

Modélisation énergétique des navires auto-déchargeurs à l'aide des méthodes d'apprentissage automatique

Par

Ginette Brenda NOUDJI LAPE

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE LA MAÎTRISE
EN GÉNIE DE LA PRODUCTION AUTOMATISÉE

M. Sc. A.

MONTRÉAL, LE 29 JUILLET 2025

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

©Tous droits réservés

Cette licence signifie qu'il est interdit de reproduire, d'enregistrer ou de diffuser en tout ou en partie, le présent document. Le lecteur qui désire imprimer ou conserver sur un autre media une partie importante de ce document, doit obligatoirement en demander l'autorisation à l'auteur.

PRÉSENTATION DU JURY

CE RAPPORT DE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Lokman Sboui, directeur de mémoire
Département de génie des systèmes à l'École de technologie supérieure

M. Amin Chaabane, codirecteur de mémoire
Département de génie des systèmes à l'École de technologie supérieure

M. Julio Cesar Montecinos, président du jury
Département de génie des systèmes à l'École de technologie supérieure

Mme Bassant Selim, membre du jury
Département de génie des systèmes à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 25 JUILLET 2025

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Il me fait énormément plaisir d'adresser mes plus vifs remerciements à l'endroit de tous ceux qui ont contribué à la réalisation de ce projet de recherche de loin ou de près, en nous accordant leurs disponibilités, leurs expériences et leurs encouragements. J'adresse tout particulièrement ma profonde gratitude respectivement envers mon directeur et mon codirecteur de mémoire, les professeurs M. Lokman Sboui et M. Amin Chaabane pour leur accompagnement, ainsi que pour toutes leurs contributions, essentielles à la qualité de ce travail, leurs conseils éclairés tout au long de mon cursus et l'opportunité qui m'est eu offert de pouvoir contribuer aux travaux de recherche sur un sujet d'actualité qui me passionne et qui cadre avec mes aspirations professionnelles.

Merci aussi au président du jury, M. Julio Cesar Montecinos, et au membre du jury, Mme Bassant Selim, d'avoir disposé du temps pour lire et corriger mon travail, ainsi que pour leurs commentaires et suggestions constructifs.

Je tiens également à remercier mon conjoint, Frank Mougoue, pour ses encouragements, son soutien, son immense patience et sa présence bienveillante qui m'ont donné la force de relever chaque défi et d'aller toujours plus loin. Sans toi, ce projet n'aurait pas été possible.

Je termine à exprimer ma profonde gratitude envers ma famille, qui a toujours été ma plus grande source de force et d'inspiration. Même éloignée physiquement, elle reste toujours présente dans mon cœur. À ma mère, Toukam Clariste, dont la sagesse, l'amour inébranlable et les valeurs qu'elle m'a transmises ont été des repères essentiels tout au long de mon parcours. À ma sœur, Hariane, dont la présence constante et le soutien indéfectible ont été pour moi une source précieuse de réconfort et de motivation. Merci d'avoir été là à chaque étape, me rappelant, par votre exemple, la valeur de la persévérance et du courage. Je vous en suis profondément reconnaissante.

Modélisation énergétique des navires auto-déchargeurs à l'aide des méthodes d'apprentissage automatique

Ginette Brenda NOUDJI LAPE

RÉSUMÉ

La consommation d'énergie lors des opérations de déchargement représente un enjeu central dans le fonctionnement des navires auto-déchargeurs, en particulier en raison du rôle important des convoyeurs à bande dans la chaîne de manutention. Ces équipements, souvent sollicités en continu sur de longues périodes, sont responsables d'une consommation excessive d'énergie qui nécessite une amélioration. Pour relever ce défi, les recherches des dernières décennies se sont principalement concentrées sur le développement de modèles d'économie basé sur les modélisations physiques ou l'utilisation des algorithmes intelligents à travers des simulations faites sur des bancs d'essai.

Dans un contexte où la réduction de la consommation énergétique constitue un enjeu stratégique tant sur le plan environnemental qu'opérationnel, ce mémoire s'intéresse à l'optimisation de la consommation électrique des convoyeurs à bande. Plus précisément, l'étude propose une double approche basée sur l'industrie 4.0 en utilisant des modèles d'apprentissage automatique, tel que les forêts aléatoires, pour prédire la consommation énergétique à partir de variables opérationnelles comme le débit de déchargement, la vitesse des moteurs, le type de cargaison ou encore les caractéristiques du port. Ces modèles sont entraînés sur des données réelles provenant de plusieurs navires, ce qui permet de capturer la complexité et la variabilité des profils de consommation observées en situation réelle. En complément, des algorithmes d'optimisation numérique, dont la méthode L-BFGS-B, sont mobilisés pour identifier les configurations de fonctionnement minimisant la consommation énergétique, tout en intégrant des contraintes pratiques, telles que la durée de déchargement.

Les résultats montrent que cette double approche permet d'atteindre des économies d'énergie significatives, en particulier pour certains navires disposant de marges de manœuvre techniques suffisantes. Elle permet d'envisager des configurations de fonctionnement plus réalistes sans compromettre l'efficacité logistique et offre également un cadre décisionnel robuste pour adapter les consignes opérationnelles en fonction du type de cargaison et du contexte portuaire. Ce travail démontre ainsi l'importance de l'intelligence artificielle appliquée à la logistique maritime en vue de renforcer la durabilité des opérations tout en optimisant les performances des équipements.

Mots clés : Prédiction, Modélisation, Convoyeur à bande, apprentissage automatique, consommation d'énergie, Industrie 4.0.

Energy consumption modeling of self-unloading ships using machine learning methods

Ginette Brenda NOUDJI LAPE

ABSTRACT

Energy consumption during unloading operations represents a central issue in the operation of self-unloading vessels, particularly due to the significant role of belt conveyors in the handling chain. This equipment, often used continuously for extended periods, is responsible for excessive energy consumption that requires improvement. To address this challenge, research in recent decades has focused primarily on the development of energy-saving models based on physical modeling or the use of intelligent algorithms through simulations performed on test benches.

In a context where reducing energy consumption is a strategic issue from both an environmental and operational perspective, this thesis focuses on optimizing the power consumption of belt conveyors. More specifically, the study proposes a dual approach based on Industry 4.0 using machine learning models, such as random forests, to predict energy consumption based on operational variables such as unloading rate, engine speed, cargo type, and port characteristics. These models are trained on real-world data from several vessels, capturing the complexity and variability of consumption profiles observed in real-world situations. In addition, numerical optimization algorithms, including the L-BFGS-B method, are used to identify operating configurations that minimize energy consumption while incorporating practical constraints such as unloading time.

The results show that this dual approach can achieve significant energy savings, particularly for certain vessels with sufficient technical flexibility. It allows for more realistic operating configurations without compromising logistical efficiency and provides a robust decision-making framework for adapting operational instructions based on cargo type and port context. This work thus demonstrates the importance of artificial intelligence applied to maritime logistics to enhance operational sustainability while optimizing equipment performance.

Keywords: Prediction, modeling, Conveyor Belt, Machine Learning, Energy Consumption, Industry 4.0.

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 CONCEPT DE BASE ET REVUE DE LITTÉRATURE	5
1.1 Introduction.....	5
1.2 Concept de base	5
1.2.1 Introduction aux navires auto-déchargeurs : convoyeurs à bandes	5
1.2.2 Fonctionnement des navires auto-déchargeurs	6
1.2.3 Industrie 4.0	8
1.2.4 Optimisation de la consommation de l'énergétique des convoyeurs.....	12
1.2.5 Apprentissage automatique (Machine Learning).....	13
1.3 Revue de littérature	15
1.3.1 Modèles de prédiction de la consommation de l'énergie.....	15
1.3.2 Techniques d'optimisation énergétique.....	17
1.3.3 Études de cas et recherches existantes	19
1.4 Limitations et défis potentiels.....	21
1.5 Objectifs de l'étude	22
1.6 Contribution et discussion.....	23
1.7 Conclusion	25
CHAPITRE 2 MÉTHODOLOGIE DE RECHERCHE	27
2.1 Introduction.....	27
2.2 Approche méthodologique adoptée	27
2.2.1 Contexte industriel.....	27
2.2.2 Justification de l'approche par apprentissage automatique	28
2.2.3 Stratégie d'optimisation de la consommation énergétique	29
2.3 Base de données	29
2.3.1 Description générale du jeu de données et source d'acquisition	29
2.3.2 Structure du jeu de données	30
2.3.3 Objectif de l'analyse des données	32
2.3.4 Prétraitement des données.....	32
2.3.5 Gestion des valeurs aberrantes.....	33
2.3.6 Séparation des données	35
2.4 Évaluation des performances du modèle	36
2.5 Environnement de développement.....	38
2.6 Stratégie de développement	40
2.7 Conclusion	42
CHAPITRE 3 CONCEPTION DU MODÈLE DE PRÉDICTION ET OPTIMISATION	43
3.1 Introduction.....	43
3.2 Développement du modèle de prédiction avec le modèle de forêt aléatoire	43

3.2.1	Conception de l'algorithme de prédiction.....	45
3.2.2	Hyperparamètres	46
3.2.3	Validation croisée	48
3.2.4	Entraînement du modèle	49
3.3	Intégration de l'algorithme d'optimisation	50
3.3.1	Sélection des variables d'optimisation explicatives	50
3.3.2	Méthode L-BFGS-B utilisée pour l'optimisation.	51
3.3.3	Formulation mathématique du problème d'optimisation.....	52
3.4	Conclusion	54
CHAPITRE 4 RÉSULTATS ET DISCUSSION		55
4.1	Introduction.....	55
4.2	Résultat des analyses exploratoires des données	56
4.2.1	Gestion des valeurs manquantes	56
4.2.2	Traitements des valeurs aberrantes.....	59
4.2.3	Analyse des relations entre les variables	63
4.3	Résultat du modèle de prédiction forêt aléatoire	67
4.3.1	Prédiction globale	68
4.3.2	Prédiction selon le type de cargo	69
4.3.3	Limite observée Cas du cargo ORE du navire 3	73
4.4	Résultat de l'algorithme d'optimisation.....	73
4.4.1	Recherche exhaustive.....	74
4.4.2	Optimisation avec L-BFGS-B.....	75
4.4.3	Analyse comparative des 2 méthodes	78
4.5	Outil de recommandation.....	83
4.6	Discussion	84
4.7	Conclusion	86
CONCLUSION ET PERSPECTIVE		89
ANNEXE I	MATRICE DE CORRÉLATION	93
ANNEXE II	DISTRIBUTION DE LA CONSOMMATION D'ÉLECTRICITÉ	95
ANNEXE III	VALIDATION CROISÉE	97
LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES		99

LISTE DES TABLEAUX

	Page	
Tableau 2.1	Caractéristiques des variables du jeu de données	31
Tableau 3.1	Hyperparamètres du modèle	48
Tableau 4.1	Métriques de performance pour les prédictions globales	68
Tableau 4.2	Résultat du modèle de forêt aléatoire pour le navire 1	70
Tableau 4.3	Résultat du modèle de forêt aléatoire pour le navire 2	70
Tableau 4.4	Résultat du modèle de forêt aléatoire pour le navire 3	71
Tableau 4.5	Résultat de la recherche exhaustive pour le navire 1	74
Tableau 4.6	Résultat de la recherche exhaustive pour le navire 2	74
Tableau 4.7	Résultat de la recherche exhaustive pour le navire 3	75
Tableau 4.8	Optimisation du navire 2 avec L-BFGS-B.....	76
Tableau 4.9	Optimisation du navire 3 avec L-BFGS-B.....	76
Tableau 4.10	Optimisation du navire 1 avec L-BFGS-B.....	77

