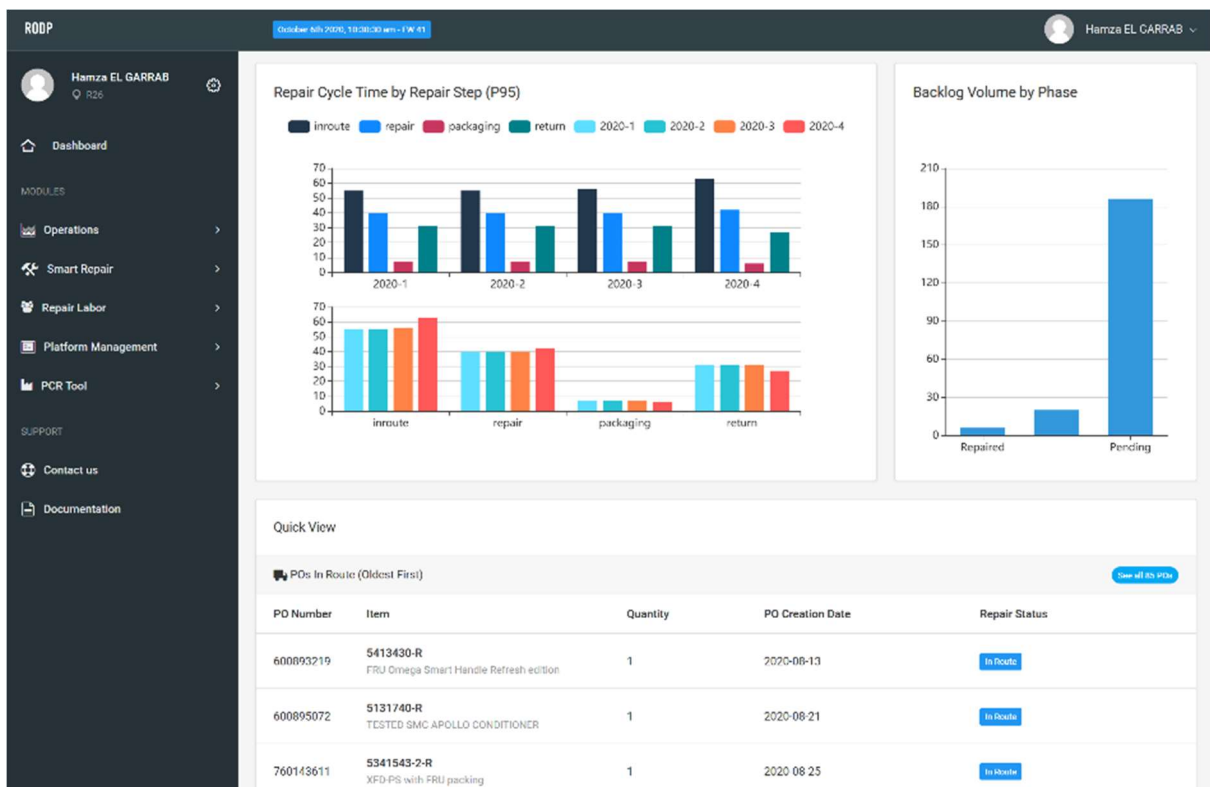


Figure VII.2 : Dashboard d'accueil de la plateforme

La figure VII.1 présente une vue des principales métriques de performance du centre de réparation. Elle montre l'historique et les tendances des réceptions dans les ateliers, les livraisons des pièces après la réparation et aussi la charge actuelle. Cette vue permet également d'avoir une vue de l'âge des commandes dans les ateliers à travers une « heat map ».

L'outil contient également les mesures de temps de réparation des pièces (Figure VII.2). Ces temps de réparation sont visualisés par trimestre, et pour chaque sous-étape de réparation. Ceci permet aux managers de détecter les étapes à améliorer dans le processus.

En plus des métriques de performance, l'outil contient les détails de chaque demande de réparation et permet le suivi de chaque pièce physique d'une façon individuelle (Figure VII.2).

La figure VII.3 montre un autre module permettant la visualisation en temps réel de la charge dans les différents ateliers du centre de réparation. Ces éléments permettent aux responsables de planifier l'activité et les ressources humaines.