LISTE DES FIGURES

	Page	
Figure 1.1	Illustration d'un convoyeur à bande	6
Figure 1.2	Fonctionnement de l'auto-déchargeur	7
Figure 1.3	Les quatre révolutions industrielles.	10
Figure 1.4	L'écosystème de l'Industrie 4.0.....	11
Figure 2.1	Visualisation de la méthode IQR	35
Figure 2.2	Séparation du jeu de données.....	36
Figure 2.3	Étapes de développement de notre solution.....	41
Figure 3.1	Étapes de prédiction de la consommation de l'énergie.....	45
Figure 4.1	Aperçu global du jeu de données du navire 1	56
Figure 4.2	Aperçu global du jeu de données du navire 3	57
Figure 4.3	Distribution de la consommation d'électricité par rapport au taux de décharge du navire 3	58
Figure 4.4	Variables du navire 4	59
Figure 4.5	Détection des valeurs aberrantes par la méthode de l'écart interquartile (IQR)	60
Figure 4.6	Comparaison des données après nettoyage.....	61
Figure 4.7	Vérification des données après traitement des valeurs aberrantes	62
Figure 4.8	Matrice de corrélation du navire 2	64
Figure 4.9	Diagramme de dispersion nettoyé de la consommation d'électricité en	65
Figure 4.10	Distribution de la consommation d'électricité	66
Figure 4.11	Comparaison de la consommation d'énergie selon le port.....	67
Figure 4.12	Comparaison des diagrammes de la consommations globale.....	69

Figure 4.13	Visualisation comparative de la prédition pour le cargo GYPSUM du navire 2.....	72
Figure 4.14	Visualisation comparative de la prédition pour le cargo SALT du navire 3	72
Figure 4.15	Visualisation comparative de la prédition du cargo ORE du navire 3	73
Figure 4.16	Consommation optimale du navire 1	78
Figure 4.17	Consommation optimale du navire 2	79
Figure 4.18	Consommation optimale du navire 3	80
Figure 4.19	Productivités optimales du navire 1	81
Figure 4.20	Productivités optimales du navire 2	81
Figure 4.21	Productivités optimales du navire 3	82
Figure 4.22	Tableau de bord itératif.....	84

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

API	Automate Programmable Industriel
BD	Base de Données
CEMA	Conveyor Equipment Manufacturers Association
CNN	Réseaux neuronaux convolutifs (Convolutional Neural Network)
CSV	Comma-Separated Values
DB	Double fond (Double Button)
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DIN	Deutsches Institut für Normung
DL	Apprentissage profond (Deep Learning)
DNN	Réseau neuronal profond (Deep Neural Network)
EAD	Analyse Exploratoire des Données
EM	Expectation-Maximization
IA	Intelligence Artificielle
IdO	Internet des Objets
ISO	Organisation de normalisation internationale (International Organization for Standardization)
JISB	Japanese Industrial Standards – Section B
KNN	K-Nearest Neighbors
L-BFGS-B	Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Bounds
LDA	Analyse discriminante linéaire (Linear Discriminant Analysis)
LSQ	Least Squares – méthode des moindres carrés
MDP	Processus de décision de Markov
ML	Apprentissage automatique (Machine Learning)
NB	Naïfs bayésiens
NIST	National Institute of Standards and Technology
PCA	Analyse en composantes principales (Principal Component Analysis)
PLC	Programmable Logic Controller
RL	Apprentissage par renforcement (Reinforcement Learning)
RLSQ	Regularized Least Squares – moindres carrés régularisés

RNN	Réseaux neuronaux récurrents (Recurrent Neural Network)
SUL	Navires auto-déchargeur (Self-Unloading ships)
SVM	Machines à vecteurs de support (Support Vector Machine)
TOU	Time Of Use
VSD	Variateurs de vitesse
WIFI	Wireless Fidelity

LISTE DES SYMBOLES ET UNITÉS DE MESURE

E	Consommation d'énergie	kWh.
IQR	Écart interquartile	
MAE	Erreure absolue moyenne	%
MAPE	Erreure absolue moyenne en pourcentage	
MSE	Erreure quadratique moyenne	
P	Puissance du moteur	kW
Rcarre	Coefficient de détermination	
RMSE	Erreure quadratique moyenne racine	
T	Température du moteur	°C
V	Vitesse	tr/min
t	Temps de déchargement	heure (h)
τ	Taux de déchargement	tonnes/heure
α	Angle d'inclinaison	degrés

INTRODUCTION

La manutention mécanique est une opération qui consiste à déplacer, transporter ou manipuler des marchandises à l'aide d'un appareil mécanique. Elle est dominante dans la plupart des domaines, comme la filière industrielle, chimique, agroalimentaire, minière, etc. Le choix du système de manutention à utiliser dans chaque application dépend des caractéristiques du produit à manipuler et des conditions de transport. Parmi les systèmes de manutention les plus utilisés, on trouve les convoyeurs à chaîne, à rouleaux, à bande, à courroies, à vis sans fin et extensibles. Chaque domaine possède ses propres besoins qui favorisent l'utilisation d'un type de systèmes par rapport aux autres. Pour le transport des matériaux en vrac, le convoyeur à bande représente le moyen mécanique le plus rentable et le plus efficace grâce à sa flexibilité et sa capacité de transporter en continu divers types de matériaux avec des débits importants et même dans des conditions difficiles. Ce qui conduit à le privilégier lorsqu'il s'agit de transporter des produits comme le phosphate, le ciment, les sels, les engrains, les agrégats, le charbon, les céréales et le minerai.

Dans un contexte où l'efficacité énergétique est devenue un enjeu central, notamment face à la hausse des coûts de l'énergie et aux objectifs de réduction des émissions de gaz à effet de serre, la maîtrise de la consommation des équipements industriels est stratégique. Les convoyeurs à bande, responsables d'une part importante de la consommation électrique à bord des navires, représentent un levier d'optimisation encore peu exploité de façon intelligente. La présente recherche se concentre donc sur l'optimisation de la consommation énergétique de ces convoyeurs, dans le contexte précis des navires auto-déchargeurs, à l'aide de techniques d'apprentissage automatique et de méthode d'optimisation.

Cette étude se limite aux systèmes de convoyeurs embarqués sur des navires SUL (*Self unloader ship*) qui représentent les navires auto-déchargeurs gravitaires opérant dans des environnements portuaires semi-industriels, avec un accent mis sur l'analyse des performances en conditions réelles de déchargement.

Bien que les approches physiques et les équations fondamentales existent pour modéliser la consommation de l'énergie dans ce type de système, la littérature présente certaines lacunes, notamment l'absence de modèles prédictifs spécifiques aux systèmes embarqués sur navires auto-déchargeurs, intégrant des contraintes opérationnelles réelles de fonctionnement et la faible utilisation de l'apprentissage automatique pour l'optimisation dans le contexte portuaire. En effet, la majorité des travaux antérieurs se concentrent plus sur l'amélioration de l'efficacité des équipements qui est effectuée soit en introduisant des équipements très efficaces, soit en améliorant l'efficacité des équipements existants. En pratique, les informations incomplètes et l'incertitude des paramètres acquises par les anciens équipements dans des environnements complexes augmentent la difficulté d'utiliser ces méthodes pour prédire la consommation d'énergie des systèmes de convoyeurs à bande. De plus, la connaissance de la prédiction de la consommation d'énergie pour les systèmes de convoyeurs à bande est un indicateur important qui a un effet sur leurs paramètres de fonctionnement.

Pour pallier ces difficultés, une question de recherche se pose: comment prédire et optimiser de manière fiable la consommation énergétique des convoyeurs à bande, en tenant compte des contraintes opérationnelles réelles de fonctionnement ? Répondre à cette question suscite également plusieurs autres interrogations. D'une part quel pourrait être le modèle prédictif basé sur l'apprentissage automatique qui permettra de reproduire fidèlement la consommation mesurée ? D'autre part, dans un environnement portuaire, quels seraient les paramètres à ajuster pour réduire la consommation d'énergie sans compromettre l'efficacité du système ? En outre, comment garantir la compatibilité des solutions proposées avec les exigences de temps et de productivité imposées par l'activité ?

Afin de répondre à cette problématique, nous avons un objectif général qui est celui de proposer une méthode basée sur l'apprentissage automatique qui permet d'optimiser la consommation de l'énergie des convoyeurs à bandes dans un contexte portuaire, afin de fournir des recommandations aux opérateurs pour effectuer des déchargements efficaces. Pour y parvenir, nous devrons au préalable :

- Identifier et analyser les facteurs techniques et environnementaux influençant la consommation énergétique des convoyeurs embarqués à l'aide de l'analyse exploratoire des données.
- Développer un modèle prédictif fiable de consommation énergétique à l'aide d'un algorithme d'apprentissage automatique adapté aux données disponibles.
- Mettre en œuvre une méthode d'optimisation L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Bound constraints) pour déterminer les paramètres de fonctionnement les plus économies en énergie, selon différents scénarios et type de cargaison.

Dans le cadre de notre recherche, nous proposons une nouvelle méthode pour prédire la consommation énergétique basée à la fois sur un modèle de forêts aléatoires et sur une méthode d'optimisation numérique (L-BFGS-B) adaptée à des contraintes industrielles spécifiques du milieu naval. L'étude utilise aussi une base de données structurées issues de mesures enregistrées par des capteurs et propose une validation croisée par cargo, rarement abordée dans la littérature. Elle se distingue par son ancrage dans une problématique opérationnelle concrète, en lien direct avec les enjeux de l'industrie maritime et de l'Industrie 4.0.

Pour parvenir aux résultats désirés, il a été question premièrement de comprendre et définir le cadre de l'étude. Cette étape commence par une revue de littérature sur les concepts d'efficacité énergétique dans les convoyeurs, d'apprentissage automatique et leurs interactions communes dans la gestion de consommation de l'énergie des convoyeurs. Ensuite nous avons fait une analyse exploratoire des données enregistrées lors du fonctionnement des navires afin d'identifier les paramètres liés à la consommation de l'énergie. Enfin des modèles de prédictions et optimisations ont été développés, plusieurs métriques de performances ont été utilisées pour valider les résultats.

Le rapport est structuré en 4 chapitres comme suit :

- Le chapitre 1 présente les fondements techniques et théoriques du sujet, avec une revue complète de la littérature.

- Le chapitre 2 détaille la méthodologie adoptée pour la modélisation et l'optimisation.
- Le chapitre 3 est consacré à la conception du modèle choisi
- Le chapitre 4 expose les résultats expérimentaux, leur validation ainsi qu'une discussion critique.

CHAPITRE 1

CONCEPT DE BASE ET REVUE DE LITTÉRATURE

1.1 Introduction

Ce chapitre s'attarde sur les recherches antérieures en lien avec la problématique étudiée. Tout d'abord, une exploration des systèmes de transport de matériaux en vrac à bord des navires auto-déchargeurs, en particulier les convoyeurs à bande est faite. Ensuite les enjeux liés à la consommation énergétique sont étudiés en détaillant les diverses stratégies d'optimisation et les outils méthodologiques observés dans les études existantes.

1.2 Concept de base

1.2.1 Introduction aux navires auto-déchargeurs : convoyeurs à bandes

Les historiens débattent de la date d'apparition des premiers convoyeurs à bande. Nombreux sont ceux qui pensent que les premiers convoyeurs à bande manuels en bois et cuir sont apparus à la fin du XVIII^e siècle. Le premier convoyeur à bande robuste est apparu en 1892 pour transporter du charbon, du minerai et des produits similaires. Les historiens attribuent leur invention à Thomas Robins, qui a développé une série de modèles qui ont finalement donné naissance au convoyeur à bande robuste (625160587-Belt-Conveyor). La technologie SUL remonte à 1908, dans les Grands Lacs (Group, 2025), où il était initialement destiné à transporter le minerai de fer, le charbon et les céréales de la zone minière et des champs vers les installations de réception et les moulins. Il existe de nombreux types de vaisseaux SUL, mais cette étude se limitera au type SUL gravitaire. Les particularités d'un convoyeur à bande sont sa facilité et son faible coût d'entretien, sa capacité de chargement et de déchargement élevée et sa capacité à transporter des matériaux denses de manière économique et très efficace sur de longues distances, permettant ainsi un mouvement relatif des matériaux.



Figure 1.1 Illustration d'un convoyeur à bande

Tirée de VALPROD (2025)

Le convoyeur à bande peut également être utilisé pour divers matériaux : abrasifs, humides, secs, collants ou sales. Les équipements de manutention tels que les convoyeurs à bande sont conçus pour charger et décharger les matériaux d'une étape de traitement à une autre de la manière la plus rapide, la plus fluide, la plus judicieuse, la plus sûre et la plus économique, avec un minimum de pertes.

Les ports d'aujourd'hui disposant des infrastructures modernes ne représentent plus juste des points de passage, mais de véritables carrefours stratégiques de la logistique mondiale où chaque instant gagné ou perdu peut avoir un impact significatif sur l'ensemble de la chaîne d'approvisionnement. Le déchargement des navires ne se situe plus comme une simple opération technique, en revanche, il constitue un point central sur lequel une attention particulière doit être apportée permettant la modernisation et la transformation des opérations portuaires. La rapidité et la précision de cette étape influencent directement la fluidité des échanges maritimes et la performance économique du secteur portuaire tout entier.

1.2.2 Fonctionnement des navires auto-déchargeurs

Le système mis à l'étude est les vraquiers auto-déchargeurs originaires des Grands Lacs. C'est un système de convoyeur qui a été installé dans les vraquiers de la taille d'un Lakers pour livrer dans les ports où les installations de manutention en vrac sont limitées ou inexistantes. Ces

navires transportaient du matériel entre les ports de la Voie maritime du Saint-Laurent et les Grands Lacs ; servant ainsi de prolongement intérieur aux principales routes océaniques à partir du Canada. Ils ne nécessitent pas de systèmes de décharge à terre ni d'infrastructure supplémentaire. De 0 à 70 mètres de quai, les auto-déchargeurs peuvent décharger directement dans une trémie, une barge, un quai, un entrepôt ou un magasin. Les auto-déchargeurs peuvent aussi effectuer des opérations de transbordement, de remplissage ou de déchargement sur des navires plus grands.

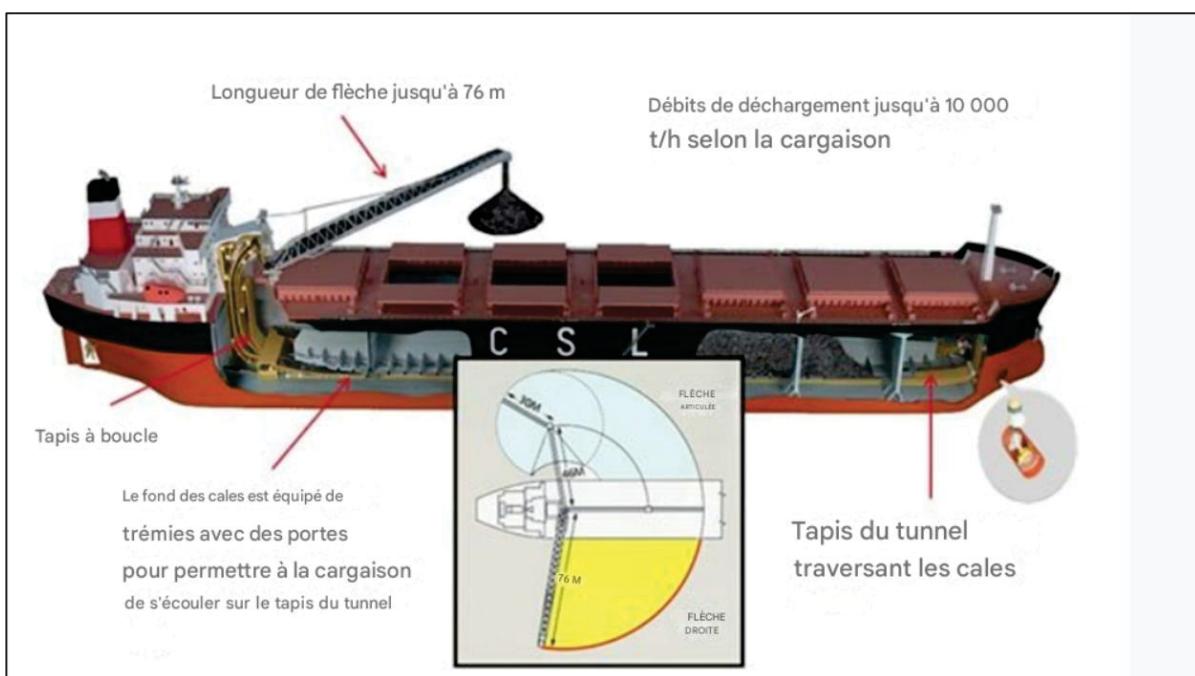


Figure 1.2 Fonctionnement de l'auto-déchargeur

Adaptée de Smith (2014)

En raison du manque d'installations de manutention en vrac dans de nombreux ports du monde, cette application a été introduite dans le commerce des eaux profondes. La principale caractéristique distinctive de ce type de navire est leur système de convoyage. Un système et une structure de convoyage de type essentiellement industriel, conçus pour une utilisation à bord, ont donné une identité spécifique à ce type de vraquier généralement appelé Self-Unloaders. Dans cette conception, les cales du navire se terminent au-dessus du sommet du

réservoir DB, formant une série de nombreuses formes coniques ouvertes appelées trémies. Ces ouvertures de trémie sont recouvertes de vannes à charnières ou d'écluses, qui sont actionnées hydrauliquement. L'espace situé entre le bas des trémies et le haut du réservoir DB, qui s'étend sur toute la longueur du navire, est appelé le tunnel. Un tapis roulant y circule pour transporter la marchandise, qui tombe grâce à la gravité en passant par des portes. Le débit est contrôlé en fonction de la quantité de marchandise présente sur la bande. Ensuite, la cargaison est remontée soit à l'aide d'ascenseurs en boucle, soit par des pentes inclinées, puis elle est acheminée vers l'estacade du navire.

Les vraquiers SUL sont autonomes, ce qui veut dire qu'ils n'ont pas besoin de machines à terre pour se décharger. Ils sont également équipés d'un convoyeur à flèche qui peut déposer la cargaison directement à terre, que ce soit sur un quai, dans une trémie, sur une barge, ou même dans un autre navire. Cette flèche est très mobile et peut bouger dans plusieurs directions, ce qui la rend très flexible.

Le convoyeur est un dispositif de conversion d'énergie typique, qui convertit l'électricité en énergie mécanique. L'optimisation vise généralement à réduire la consommation énergétique, à maximiser le débit ou à préserver la durée de vie des équipements. L'économie d'énergie sur un convoyeur à bande peut être réalisée à quatre niveaux : efficacité, fonctionnement, équipement et technologie.

1.2.3 Industrie 4.0

L'industrie 4.0 permet aux entreprises de gérer plus efficacement des processus complexes en fournissant des capacités sous forme de services encapsulés (Hermann et Al, 2020). Dans le contexte où l'optimisation des opérations de déchargement devient une préoccupation importante, l'industrie 4.0 s'impose comme une solution efficace et incontournable en intégrant les technologies qu'elle englobe. Elle est définie de manière polyvalente dans la littérature et fait son apparition pour la première fois lors de la foire de Hanovre en 2011. Elle sera par la suite formellement annoncée en 2013 comme une orientation stratégique visant à jouer un rôle de précurseur dans les industries qui révolutionnent actuellement le secteur manufacturier.

L'expression "Industrie 4.0" fait référence à la quatrième révolution industrielle qu'a connue le monde jusqu'aujourd'hui. Elle repose sur une nouvelle conception des systèmes de production intégrant les technologies numériques, l'intelligence artificielle, l'Internet des objets et la connectivité en temps réel. Cette révolution s'inscrit dans la continuité de trois grandes étapes historiques. La première révolution industrielle apparue à la fin du 18^e siècle avec l'introduction de la machine à vapeur comme moyen de production mécanique. Par la suite nous observons une deuxième révolution industrielle marquée par l'électrification et la production en masse jusqu'au début du 20^e siècle. À la fin du 20^e siècle vers les années 1970, une troisième révolution industrielle fait son apparition, encore appelée "révolution numérique". Celle-ci est particulièrement observée par l'apparition des progrès électroniques et de l'informatique grâce aux automates programmables industriels (API). La production devient donc automatisée et soutenue par les technologies informatiques. L'Industrie 4.0 prolonge cette évolution en connectant intelligemment les machines, les systèmes et les données pour créer des usines plus flexibles, réactives et efficientes.

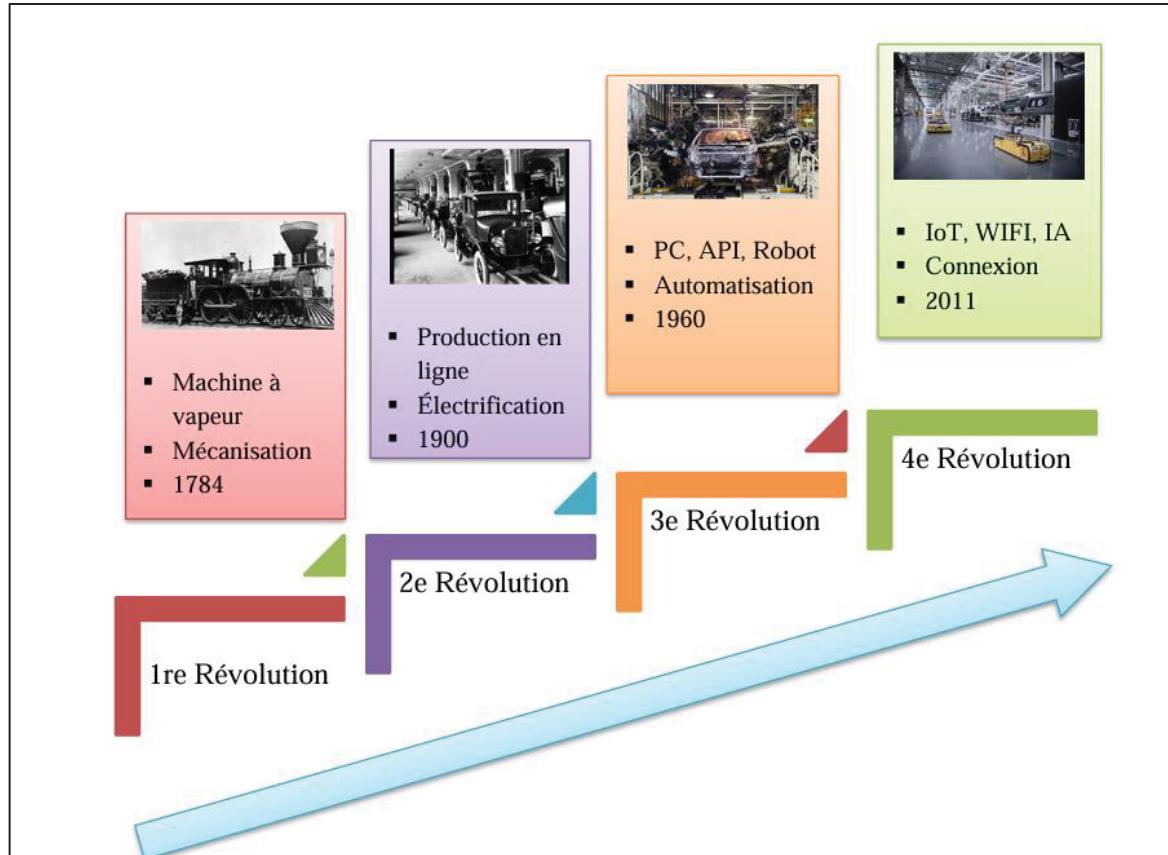


Figure 1.3 Les quatre révolutions industrielles.