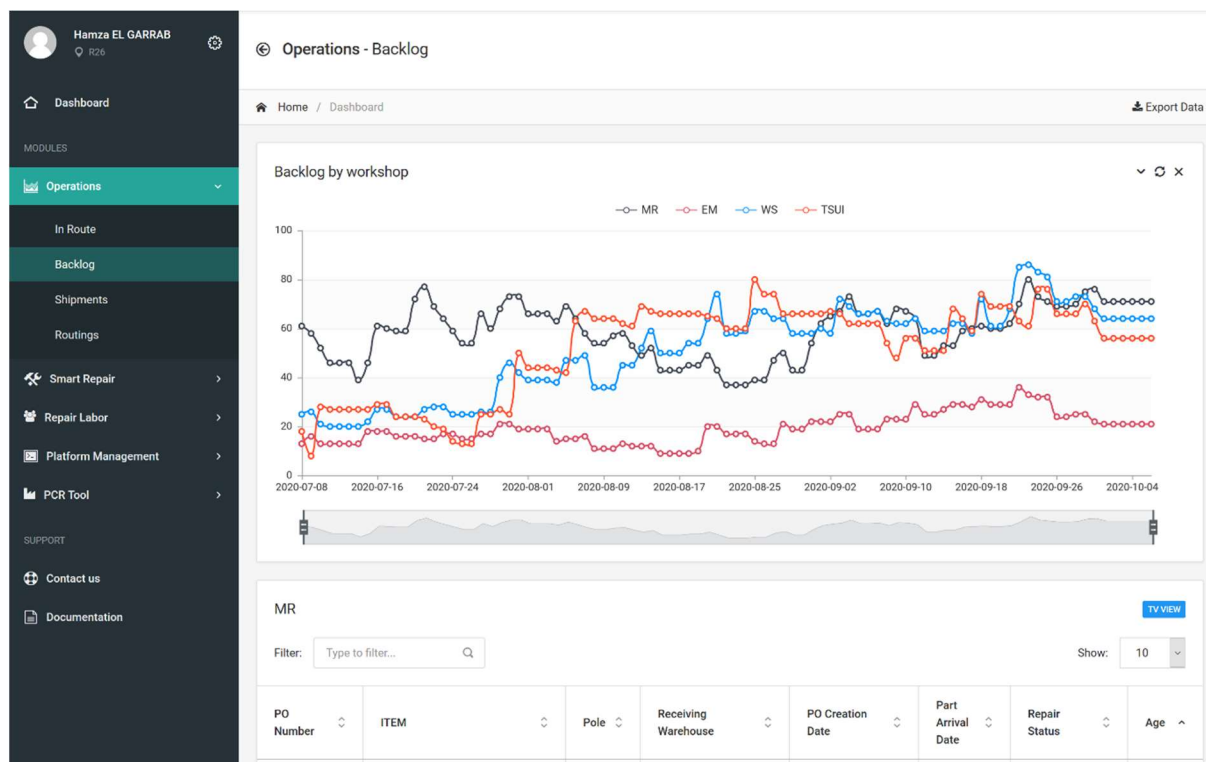


Figure VII.3 : Visualisation de l'historique de la charge des différents ateliers du centre de réparation

La figure VII.4 présente le module permettant la gestion de l'équipe de réparation et des absences. Il permet la visualisation et la saisie des absences des équipes de réparation, ce qui facilite la tâche de gestion du personnel aux managers.