Tirée de HADDOUCHE (2023)

Elle se caractérise par la transformation globale et l'intégration des nouvelles technologies numériques dans le processus de production. Son principe repose sur la connectivité intelligente des machines, des systèmes et des objets. Plusieurs technologies accompagnent cette révolution, parmi lesquelles nous pouvons citer l'Internet des objets (IoT), l'intelligence artificielle (IA), les mégadonnées, l'impression 3D, les jumeaux numériques, la robotique collaborative, ainsi que l'usage croissant de l'informatique en nuage (cloud computing).

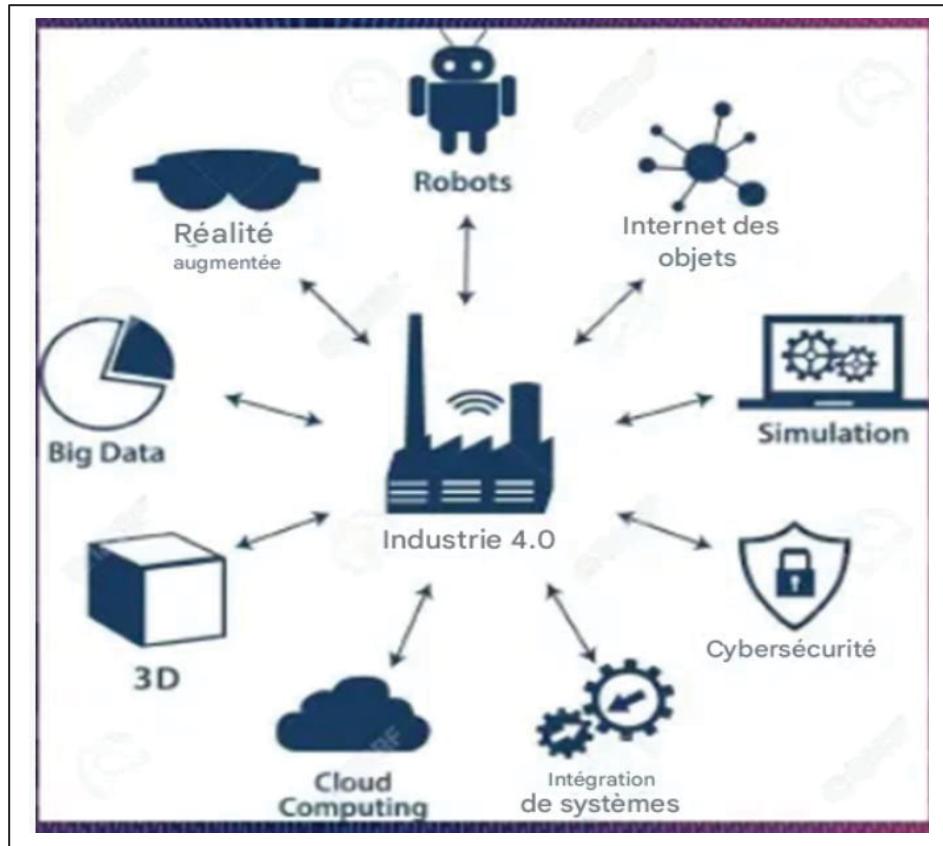


Figure 1.4 L'écosystème de l'Industrie 4.0

Adaptée de Królikowski, T. (2022)

Internet des objets (IdO) ou encore dans sa version anglaise « Internet of Things (IoT) » : il fait son apparition en Amérique du Nord grâce à la société General Electric vers fin 2012. Ce concept intègre étroitement les objets connectés via l'analyse des mégadonnées (big data) en assurant une convergence entre le monde physique et celui du numérique. Il couvre un champ d'application significativement important dans l'industrie 4.0 permettant une exploitation intelligente des informations issues du terrain. Son application trouve des intérêts dans tous les domaines, comme la production industrielle, la santé, le secteur public, le domaine du transport et même celui de l'exploitation d'industrie minière. Son apport dans l'industrie 4.0 est phénoménal, notamment à travers l'utilisation extensive d'internet, non seulement comme un canal à faible coût pour connecter des machines, des dispositifs, des capteurs et des personnes, mais aussi facilite la création des nouvelles fonctions et offre la possibilité d'améliorer des

caractéristiques de produits liées à la capacité d'utiliser l'internet comme source d'information. De manière plus simple, l'IdO est une interface de données disponibles qui permet à différents logiciels ou applications de partager ou échanger les informations entre eux grâce à des protocoles de communications standards.

Mégadonnées (Big Data en anglais) : On parle de mégadonnées lorsque les volumes, la vitesse et la variété des données dépassent les capacités traditionnelles de traitement. Elle est obtenue en collectant de grandes quantités de données issues de plusieurs sources, comme les sites web, les réseaux sociaux et généralement à l'aide l'IdO. Dans les systèmes industriels, tels que les convoyeurs à bande embarqués sur les navires auto-déchargeurs, les capteurs installés génèrent en continu des données massives : vitesses de moteurs, températures, taux de charge, vibrations, inclinaison des bras de déchargement, ou encore données environnementales. La collecte et l'analyse des données aux jours d'aujourd'hui s'inscrivent comme une norme pour appuyer la prise de décision en temps réel.

L'infonuagique (Cloud computing) : est une technologie qui permet d'utiliser des services informatiques comme le stockage des fichiers, des logiciels, des réseaux, des bases de données à travers internet sans avoir besoin de tout installer sur un ordinateur ou un serveur. Il permet d'éviter le coût d'investissement lié à l'achat de logiciels et matériels.

1.2.4 Optimisation de la consommation de l'énergétique des convoyeurs

Aujourd'hui, limiter ou réduire la consommation énergétique dans les systèmes de fabrication de l'industrie 4.0 constitue un enjeu majeur pour des entreprises industrielles. Dans les industries fortement automatisées, comme les systèmes de convoyages, la consommation énergétique constitue une part significative des coûts de production, ce qui rend sa gestion en temps réel indispensable pour atteindre une performance énergétique optimale. Toutefois, modéliser précisément cette consommation reste complexe. Ainsi, l'Industrie 4.0 permet une analyse fine des flux énergétiques à travers des capteurs intelligents et des systèmes de surveillance connectés. Ces systèmes collectent en continu des données, telles que la

température, la vitesse, la quantité de charges transportée, etc., qui sont ensuite exploitées par des méthodes d'IA pour détecter les anomalies, ajuster automatiquement les paramètres de fonctionnement, ou proposer des stratégies d'optimisation. Par exemple, les réseaux neuronaux profonds (DNN) permettent d'exploiter les mégadonnées issues de la surveillance énergétique pour identifier des schémas de consommation et adapter, sans intervention humaine, le comportement des équipements en vue de réduire leur impact énergétique.

Se doter d'un système industriel efficace du point de vue énergétique en utilisant la technologie de l'industrie 4.0 consiste tout d'abord par l'intégration d'équipements adaptés qui permettent une gestion optimisée de la consommation. Cela implique notamment l'usage de capteurs, d'actionneurs et de machines à faible consommation, interconnectés via des réseaux sans fil industriels. Dans ce contexte la technologie dite « intelligente » remplace progressivement les systèmes traditionnels, qu'elles soient analogiques ou numériques. Par exemple, à la place des capteurs classiques, on aura des capteurs intelligents capables non seulement de collecter des données, mais aussi d'apprendre du processus, de mémoriser des informations et de s'organiser automatiquement. Conçus comme des objets connectés (IoT), ils transmettent des données numériques en temps réel vers des passerelles via des réseaux sans fil et peuvent ainsi analyser le comportement du système, s'adapter à l'environnement et prendre des décisions localement. La communication entre les dispositifs peut se faire soit par réseau WIFI industriel ou par une communication RFID.

1.2.5 Apprentissage automatique (Machine Learning)

L'apprentissage automatique est une branche de l'intelligence artificielle en constante évolution qui permet de reproduire certains aspects de l'intelligence humaine en apprenant à partir de l'environnement et des données disponibles. L'on retrouve leurs applications dans divers domaines, comme la reconnaissance de formes, de la vision par ordinateur, de l'ingénierie aérospatiale, de la finance, du divertissement, etc. Cependant il existe 3 catégories principales de l'apprentissage selon la manière, donc les données sont enregistrées. Nous avons :

Apprentissage supervisé : Dans le ML supervisé, les algorithmes sont entraînés à partir de jeux de données étiquetées, c'est-à-dire que chaque donnée d'entrée est associée à une réponse connue (ou label). Autrement dit, l'algorithme reçoit des exemples déjà interprétés, ce qui lui permet d'apprendre à faire des prédictions précises. Par exemple, un algorithme peut entraîner un modèle avec des images de fleurs, chacune étant accompagnée du nom de l'espèce correspondante. Grâce à ces étiquettes, le modèle apprend à reconnaître les caractéristiques de chaque type de fleur et devient capable d'identifier correctement une fleur lorsqu'une nouvelle image lui est présentée. Le ML supervisé est souvent utilisé pour créer des modèles pour prédiction (régression) et classification. Ils comprennent principalement des modèles linéaires, des k plus proches voisins (KNN), des machines à vecteurs de support (SVM), des modèles naïfs bayésiens (NB), des modèles basés sur des arbres de décision et forêt aléatoire.

Apprentissage non-supervisé : Ce type d'apprentissage utilise des ensembles de données non étiquetées pour entraîner les algorithmes. Ici, l'algorithme est entraîné par des données qui ne comportent pas d'étiquettes : on lui fournit des données sans réponse, ce qui l'oblige à découvrir des modèles par lui-même, sans aucune aide extérieure. Le regroupement (clustering) et la réduction de la dimensionnalité sont des modèles communs d'apprentissage non-supervisé. Ils peuvent être divisés en algorithmes de clustering et de réduction de la dimensionnalité en fonction de leurs objectifs. Les méthodes de clustering comprennent principalement les k-moyennes, le clustering spatial basé sur la densité des applications avec bruit (DBSCAN), le clustering hiérarchique et l'anticipation-maximisation (EM). De plus, il existe plusieurs méthodes d'apprentissage d'ensemble qui combinent différents modèles singuliers pour améliorer encore les performances des modèles, comme le vote, l'ensachage et AdaBoost.

L'apprentissage profond (DL) est une technique de ML importante qui continue à faire ses preuves ces derniers temps. C'est une forme d'apprentissage automatique qui permet aux ordinateurs d'apprendre en interagissant avec un environnement de l'expérience ou en construisant progressivement des représentations de plus en plus abstraites à partir des données. Les algorithmes ML traditionnels, quant à eux, sont limités dans leurs applications

pour le traitement de représentations d'ensembles de données complexes, car ils dépendent d'une extraction de caractéristiques adéquate pour apprendre et effectuer diverses tâches. Il est intéressant de noter que les algorithmes DL, inspirés par l'imitation du système nerveux biologique, peuvent effectuer l'extraction et la transformation de caractéristiques à partir de données non étiquetées ou non structurées(C. Zhang, Patras, & Haddadi, 2019). Les modèles DL comprennent plusieurs couches interconnectées de neurones, chaque couche utilisant la sortie de la couche précédente comme entrée. De vastes applications des algorithmes DL ont été enregistrées dans différents domaines, tels que la reconnaissance de formes, le traitement du langage naturel, la vision par ordinateur, la régression, etc. Les modèles DL populaires étendus aux réseaux IoT comprennent les réseaux neuronaux convolutifs (CNN), les réseaux neuronaux récurrents (RNN), les réseaux de croyances profondes et les auto-encodeurs.