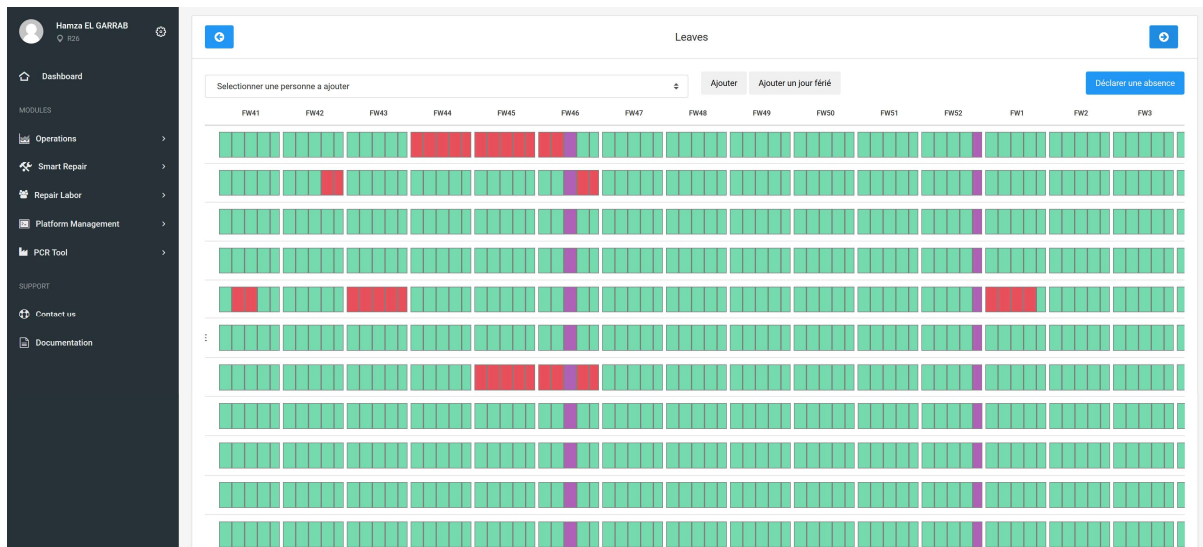


Figure VII.4 : Module de gestion d'équipe du centre de réparation

L'ensemble de ces modules et autres sont actuellement en cours d'utilisation par le centre de réparation européen. Le module de prévision dans lequel j'ai implémenté les différents modèles permet uniquement de générer des résultats dans des fichiers CSV et n'a pas encore d'interface graphique. Ce module est encore dans une branche de développement, et pas encore mis en production pour des raisons de priorités opérationnelles.

3 Perspectives

Dans ce projet de thèse, nous avons proposé plusieurs méthodes d'estimation de la charge dans les centres de réparation en appliquant des méthodes d'apprentissage.

Nous avons également implémenté une « bibliothèque » de modèles de prévision qui sont proposés dans une plateforme d'aide à la décision au sein du service de réparation des pièces de rechange de GEHC.

Ces prévisions permettent aux responsables d'avoir une visibilité fiable des volumes qui seront reçus dans chaque atelier pour les six prochains mois. Cependant, nous n'avons pas traité l'exploitation de ces prévisions pour l'optimisation des opérations dans le centre de réparation.

Ainsi, un premier axe de recherche serait d'utiliser cette prévision dans les choix décisionnels liés aux activités et à l'organisation des opérations au sein des ateliers de réparation. Cette problématique englobe deux aspects principaux :

- ***Les approvisionnements des composants et de la matière première*** : cette première partie concerne les approvisionnements des composants et des produits chimiques et nécessaires pour réparer les pièces de rechange. Cette étape est importante pour assurer un bon compromis inventaire vs. disponibilité. La complexité de la tâche vient de la

multitude des pannes possibles des diverses références, de l'interchangeabilité de certains composants compatibles avec plusieurs pièces et de la difficulté d'approvisionnement pour certains composants en fin de vie.

- ***L'organisation des ateliers*** : cette partie de la problématique concerne la décision de la répartition des activités dans les différents ateliers, en fonction de la charge dans chacun. C'est-à-dire la planification des opérations et l'ordonnancement des tâches sur les lignes de réparation en fonction du volume et des composants disponibles dans l'inventaire.

Un des apports importants du travail réalisé dans le cadre de cette thèse CIFRE est la création d'une plateforme informatique qui va au-delà de la problématique de prévision de la charge. Cet outil contient d'ores et déjà un ensemble de modules opérationnels auxquels nous pouvons ajouter des modèles pour répondre à de futures problématiques.