Apprentissage par renforcement (RL) : Contrairement à l'apprentissage supervisé et à l'apprentissage non supervisé, RL est conçu pour permettre à plusieurs agents (par exemple, des appareils IoT) d'apprendre en interagissant avec leur environnement et est entraîné en fonction de la récompense (positive ou négative) obtenue par une approche d'essai et d'erreur (Alamu, Olwal, & Migabo, 2025). Les algorithmes d'apprentissage par renforcement visent à apprendre la politique optimale c'est-à-dire la cartographie optimale des situations en actions afin de maximiser la récompense cumulative au fil du temps. Ceci est réalisé en trouvant un équilibre approprié entre l'exploration de nouvelles actions et l'exploitation d'actions connues (Alamu et al., 2025). Parce que les algorithmes RL peuvent échanger des récompenses immédiates contre des gains à long terme, ils sont connus pour être adaptés à la prise de décision, à la planification et à l'ordonnancement dans des environnements de réseau stochastiques.

1.3 Revue de littérature

1.3.1 Modèles de prédiction de la consommation de l'énergie

Autrefois, le déchargement des navires était effectué de manière manuelle où l'effort physique primait sur la technologie. Mais aujourd'hui, grâce aux différentes avancées technologiques

les opérations sont de plus en plus automatisées, augmentant la rapidité et la précision sans oublier la diminution des risques pour le personnel. Les ports modernes sont devenus de véritables centres technologiques, où des logiciels intelligents contrôlent chaque mouvement : celui des grues, des conteneurs et des camions. Grâce à ces outils, les erreurs humaines sont réduites, l'attente est mieux gérée, et l'espace sur les quais est utilisé de façon plus efficace.

Cependant, ils consomment également de grandes quantités d'énergie en raison de divers impacts qui entraînent un gaspillage d'énergie. La bonne gestion de ces mécanismes nécessite une prédition précise et fiable de la courbe de consommation d'énergie. Quantifier la demande d'énergie à l'intérieur des systèmes de convoyeurs à bande est, en général, un processus complexe. Plus précisément, plusieurs facteurs influencent la consommation d'énergie d'un système de convoyeur à bande, tels que les variables de fonctionnement, la construction du système, le taux de charge de la bande et ses caractéristiques dynamiques, l'utilisation des pilotes et leurs performances, etc. Une telle prédition devra être faite en tenant compte des facteurs les plus influents sur la consommation d'énergie. De plus, la prédition de la consommation d'énergie d'un système de convoyeur à bande serait également propice à la détermination des fonctions objectives à contrôler.

De nombreuses techniques ont été appliquées à la tâche de prédition, telles que l'ingénierie, les statistiques et les méthodes d'intelligence artificielle. Les méthodes d'ingénierie utilisent des principes physiques pour calculer la consommation d'énergie du système de convoyeur à bande en prenant en compte les paramètres de structure mécanique et les variables de fonctionnement (distance de transport, le taux de transfert, vitesse de la bande). Il existe de nombreux modèles énergétiques pour les convoyeurs à bande basée sur des normes reconnues, tels que l'ISO 5048, la DIN 22101, la JISB 8805 ou encore la CEMA (Shirong & Xiaohua, 2009). Toutefois, ces modèles présentent des limites communes. En effet, ces normes servent principalement de guides pour la conception des systèmes de convoyage, mais plusieurs paramètres requis sont issus de l'expérience pratique et sans aucune grande précision. De plus ces standards ne permettent pas d'optimiser de manière inverse les performances énergétiques des systèmes une fois conçus. Or, dans le cadre de cette étude, la consommation d'énergie et

les coûts associés constituent des critères clés à optimiser. Il existe un modèle de calcul de l'énergie modifié reliant les deux catégories différentes des modèles de calcul de l'énergie existants (Zhang et Al, 2009). Bien que ces méthodes soient précises, elles nécessitent un ensemble complet et détaillé de paramètres du système de convoyeur à bande et des paramètres de fonctionnement, et ces détails ne sont pas toujours disponibles.

Les méthodes statistiques, quant à elles, tentent de trouver une corrélation entre la consommation d'énergie et les variables les plus influentes. Ces modèles empiriques sont développés à partir des données historiques qui sont utilisées pour entraîner le modèle. En pratique, les informations incomplètes et l'incertitude des paramètres acquises par les anciens équipements dans des environnements complexes augmentent la difficulté d'utiliser ces méthodes pour prédire la consommation d'énergie des systèmes de convoyeurs à bande (Zeng & Wu, 2015).

Les méthodes d'intelligence artificielle sont un axe de recherche qui a fait l'objet d'une attention croissante au cours des dernières années en raison de leur bonne adéquation à ce type de problèmes. Contrairement aux approches déterministes classiques fondées sur des lois physiques, l'IA propose une modélisation empirique, car elle apprend directement à partir des données historiques les relations entre variables techniques et consommation d'énergie. Dans les systèmes de convoyage industriel, la consommation énergétique dépend d'un ensemble de facteurs et paramètres hétérogènes qui interagissent de manière non linéaire et varient selon le contexte opérationnel. L'apprentissage automatique est particulièrement adapté à ce type de problématique, car il permet de capter des relations complexes et d'ajuster dynamiquement les prédictions en fonction des données observées. Fei ZENG et al. (2015) dans leurs travaux, ont présenté un modèle de prédiction de la consommation d'énergie d'un système de convoyeur à bande basé sur un réseau de neurones.

1.3.2 Techniques d'optimisation énergétique.

Le convoyeur à bande est un bon exemple de conversion de l'énergie électrique en énergie mécanique. Certains auteurs ont démontré que les convoyeurs à bande sont responsables de 50

à 70 % de la consommation totale d'électricité dans un terminal de vrac sec. Cependant, la majeure partie de l'électricité est encore produite aujourd'hui en brûlant des combustibles fossiles, comme le charbon et le pétrole. Le charbon demeure une source majeure d'énergie à l'échelle mondiale, assurant la production d'électricité. Toutefois, il est également à l'origine des émissions mondiales de CO₂ issues de la combustion de combustibles fossiles. Par conséquent, en tenant compte des défis économiques et sociaux pertinents, il est urgent d'améliorer l'efficacité énergétique des convoyeurs à bande et de réduire considérablement l'empreinte carbone (Mhlongo, Nnachi, Nnachi, & Adesola, 2020). Selon le ministère de l'Énergie et des Ressources naturelles du Québec, l'efficacité énergétique se définit par " Faire la meilleure utilisation possible de l'énergie disponible pour obtenir un meilleur rendement énergétique (HADDOUCHE, 2023). Elle est améliorée lorsque moins d'énergie est utilisée pour produire le même bien ou service ". L'efficacité énergétique des convoyeurs comme tout autre système industriel pourrait se faire selon quatre niveaux : efficacité, fonctionnement, équipement et technologie.

Dans la littérature, ces différentes méthodes sont étudiées de manières isolées. L'efficacité du fonctionnement et de l'équipement peut être améliorée dans la plupart des systèmes, y compris les convoyeurs à bande. L'efficacité des performances dépend de l'efficacité du fonctionnement et de l'équipement, où l'efficacité des performances est reflétée par de nombreux indicateurs externes, comme la consommation et le coût de l'énergie. Les mesures d'économie d'énergie sur les convoyeurs à bande se concentrent principalement sur l'efficacité de l'équipement. Cela a été fait soit en introduisant des équipements très efficaces, soit en améliorant l'efficacité des équipements existants, tels que la roue folle, la courroie, le système d'entraînement, etc. Il peut également être réalisé en améliorant les dispositifs d'entraînement lors de son démarrage et de son fonctionnement, tels que la technologie de démarrage progressif, les variateurs de vitesse (VSD).

Au niveau opérationnel, les économies d'énergie peuvent être réalisées par la coordination des sous-systèmes ou les opérations de partage du temps des appareils. Par exemple, nous pouvons ajuster l'état des appareils, allumé ou éteint, pour réduire l'indice de coût de la consommation

d'énergie en fonction du temps d'utilisation (TOU) de la région. Cependant, dans le processus industriel réel, nous ajustons principalement la vitesse de la bande pour correspondre à la vitesse d'alimentation du matériau afin d'atteindre le transport maximal de matériaux de masse dans une période unitaire (Chen, Zeng, Du, & Zhou, 2015).

1.3.3 Études de cas et recherches existantes

Au cours des dernières décennies, plusieurs chercheurs se sont lancés dans la recherche sur la façon d'atténuer la consommation d'énergie des convoyeurs à bande et certains résultats importants ont été obtenus. Sur la base de la norme DIN 22101, (Hiltermann et al., 2011) ont proposé une technique de calcul des économies d'énergie réalisées par le contrôle de la vitesse. La norme permet de prédire les économies réalisées sur le contrôle de la vitesse des convoyeurs à bande. Cependant, l'étalonnage est nécessaire en déduisant le coefficient de frottement applicable à partir des mesures. De même, (Mhlongo et al., 2020) propose de modéliser et de simuler le comportement dynamique et le contrôle des systèmes de bandes transporteuses pour des études d'efficacité énergétique à l'aide de Matlab/Simulink. La nouveauté de ce travail réside dans les calculs et simulations détaillés du système de convoyeur à bande en tenant compte de certaines constantes du convoyeur à bande et de différents aspects de la puissance consommée par le convoyeur à bande au démarrage, à charge maximale, à vide et à différents volumes de flux de matériaux.

(S. Zhang & Xia, 2011) se concentrent également sur l'économie d'énergie des convoyeurs à bande grâce à l'amélioration de l'efficacité des opérations, lorsque l'optimisation est utilisée. Un modèle énergétique analytique, issu de l'ISO 5048, est d'abord proposé, qui regroupe tous les paramètres en quatre coefficients. L'estimation des paramètres hors ligne, basée sur la LSQ, et l'estimation des paramètres en ligne, basée sur la RLSQ, sont proposées respectivement pour l'identification de ce modèle énergétique. Enfin, l'optimisation est effectuée pour obtenir la meilleure efficacité de fonctionnement des convoyeurs à bande sous diverses contraintes. Six problèmes d'optimisation d'un système de convoyeur à bande typique sont formulés, respectivement, avec des solutions en simulation pour une étude de cas. Un système de

convoyeur à bande typique est utilisé pour la simulation. Il est démontré par la simulation que le contrôle de la vitesse variable des convoyeurs à bande peut en effet économiser de l'énergie.

Bien que ces méthodes soient précises, elles nécessitent un ensemble complet et détaillé de paramètres du système de convoyeur à bande et des paramètres de fonctionnement, et ces détails ne sont pas toujours disponibles. D'où l'importance de l'apprentissage automatique dans le processus de prédiction. (Zeng & Wu, 2015) ont proposé un modèle de prédiction de la consommation de l'Énergie d'un système de convoyeur à bande basée sur les réseaux de neurones. Ils présentent la conception et la vérification d'un modèle mathématique à court terme destiné à la prédiction de la consommation d'énergie d'un système de convoyeur à bande à l'aide de la méthode des réseaux de neurones. Le modèle obtenu correspond aux conditions réelles de fonctionnement. Il peut tirer parti des paramètres de fonctionnement obtenus en fonction de la mesure expérimentale réelle. Le concept du banc d'essai de l'Université de technologie de Wuhan en Chine est conçu de manière à représenter un système de convoyeur à bande de 3,5 m de long sur lequel des expériences d'acquisition de paramètres de fonctionnement peuvent être menées. Les expériences réalisées montrent une prédiction rapide avec des résultats finaux acceptables pour des données réelles avec un horizon de prédiction à court terme égal à 60 min et avec une erreur moyenne de 4,8 %. Le modèle obtenu est utile pour analyser la vitesse optimale de la bande dans des conditions opérationnelles réelles et pour optimiser les procédures de fonctionnement des systèmes de convoyeur à bande.

Les auteurs (Schmidl, Fischer, Steindl, Wenk, & Franke, 2021) utilisent le machine learning pour présenter un système intelligent capable de s'adapter de manière autonome aux besoins spécifiques de l'usine et de prendre des décisions sur l'état énergétique optimal de chaque composant. Ce système utilise des algorithmes d'apprentissage par renforcement formés sur des modèles d'usine virtuels, permettant à l'agent d'apprendre et de déduire des stratégies optimales pour minimiser la consommation d'énergie sans compromettre le flux de production.