Références

1. Oliver, R. Keith, and Michael D. Webber. "Supply-chain management: logistics catches up with strategy." *Outlook* 5.1 (1982): 42-47.
2. Fleischmann, Moritz, Jo AEE Van Nunen, and Ben Gräve. "Integrating closed-loop supply chains and spare-parts management at IBM." *Interfaces* 33.6 (2003): 44-56.
3. Boldoczki, Sandra, Andrea Thorenz, and Axel Tuma. "The environmental impacts of preparation for reuse: A case study of WEEE reuse in Germany." *Journal of Cleaner Production* 252 (2020): 119736.
4. Mentzer, John T., et al. "Defining supply chain management." *Journal of Business logistics* 22.2 (2001): 1-25.
5. Lazrak, Adnane. *Amélioration des processus de prévision et de gestion des stocks dans le cas d'une chaîne logistique des pièces de rechange*. Diss. Nantes, Ecole des Mines, 2015.
6. Cooper, Martha C., Douglas M. Lambert, and Janus D. Pagh. "Supply chain management: more than a new name for logistics." *The international journal of logistics management* 8.1 (1997): 1-14.
7. Tanyas, P. M. "The role of optimization in supply chain management". (2014, November 06).
8. Botta-Genoulaz, Valerie, P-A. Millet, and Bernard Grabot. "A survey on the recent research literature on ERP systems." *Computers in industry* 56.6 (2005): 510-522.
9. El Miloudi, F., Semma, H., & Riane, F. (2015, December). *Enjeux financiers dans la gestion des chaînes logistiques*.
10. Ellram, Lisa M., Wendy L. Tate, and Corey Billington. "Understanding and managing the services supply chain." *Journal of Supply Chain Management* 40.3 (2004): 17-32.
11. Hofmann E., & Belin O., *Supply Chain Finance Solutions relevance-propositions-market value*, Heidelberg: Springer-verlag 2011, p 1-100.
12. Peng, Hui, et al. "Uncertainty factors, methods, and solutions of closed-loop supply chain—A review for current situation and future prospects." *Journal of Cleaner Production* (2020): 120032.

13. Homrich, Aline Sacchi, et al. "The circular economy umbrella: Trends and gaps on integrating pathways." *Journal of Cleaner Production* 175 (2018): 525-543.
14. Van Wassenhove, Luk N., and V. Daniel R. Guide. *The evolution of closed-loop supply chain research*. INSEAD, 2008.
15. Khor, Kuan Siew, and Zulkifli Mohamed Udin. "Impact of reverse logistics product disposition towards business performance in Malaysian E&E companies." *Journal of Supply Chain and Customer Relationship Management* 2012 (2012): 1.
16. Huiskonen, J. (2001). Maintenance spare parts logistics: Special characteristics and strategic choices. *International journal of production economics*, 71(1-3), 125-133.
17. Tapia-Ubeda, Francisco J., et al. "Supplier Selection for Spare Parts Supply Chain Networks." *IFAC-PapersOnLine* 52.13 (2019): 2237-2242.
18. Janvier-James, Assey Mbang. "A new introduction to supply chains and supply chain management: Definitions and theories perspective." *International Business Research* 5.1 (2012): 194-207.
19. Kim, HyungTae, and SuHwan Kim. "Data Mining based army repair parts demand forecast." *The Korean Data & Information Science Society* 30.2 (2019): 429-444.
20. Dekker, Rommert, et al. "On the use of installed base information for spare parts logistics: A review of ideas and industry practice." *International Journal of Production Economics* 143.2 (2013): 536-545.
21. Shinohara, Yukio. "Method for predicting a demand for repair parts." U.S. Patent No. 6,647,371. 11 Nov. 2003.
22. Van Jaarsveld, Willem, and Rommert Dekker. "Spare parts stock control for redundant systems using reliability centered maintenance data." *Reliability Engineering & System Safety* 96.11 (2011): 1576-1586.
23. Guo, Feng, et al. "A double-level combination approach for demand forecasting of repairable airplane spare parts based on turnover data." *Computers & Industrial Engineering* 110 (2017): 92-108.
24. Kareem, B., and A. S. Lawal. "Spare parts failure prediction of an automobile under criticality condition." *Engineering Failure Analysis* 56 (2015): 69-79.
25. Baryannis, George, et al. "Supply chain risk management and artificial intelligence: state of the art and future research directions." *International Journal of Production Research* 57.7 (2019): 2179-2202.