L'étude se concentre sur les chaînes de montage équipées de systèmes de convoyage et de manipulation. L'approche vise à atteindre théoriquement la consommation d'énergie la plus basse en tenant compte de l'énergie consommée pour le transport des porte-pièces et lors des

processus de manipulation individuels. Le système ajuste dynamiquement l'état de chaque composant en fonction des conditions actuelles de l'usine, ce qui lui permet d'économiser de l'énergie pendant les pauses planifiées 15 et non planifiée ainsi que pendant les pauses de processus. Le cadre d'apprentissage par renforcement est basé sur le processus de décision de Markov (MDP), où l'agent interagit avec le modèle de simulation pour apprendre la fonction de valeur d'action optimale (Q-valeurs) qui maximise la récompense. Les états énergétiques des composants déterminent les actions de l'agent, et il utilise l'exploration de Boltzmann pour trouver un équilibre entre l'exploration et l'exploitation lors du processus d'apprentissage. Le système présenté a été testé sur une usine réelle, ainsi que sur son modèle de simulation, et les résultats ont montré des économies d'énergie significatives. Pendant la production, le système a réalisé environ 20 % d'économies d'énergie, tandis que, pendant les pauses plus courtes allant jusqu'à cinq minutes, les économies ont atteint 70 %, et, pendant les périodes d'arrêt plus longues, les économies se sont élevées à 90 %. Cette approche a surpassé les stratégies simples programmées dans le PLC, démontrant le potentiel d'une optimisation supplémentaire pour se rapprocher de la consommation d'énergie minimale théorique.

La présente étude s'inscrit dans cette dynamique. Elle propose une modélisation prédictive de la consommation énergétique des convoyeurs à bande à l'aide d'un modèle de forêt aléatoire, entraîné sur des données mesurées en temps réel à bord de navires auto-déchargeurs. À la différence des approches classiques, cette méthode n'exige pas de connaissance préalable de l'architecture mécanique. Elle permet en outre d'intégrer des contraintes pratiques, telles que le temps de déchargement maximal ou la plage de vitesses autorisée, dans une démarche d'optimisation réaliste et opérationnelle.

1.4 Limitations et défis potentiels

Malgré la richesse des approches théoriques et appliquées dans la littérature, plusieurs limites récurrentes sont identifiables. Les modèles physiques complexes, souvent perçus comme rigides, posent des défis pratiques majeurs. En effet, les méthodes s'appuyant sur des normes, telles que l'ISO 5048, la DIN 22101 ou la CEMA, utilisent des équations fixes et dépendant de paramètres mécaniques qui sont parfois difficiles à mesurer précisément sur le terrain. Ces

modèles, reconnus comme essentiels en phase de conception, montrent leurs limites dès qu'on cherche à optimiser en temps réel des convoyeurs en exploitation.

Un autre obstacle réside dans l'accès aux données opérationnelles. Il est fréquent que les études techniques fassent appel à des jeux de données issus de simulations ou de bancs d'essai, lesquels ne capturent pas pleinement la variabilité des situations réelles. On observe notamment que les environnements portuaires, avec leurs contraintes spécifiques et aléas quotidiens, sont rarement bien représentés dans ces approches. Cette lacune complique l'adaptation des solutions théoriques aux besoins concrets des exploitants selon les conditions d'exploitation réelles, telles que les variations de charge, de vitesse, ou les défaillances mécaniques sont rarement prises en compte.

Nous avons également observé une limitation lors de l'optimisation qui est en général effectuée de manière théorique et rarement opérationnelle. Les stratégies d'optimisation, bien que prometteuses sur le plan algorithmique, comme les réseaux de neurones, apprentissage par renforcement, sont souvent testées dans des environnements simulés ou limités à des cas spécifiques. Elles négligent parfois les contraintes pratiques, telles que la durée de déchargement, la capacité des moteurs ou les caractéristiques spécifiques des cargaisons. La faible intégration des contraintes réelles du métier est perçue, car très peu de travaux intègrent simultanément des objectifs énergétiques, des contraintes de temps, et des spécificités opérationnelles liés au type de navire ou de cargaison. Cette absence de vision holistique réduit la portée opérationnelle des modèles.

1.5 Objectifs de l'étude

L'objectif général de cette étude est de proposer une méthode fondée sur l'apprentissage automatique, visant à optimiser la consommation énergétique des convoyeurs à bande utilisés dans un environnement portuaire. Cette démarche permettra de fournir aux opérateurs des recommandations précises et applicables afin de leur apporter une aide à la décision lors des opérations de déchargement. Pour y parvenir, il sera question tout d'abord d'identifier et

analyser les facteurs influençant la consommation énergétique. Plus précisément il s'agit de réaliser une analyse exploratoire approfondie des données collectées sur les convoyeurs à bande, incluant des variables techniques (vitesse du convoyeur, charge transportée, etc.) ainsi que des variables catégorielles comme le type de cargaison et les différents ports où s'effectuent ces opérations. L'objectif ici est de détecter les relations qui pourraient exister entre ces variables, les corrélations significatives et les tendances générales qui influencent la consommation d'énergie. Cette analyse servira de fondement à la construction du modèle prédictif, en identifiant les variables les plus influentes à prendre en compte.

Par la suite, sur la base des résultats de l'analyse exploratoire, un modèle d'apprentissage automatique sera élaboré afin de prédire de manière fiable la consommation énergétique des convoyeurs en fonction des conditions réelles opérationnelles. Le choix de l'algorithme forêt aléatoire est justifié par la nature des données disponibles (quantité, qualité, distribution, linéarité) ainsi que par la capacité du modèle à généraliser ses prédictions sur de nouveaux jeux de données.

La dernière étape visera à intégrer au processus, un algorithme d'optimisation numérique, en l'occurrence la méthode L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Bound constraints), afin de déterminer les configurations optimales des paramètres de fonctionnement (vitesse du convoyeur, taux de décharge) minimisant la consommation d'énergie. Cette méthode, qui permet de gérer efficacement les contraintes aux limites sur les variables, sera appliquée à divers cas d'usage représentatifs notamment différents types de cargaisons, pour produire des recommandations adaptées au contexte. Les résultats qui découleront de cette optimisation seront traduits en consignes opérationnelles simples pour les opérateurs, favorisant une exploitation plus durable et performante des équipements.

1.6 Contribution et discussion

Dans ce contexte, la présente étude se distingue par plusieurs apports originaux permettant de pallier les limites observées dans la revue de littérature. Nous proposons l'application concrète

d'un modèle d'apprentissage supervisé sur données réelles en utilisant l'algorithme de forêts aléatoires, l'étude construit un modèle prédictif directement entraîné sur les mesures opérationnelles collectées à bord de navires auto-déchargeurs, sans nécessiter de schéma mécanique détaillé. Par la suite, nous intégrons des contraintes opérationnelles dans l'optimisation. La méthodologie proposée combine apprentissage machine et optimisation numérique (méthode L-BFGS-B) pour déterminer les réglages minimisant la consommation d'énergie tout en respectant des seuils concrets, comme un temps de déchargement maximal de 3 heures et des plages admissibles de vitesse.

L'adaptabilité au type de cargaison, contrairement aux approches génériques, permet de prendre en compte les spécificités de chaque cargaison pour ajuster les paramètres optimaux, ce qui reflète mieux la variabilité réelle du transport maritime. Enfin, l'étude illustre comment les technologies de l'Industrie 4.0 (capteurs intelligents, mégadonnées, optimisation par IA) peuvent être mobilisées de manière concrète pour améliorer la performance énergétique d'un système industriel portuaire.

La littérature démontre de manière croissante l'intérêt d'optimiser la consommation énergétique des systèmes de convoyage. Toutefois, les approches diffèrent profondément selon qu'elles relèvent de la modélisation physique, de l'ingénierie empirique ou de l'intelligence artificielle. Les modèles traditionnels généralement basés sur des normes restent pertinents pour la conception initiale, mais manquent de souplesse pour s'adapter en temps réel aux fluctuations des conditions d'exploitation. Leur granularité et leur dépendance à des paramètres techniques limitent leur usage pour l'optimisation dynamique. À l'inverse, les approches par apprentissage automatique, bien plus souples, souffrent parfois d'un manque de robustesse lorsque les données sont bruitées, incomplètes ou trop contextuelles. De plus, certaines études s'appuient sur des données expérimentales dans des situations de banc d'essai, qui ne sont pas généralement communes aux conditions réelles.

L'étude présentée se situe à l'intersection entre précision, opérationnalité et robustesse. En alliant la fiabilité d'un modèle de prédiction aux contraintes métiers concrètes par optimisation

numérique, elle propose une réponse efficace, réaliste et déployable à court terme dans les systèmes portuaires existants.

1.7 Conclusion

Ce chapitre a posé les bases théoriques et techniques de l'étude, en mettant en avant le rôle des convoyeurs à bande dans la manutention de vrac. Il a souligné les enjeux liés à leur consommation énergétique, devenue un enjeu stratégique. La revue de littérature a exposé les limites des approches classiques de modélisation. Ces constats justifient le recours à une modélisation par forêts aléatoires combinée à une optimisation par L-BFGS-B. L'objectif est de proposer une gestion énergétique plus efficace et adaptée aux contraintes réelles.

Le chapitre suivant sera consacré à la présentation détaillée de la méthodologie utilisée pour répondre à la problématique.

CHAPITRE 2

MÉTHODOLOGIE DE RECHERCHE

2.1 Introduction

Ce chapitre présente la méthodologie utilisée pour atteindre l'objectif principal de ce travail, qui est de concevoir un modèle de prédiction et d'optimisation à l'aide des algorithmes intelligents pour minimiser la consommation d'énergie des convoyeurs à bande tout en respectant les contraintes opérationnelles. La démarche adoptée repose sur une combinaison d'analyse des données, d'algorithme de prédiction et d'optimisation à l'aide des méthodes d'apprentissage automatique. Ce chapitre est structuré en plusieurs sections : l'approche de recherche, une description détaillée de la base donnée, le choix des algorithme et technique d'évaluation. Chaque étape est détaillée pour permettre une compréhension claire et transparente des méthodes employées.

2.2 Approche méthodologique adoptée

2.2.1 Contexte industriel

Le contexte industriel dans lequel s'inscrit ce projet impose des exigences à la fois énergétiques, opérationnelles et techniques. Les navires auto-déchargeurs SUL utilisent un système de convoyage à bande dont la consommation de l'énergie est influencée par plusieurs paramètres techniques, tels que la vitesse de la bande, le taux de charge, la puissance des moteurs du convoyeur, l'angle d'inclinaison de la flèche et également par les paramètres opérationnels comme le type de cargo, la nature du port. Nous constatons que la relation entre ces différents paramètres influençant cette consommation est non linéaire, ce qui rend impossible une modélisation par les équations physiques. Il est également impossible pour nous dans le contexte industriel actuel d'avoir des informations originales de conceptions du système sans oublier toutes les opérations de maintenances qui ont été effectuées jusqu'à ce jour.

2.2.2 Justification de l'approche par apprentissage automatique

Face à cette complexité, le recours à des techniques d'apprentissage automatique (machine learning) s'est imposé comme une approche pertinente et particulièrement adapté à ce problème, car il permet de capturer des relations complexes entre variables d'entrée et sortie, sans nécessiter une modélisation physique explicite du système. L'apprentissage supervisé, en particulier, permet d'apprendre à partir d'exemples historiques disponibles, en associant des entrées multivariées à une variable cible continue, qui est la consommation d'électricité du convoyeur. Cette méthode présente l'avantage de s'adapter à des volumes importants de données, de prendre en compte des interactions complexes entre variables, et de généraliser la prédiction à de nouveaux contextes opérationnels.

Une étude précédente a été effectuée sur ce système avec ce même ensemble de données et les résultats ont permis de vérifier que le modèlent forêt aléatoire comme le meilleur algorithme de régression pour la prédiction de la consommation de l'Énergie par sa robustesse face au bruit, sa capacité à modéliser des relations non linéaires, du faible risque de surapprentissage si correctement paramétré, et interprétabilité relative des résultats. Ce modèle a permis d'atteindre des performances prédictives satisfaisantes (R^2 élevé, RMSE faible) sur l'ensemble des données.

Ces résultats nous ont fourni un support de base et une orientation claire vers le choix de notre modèle. Une fois l'algorithme de prédiction établi, l'approche de recherche adoptée dans ce projet étant de minimiser la consommation énergétique des convoyeurs à bande tout en respectant les contraintes opérationnelles, plusieurs algorithmes d'optimisation sont évalués afin de choisir lequel s'adaptera mieux aux données et permettra une minimalisation de la consommation de l'énergie tout en gardant le système performant.

2.2.3 Stratégie d'optimisation de la consommation énergétique

Enfin, pour enrichir l'analyse et mieux comprendre les comportements propres à chaque type de marchandise, le modèle a été décliné sous deux formes : une version généraliste (tous cargos confondus), et des modèles spécialisés entraînés uniquement sur les données de chaque type de cargo. Cette distinction permet d'évaluer l'intérêt d'un ajustement fin du modèle aux spécificités opérationnelles de chaque cargaison.

Afin de mieux encadrer notre projet, la méthodologie repose sur une méthodologie quantitative et orientée par les données, intégrant plusieurs étapes clés :

Étape 1 : cette étape est consacrée à la définition du cadre de la recherche et repose sur une étude bibliographique approfondie, visant d'abord à clarifier les notions clés, puis à analyser leurs interactions. Elle inclut une revue détaillée des travaux scientifiques en lien direct avec notre problème, et se conclut par des exemples concrets d'applications similaires dans d'autres secteurs industriels utilisant les mêmes concepts.

Étape 2 : conception et entraînement du modèle d'apprentissage automatique. Cette étape comprend l'analyse exploratoire des données, la sélection des variables les plus influentes sur la consommation de l'énergie et la conception de l'algorithme de prédiction et d'optimisation

Étape 3 : évaluation des performances.

2.3 Base de données

2.3.1 Description générale du jeu de données et source d'acquisition

Les données utilisées dans cette étude ont été collectées du 1/1/2020 au 1/1/2023 avec un intervalle de 5min par plusieurs capteurs en temps réel et stockées dans des fichiers CSV et Excel. Ces capteurs comprennent des capteurs de proximité, qui détectent la présence de la cargaison sur la bande transporteuse ; des cellules de charge, qui mesurent le poids de la cargaison ou la force appliquée ; des capteurs de vitesse, qui fournissent des données sur les taux de déplacement de la bande transporteuse ; des capteurs de température, qui surveillent les températures des composants pour un fonctionnement sûr ; des capteurs de vibration, qui

déetectent les problèmes mécaniques par les vibrations ; des capteurs optiques, qui identifient les objets et mesurent les distances ; des capteurs de pression, qui surveillent les changements de pression dans le système ; et des capteurs de position, qui fournissent des informations sur l'alignement et le mouvement des différents composants du convoyeur.

À travers ses capteurs, nous avons pu obtenir des données sur les paramètres qui nous permettront de faire une analyse de la consommation de l'Énergie, notamment le tunnel, le transfert, la boucle et la flèche qui représentent les étapes de déchargement dans les convoyeurs. Chaque étape comporte des moteurs maîtres et esclaves. Les paramètres utilisés dans ces moteurs comprennent la puissance du moteur (kW), la vitesse (tr/min) et les températures (degrés Celsius). Nous avons également la puissance des générateurs du navire, AE1, AE2 et AE3. Les données contiennent également le type de cargaison en cours de déchargement, le port où le navire est en cours de déchargement, le taux de déchargement (t/h) utilisé pendant le processus de déchargement, l'angle de la flèche (degrés) au moment du déchargement et le numéro de voyage.

2.3.2 Structure du jeu de données

Le jeu de données comprend en moyenne 35 colonnes ou variables et environ 31,758 lignes ou observations, représentant les cycles de fonctionnement des systèmes de convoyage à travers plusieurs navires et contextes opérationnels. Pour cette étude nous avons les données de 4 navires à savoir : Navire 1, Navire 2, Navire 3, et Navire 4. Ces fichiers sont configurés de la même manière dont les principales variables sont décrites comme suit :

- Variables temporelles : La colonne date fournit des "timestamps" qui représentent le temps d'enregistrement de l'événement.
- Paramètres des moteurs : Variables comme « boom master motor power » (puissance du moteur maître), ou « boom slave motor power » (puissance du moteur esclave), boom master motor temp (température des moteurs), etc.
- Angle : Variable « boom inclination angle » qui représente l'angle d'inclinaison de la flèche.

- Vitesses : Variables comme « boom master motor speed » qui représentent les vitesses des moteurs.
- Consommation énergétique : « Electricity consumption » qui est la consommation du convoyeur et « House usage » (ou encore consommation énergétique de l'habitacle).
- Autres attributs : Variables catégoriques telles que « Cargo » qui représente le type de matériaux décharge et Port.

L'ensemble de ces variables sont résumées dans le Tableau 2.1 ci-après avec leur type de données

Tableau 2.1 Caractéristiques des variables du jeu de données

Nom de la variable	Description	Type de données
Boom master motor power	Puissance en kW du moteur maître de la flèche du convoyeur.	Continue (numérique)
Boom slave motor power	Puissance en kW du moteur esclave de la flèche du convoyeur.	Continue (numérique)
boom master motor temp	Température en °C du moteur maître.	Continue (numérique)
Boom inclination angle	Angle d'inclinaison de la flèche du convoyeur en degrés.	Continue (numérique)
Transfer master/slave motor	Puissance et température des moteurs principaux du système de transfert.	Continue (numérique)
Tunnel master/slave motor	Paramètres des moteurs dans le système de tunnel (puissance/température).	Continue (numérique)
Discharge rate	Taux de déchargement en tonnes/heure.	Continue (numérique)
Electricity consumption	Consommation électrique globale en kWh.	Continue (numérique)
Cargo	Type de cargaison (ex. AGRÉGATS, CIMENT, GRAINS).	Catégorielle (texte)
Port	Port où les opérations de déchargement sont réalisées.	Catégorielle (texte)

L'analyse de ces données est cruciale, car elle permet non seulement de comprendre les facteurs qui influencent la consommation énergétique afin d'identifier des leviers d'optimisation, mais

aussi d'éliminer les variables non pertinentes, rendant ainsi le jeu de données mieux structuré pour l'apprentissage automatique.

2.3.3 Objectif de l'analyse des données

L'analyse des données est une étape cruciale pour identifier les modèles et les tendances dans les données, afin de mieux comprendre le problème et les solutions possibles. Cela permet également d'acquérir des connaissances supplémentaires. Pendant cette étape, différentes visualisations, notamment des histogrammes et des diagrammes en boîte, sont utilisées pour interpréter les relations qui existent entre les variables. Ces graphiques permettent de mieux comprendre les comportements et la distribution de chaque variable.

Par la suite, nous allons utiliser une matrice de corrélation qui est un outil statistique permettant de montrer la force et la direction de la relation entre deux ou plusieurs variables. L'objectif est d'analyser la relation entre la consommation d'énergie et des variables, telles que la vitesse (“Speed”), la charge transportée (“discharge rate”), et les conditions extérieures, telles que le port ou le type de cargo. Comprendre comment ces variables influencent la consommation d'énergie est essentiel, car elles constituent des facteurs déterminants dans l'optimisation énergétique. Elle permet également de visualiser les variables qui sont fortement corrélées entre elles afin de réduire les redondances. La représentation temporelle de la consommation d'énergie met en évidence les variations liées à des facteurs opérationnels et environnementaux, permettant ainsi de visualiser les périodes où la consommation est la plus élevée. Ces observations serviront à orienter la prédiction et l'optimisation.

L'analyse s'inscrit dans une logique d'optimisation énergétique, de réduction des coûts et d'amélioration des techniques opérationnelles.

2.3.4 Prétraitement des données

La performance d'un modèle d'apprentissage automatique repose en grande partie sur la qualité des données et sur leur préparation rigoureuse. Le prétraitement des données comprend plusieurs étapes clés, notamment le traitement des données manquantes, la détection et la

suppression des valeurs aberrantes, l'analyse des relations entre les variables et leur division en ensembles distincts. La section suivante détaille les méthodes appliquées pour préparer le jeu de données avant son intégration dans les modèles d'intelligence artificielle.

Par ailleurs, le langage de programmation Python, accompagné de bibliothèques puissantes telles que Pandas, NumPy et Seaborn, a été largement utilisé pour manipuler et analyser les données. Les fonctions précises des enjeux de bibliothèque mobilisées à chaque étape du processus de prétraitement seront également présentées.

Pour effectuer ce traitement, nous allons premièrement charger notre fichier via des outils Pandas (Python) qui nous permettra de préparer les données pour garantir leur qualité. Nous utilisons les fonctions `print(data.isnull().sum())` pour quantifier les Valeurs manquantes `print(data.describe())` pour avoir un aperçu des Statistiques descriptives.

2.3.5 Gestion des valeurs aberrantes

Dans de nombreuses applications, le traitement et l'analyse de grands volumes de mesures font face à l'apparition de données anormales (valeurs aberrantes). Dans ce cas, il est important de déterminer si une nouvelle observation anormale appartient à la même distribution que les observations existantes ou si elle doit être considérée comme une manifestation de nouvelles propriétés ou de nouveaux phénomènes de l'objet étudié (Illyushin, 2024). Les valeurs aberrantes anormales peuvent être dues à une erreur de mesure ou au bruit. Lorsque les données sont analysées, ces valeurs aberrantes peuvent conduire à une erreur exagérée. Ces valeurs aberrantes devraient être identifiées et écartées dans une analyse plus approfondie. Pour effectuer l'analyse de ces valeurs aberrantes, nous avons appliqué la méthode de l'écart interquartile (IQR). En traitant des données expérimentales par cette méthode, on peut filtrer les mesures erronées dans les queues des distributions. Comme indiqué précédemment, des mesures erronées peuvent introduire une erreur significative dans le calcul des caractéristiques statistiques. (Swami et al., 2023) ont utilisé la technologie IQR pour détecter efficacement les

attaques par inondation SYN en SDN. Ils ont vérifié l'efficacité de la détection des anomalies IQR dans plusieurs scénarios.

L'IQR qui a été développée par John Tukey n'a besoin de calculer que les quartiles de données, sans itérations complexes ni ajustements de paramètres, afin de pouvoir déterminer rapidement les limites supérieures et inférieures des valeurs aberrantes. Le calcul de l'IQR est le suivant : soit un ensemble de données triées $\mathbf{X} = [x_1, x_2, \dots, x_n]$. Tout d'abord, les positions théoriques du quartile inférieur Q_1 et du quartile supérieur Q_3 sont calculées selon l'équation (2-1). Puis Q_1 et Q_3 sont calculés à l'aide de l'équation (2-2) et, enfin, la valeur IQR est obtenue par l'équation (2-3). (Yue, Yan, Han, & Wu, 2025) proposent une plage de 1,5 fois IQR de comme règle IQR pour marquer les valeurs aberrantes possibles. D'après le calcul, les équations (2-4), (2-5) sont obtenues comme les limites inférieure et supérieure des valeurs aberrantes, respectivement.