26. Ngai, E. W. T., et al. "Decision support and intelligent systems in the textile and apparel supply chain: An academic review of research articles." *Expert Systems with Applications* 41.1 (2014): 81-91.
27. Giri, Chandadevi, et al. "A Detailed Review of Artificial Intelligence Applied in the Fashion and Apparel Industry." *IEEE Access* 7 (2019): 95364-95384.
28. Min, Hokey. "Artificial intelligence in supply chain management: theory and applications." *International Journal of Logistics: Research and Applications* 13.1 (2010): 13-39.
29. Ko, Mark, Ashutosh Tiwari, and Jörn Mehnen. "A review of soft computing applications in supply chain management." *Applied Soft Computing* 10.3 (2010): 661-674.
30. Gillian, Nicholas Edward. *Gesture recognition for musician computer interaction*. Diss. Queen's University Belfast, 2012.
31. Kautz, Oliver, Alexander Roth, and Bernhard Rumpe. "Achievements, Failures, and the Future of Model-Based Software Engineering." (2018): 221-236.
32. Ouzayd, Fatima, and J. Saadi. "Proposition d'une approche de modélisation du système logistique hospitalier combinant UML et les Réseaux de Petri temporisés : Cas de la chaîne logistique médicamenteuse globale au Maroc." (2011).
33. Di Martinelly, Christine, Fouad Riane, and Alain Guinet. "A Porter-SCOR modelling approach for the hospital supply chain." *International Journal of Logistics Systems and Management* 5.3-4 (2009): 436-456.
34. Moussa, Mohamed, and Khaled Belkadi. "Modélisation et Simulation de Flux Dans un Service d'Imagerie de l'HMRUO." *5th International Conference : Sciences of Electronic, Technologies of Information and Telecommunications, March. 2009*.
35. Belkadi K., Tangy A., « Modélisation et simulation de l'unité d'Hémobiologie du Laboratoire Central d'ORAN Algérie », 8th International Conference of Modeling and Simulation (MOSIM'10), Tunisia, May 10-12, 2010
36. Lemoine, David. *Modèles génériques et méthodes de résolution pour la planification tactique mono-site et multi-site*. Diss. 2008.
37. Huet J.C., Paris J.L., Gourgand M., Kouiss K., « Modèle de connaissance générique du circuit du Médicament dans un hôpital », 8th International Conference of Modeling and Simulation MOSIM'10), Tunisia, May 10-12, 2010
38. Torre, Damiano, et al. "How consistency is handled in Model Driven Software Engineering and UML: a survey of experts in academia and industry." (2018).

39. Savic, Dragan A., and Witold Pedrycz. "Evaluation of fuzzy linear regression models." *Fuzzy sets and systems* 39.1 (1991) : 51-63.
40. Prairie, Yves T. "Evaluating the predictive power of regression models." *Canadian Journal of Fisheries and Aquatic Sciences* 53.3 (1996) : 490-492.
41. Turing, Alan M. "Computing machinery and intelligence (1950)." *The Essential Turing: The Ideas that Gave Birth to the Computer Age*. Ed. B. Jack Copeland. Oxford: Oxford UP (2004): 433-64.
42. Gao, Kaifeng, et al. "Julia language in machine learning: Algorithms, applications, and open issues." *Computer Science Review* 37 (2020): 100254.
43. Praveena, M., and V. Jaiganesh. "A literature review on supervised machine learning algorithms and boosting process." *International Journal of Computer Applications* 169.8 (2017): 32-35.
44. Sutton-Charani, Nicolas. *Apprentissage à partir de données et de connaissances incertaines : application à la prédiction de la qualité du caoutchouc*. Diss. 2014.
45. Chen, Yanbei, Xiatian Zhu, and Shaogang Gong. "Person re-identification by deep learning multi-scale representations." *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017.
46. Goodfellow, Ian, et al. *Deep learning*. Vol. 1. Cambridge: MIT press, 2016.
47. Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: theory and applications." *Neurocomputing* 70.1-3 (2006): 489-501.
48. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436-444.
49. Sauvanaud, Carla. *Monitoring et détection d'anomalie par apprentissage dans les infrastructures virtualisées*. Diss. Toulouse, INSA, 2016.
50. Roy, Jean-Francis. "Apprentissage automatique avec garanties de généralisation à l'aide de méthodes d'ensemble maximisant le désaccord." (2018).
51. Lebarbier, Emilie, and Tristan Mary-Huard. "Le critère BIC : fondements théoriques et interprétation." (2004).
52. Gamal, Donia, et al. "Implementation of Machine Learning Algorithms in Arabic Sentiment Analysis Using N-Gram Features." *Procedia Computer Science* 154 (2019): 332-340.
53. Gallagher, Colm V., et al. "Development and application of a machine learning supported methodology for measurement and verification (M&V) 2.0." *Energy and Buildings* 167 (2018): 8-22.

54. Lecomte, Sébastien. *Classification partiellement supervisée par SVM: application à la détection d'événements en surveillance audio*. Diss. Troyes, 2013.
55. Olivares, Javier. Scaling out-of-core k-nearest neighbors computation on single machines. Diss. 2016.
56. Gudivada, Venkat, Amy Apon, and Junhua Ding. "Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations." *International Journal on Advances in Software* 10.1 (2017): 1-20.
57. Le Guen, Monique. "La boîte à moustaches pour sensibiliser à la statistique." *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 73.1 (2002): 43-64.
58. Tukey, John W. *Exploratory data analysis*. Vol. 2. 1977.
59. Chesneau, Christophe. "Modèles de régression." (2017).
60. Tibshirani, Robert. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (1996): 267-288.
61. Tsai, Chih-Fong, and Ming-Lun Chen. "Credit rating by hybrid machine learning techniques." *Applied soft computing* 10.2 (2010): 374-380.
62. Banihashemi, Saeed, Grace Ding, and Jack Wang. "Developing a hybrid model of prediction and classification algorithms for building energy consumption." *Energy Procedia* 110 (2017): 371-376.
63. Nguyen, The-Duong, Thu-Hien Tran, and Nhat-Duc Hoang. "Prediction of interface yield stress and plastic viscosity of fresh concrete using a hybrid machine learning approach." *Advanced Engineering Informatics* 44 (2020): 101057.
64. Belson WA. A technique for studying the effects of television broadcast. *J R Stat Soc* 1956, 5:195.
65. Quinlan JR. Discovering rules by induction from large collections of examples. In: Michie D, ed. *Expert Systems in the Micro-electronic Age*. Edinburgh: Edinburgh University Press; 1979, 168–201.
66. Breiman, L. (2001). Random forests. *Machine learning*, 45(1) :5–32.
67. Quinlan, J. Ross. "Induction of decision trees." *Machine learning* 1.1 (1986): 81-106.

Titre : Amélioration de la chaîne logistique de pièces de rechange en boucle fermée : Application des modèles d'apprentissage

Mots clés : chaîne logistique, pièces de rechange réparables, prévision, Machine Learning

Résumé : Dans le domaine de service après-vente et particulièrement dans la maintenance, l'intervention rapide et la réparation du bien du client est un élément clé pour sa satisfaction et pour la création de l'image de marque dans le marché.

Le travail présenté dans cette thèse propose une approche de type Machine Learning pour l'amélioration du flux informationnel de la chaîne logistique de pièces de rechange.

Notre contribution se focalise sur la prévision de la charge dans les centres de réparation des pièces de rechange, qui sont les fournisseurs principaux des pièces utilisées pour réparer les systèmes des clients.

La grandeur de la chaîne logistique et sa complexité, le grand nombre des références de pièces ainsi que la multitude des cas spéciaux (pays avec de lois spécifiques, pièces particulières...) fait que les approches classiques n'offrent pas des prévisions fiables pour les services de réparation.

Dans ce projet, nous proposons des algorithmes d'apprentissage permettant la construction de la connaissance à partir de grands volumes de données, au lieu des méthodes de prévision classiques.

Nous allons voir les modèles dans la littérature, présenter notre méthodologie, et ensuite implémenter les modèles et évaluer leur performance en comparaison avec les algorithmes existants.

Title: The improvement of the closed loop spare parts supply chain: Applying Machine Learning models

Keywords: supply chain, repairable spare parts, forecasting, Machine Learning

Abstract: In the field of after-sales service and particularly in maintenance, rapid intervention and repair of the customer's property is a key element for his satisfaction and for the creation of the brand image in the market.

The work presented in this thesis proposes an inspired Machine Learning approach for improving the information flow of the spare parts supply chain.

Our contribution is focused on forecasting the load in spare parts repair centers, which are the primary suppliers of spare parts used to repair customer systems.

The size of the supply chain and its complexity, the large number of part references as well as the multitude of special cases (countries with specific laws, particular parts, etc.) mean that traditional approaches do not offer reliable forecasts for the repair teams.

In this project, we propose learning algorithms allowing the construction of knowledge from large sets of data, instead of classical forecasting methods.

We will study models in the literature, present our methodology, and then implement the models and evaluate their performance in comparison with existing models.