Formule :

$$\begin{cases} P_1 = (n + 1) * 0.25 \\ P_3 = (n + 1) * 0.75 \end{cases} \quad (2-1)$$

$$\begin{cases} Q_1 = x_{\lfloor P_1 \rfloor} + (P_1 - \lfloor P_1 \rfloor) \times (x_{\lceil P_1 \rceil} - x_{\lfloor P_1 \rfloor}) \\ Q_3 = x_{\lfloor P_3 \rfloor} + (P_3 - \lfloor P_3 \rfloor) \times (x_{\lceil P_3 \rceil} - x_{\lfloor P_3 \rfloor}) \end{cases} \quad (2-2)$$

$$IQR = Q_3 - Q_1 \quad (2-3)$$

$$T_{low} = Q_1 - 1.5 * IQR \quad (2-4)$$

$$T_{up} = Q_3 + 1.5 * IQR \quad (2-5)$$

Toutes les données se trouvant à l'extérieur de ces limites sont considérées comme valeurs aberrantes. La Figure 2.1 suivante nous permet de visualiser les positions des limites et de repérer les valeurs aberrantes.

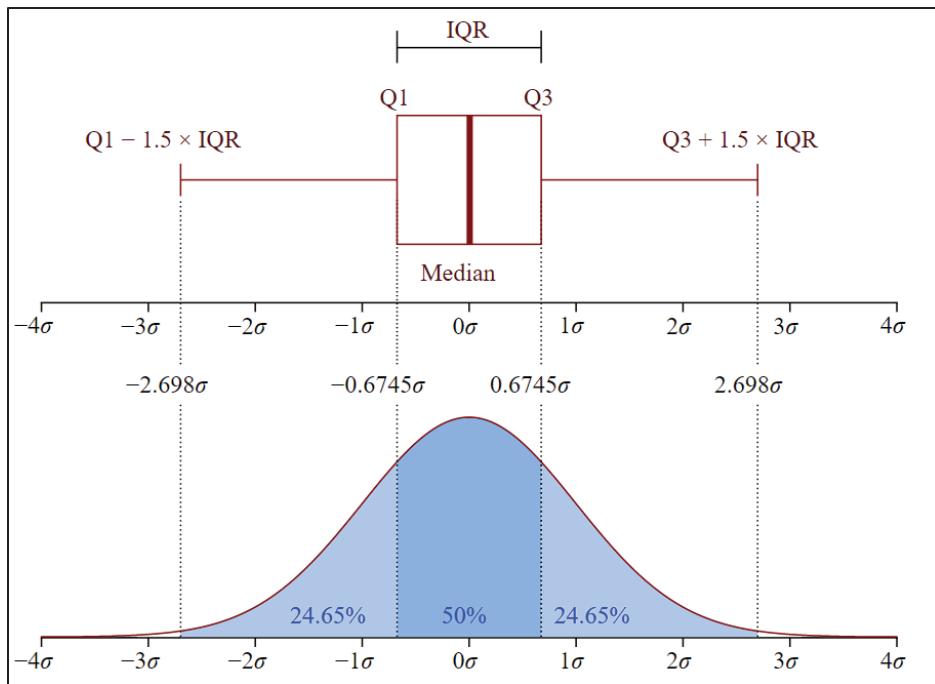


Figure 2.1 Visualisation de la méthode IQR
(Klacar, 2023)

2.3.6 Séparation des données

Il est important d'évaluer la précision des prédictions en utilisant des valeurs réelles. Par conséquent, la précision des prédictions ne peut être déterminée qu'en examinant les performances d'un modèle sur de nouvelles données qui n'ont pas été utilisées lors de l'apprentissage du modèle. Généralement, 3 types d'ensembles de données sont utilisés : ensembles d'entraînement, ensemble de validation, et l'ensemble test. Cette division garantit que le modèle est évalué sur des données qu'il n'a pas rencontrées lors de l'entraînement. Il s'agit d'un processus itératif où le modèle apprend à partir des données d'entraînement, puis est validé et amélioré sur les données de validation. Les ensembles de données de tests sont utilisés pour évaluer l'adéquation finale du modèle, en utilisant des données du monde réel non-vues auparavant. Ils permettent d'évaluer comment le modèle se comporterait lorsqu'il rencontre de nouvelles données dans un environnement réel. Dans la plupart des cas, 80 % des données sont utilisées pour entraîner et valider le modèle et les 20 % restants sont utilisés pour

un test de généralisation complètement indépendant du modèle choisi. Pour éviter tout biais involontaire envers les différents ensembles, la séparation est faite de manière aléatoire.

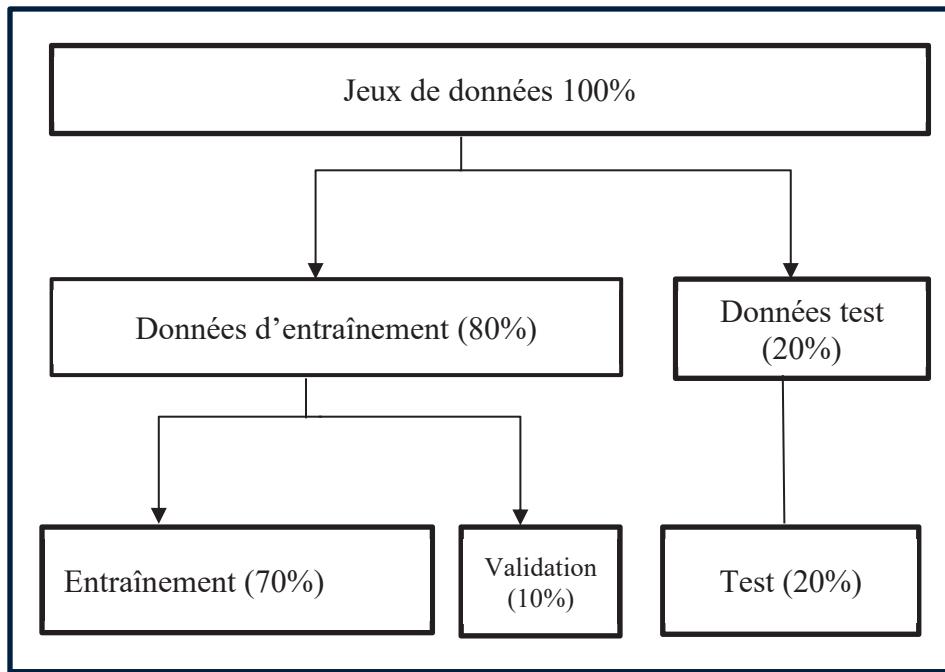


Figure 2.2 Séparation du jeu de données

Rappelons qu'un modèle est dit en surapprentissage lorsqu'un modèle d'apprentissage automatique apprend trop bien les données d'entraînement, au point de mémoriser les détails, les bruits ou les anomalies spécifiques à ces données au lieu de faire une généralisation à de nouvelles données. On observe cependant de bonnes performances durant la phase l'entraînement, mais au moment de la validation, elles se dégradent énormément.

2.4 Évaluation des performances du modèle

L'évaluation des modèles est un processus crucial de la modélisation. Les mesures d'évaluation permettent d'optimiser les paramètres d'apprentissage des modèles et de définir les plus performants parmi leurs pairs. Cet effort aboutit à l'identification du modèle qui présente les performances élevées pour la prédiction de la consommation d'énergie. Il existe différentes mesures pour évaluer les résultats des modèles d'optimisation. Parmi celles-ci,

l'erreur absolue moyenne "Mean Average Error" (MAE), l'erreur quadratique moyenne "Mean Square Error" (MSE) et la racine de l'erreur quadratique moyenne "Root Mean Square Error" (RMSE) sont des mesures conventionnelles et largement utilisées. En outre, le pourcentage d'erreur absolue moyenne "Mean Average Percentage Error"(MAPE) est souvent utilisé pour exprimer le pourcentage d'erreur, ce qui est plus intuitif. Enfin, R-carré (R^2) indique si le modèle est bien adapté aux valeurs observées et à quel point il l'est. Une brève description de ces mesures d'évaluation est présentée dans ce qui suit:

MAE, MSE, RMSE: ces trois mesures sont largement utilisées pour évaluer les résultats de la régression. Elles sont utiles pour comparer différentes méthodes appliquées à un même ensemble de données. Leurs expressions mathématiques sont données dans les équations suivantes :

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2-6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2-7)$$

$$RMSE = \sqrt{MSE} \quad (2-8)$$

MAE (Mean Absolute Error) mesure la moyenne des erreurs absolues entre les valeurs prédites et les valeurs réelles, exprimées dans la même unité que les données d'origine. Un MAE égal à zéro signifie que les prédictions sont parfaitement exactes. L'utilisation de la valeur absolue évite que les erreurs positives et négatives ne se compensent. Sa principale caractéristique est de traiter toutes les erreurs de manière égale, qu'elles soient faibles ou importantes.

MAPE : est une version avancée de MAE qui supprime la dépendance à l'égard de l'échelle. Comme il s'agit d'une erreur en pourcentage, elle donne une bonne idée de l'erreur relative et est donc fréquemment utilisée pour comparer les performances des prédictions entre les ensembles de données. Son expression mathématique est donnée dans l'équation ci-dessous :

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{\hat{y}_i} * 100 \quad (2-9)$$

R-carre (R^2): représente la part de la variance des valeurs réelles (y) qui est expliquée par les variables d'entrée du modèle. Il offre ainsi une mesure de la qualité des prédictions, notamment lorsqu'on applique le modèle à de nouvelles données de test. Une valeur élevée signifie que la corrélation entre les valeurs observées et prédictes est élevée. Son expression mathématique est donnée dans l'équation :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2-10)$$

Lorsque les prédictions sont proches des observations réelles, le coefficient R^2 tend vers 1, ce qui indique une bonne qualité de prédiction. À l'inverse, des prédictions éloignées des valeurs observées réellement entraînent un R^2 proche de 0 et peuvent aussi prendre des valeurs négatives indiquant les performances médiocres en matière de prédictions. Toutefois, cette métrique présente une limite importante, car elle a tendance à augmenter avec l'ajout de nouvelles variables d'entrée, même si celles-ci ne sont pas réellement pertinentes. Cela peut conduire à un surapprentissage, où le modèle affiche un R^2 artificiellement élevé, tout en perdant sa capacité de généralisation et sa précision réelle sur de nouvelles données.

Dans les équations suscités, y_i est la valeur réelle, \hat{y}_i est la valeur de prédiction indexée avec i , \bar{y} est la moyenne des valeurs observées, et n est le nombre de points de données, respectivement.

Limites de la Méthodologie

- **Données imparfaites** : Certaines variables, comme les conditions climatiques ou l'état des convoyeurs n'ont pas été incluses faute de données disponibles.
- **Simplifications** : Les relations entre variables sont modélisées de manière linéaire dans certaines parties du modèle.

2.5 Environnement de développement

Il s'agit de l'ensemble des outils qui nous ont permis de développer notre solution. Nous pouvons citer entre autres :

Google Colab : Colab, abréviation de *Colaboratory* est une plateforme gratuite proposée par Google permettant d'écrire et d'exécuter du code Python directement depuis un navigateur, sans installation préalable. Basé sur Jupyter Notebook, Colab est particulièrement adapté à la formation et à la recherche en apprentissage automatique. Il offre la possibilité d'entraîner des modèles d'apprentissage automatique dans le nuage, tout en facilitant la pratique du Python et l'utilisation de bibliothèques populaires comme TensorFlow, Keras, PyTorch ou OpenCV. Sa particularité est qu'il permet d'exécuter les sous-blocs de code permettant une compréhension facile et une détection des erreurs rapides.

Python s'est imposé ces dernières années comme le langage de programmation de référence dans de nombreux domaines. Grâce à sa simplicité syntaxique, il permet aux développeurs de se concentrer sur la logique métier plutôt que sur des détails techniques. À l'aide de ce langage, il est possible d'avoir une bibliothèque qui représente un ensemble de modules préconstruits avec des fonctions, méthodes et classes directement prêt à l'emploi, ce qui vient simplifier le processus de codage. Il couvre un large éventail, comme la science des données, les opérations mathématiques avancées, et permet même la création des graphiques.

NumPy : dans le cadre de ce projet, nous avons utilisé cette bibliothèque pour sa référence en matière de calcul scientifique. Elle permet de manipuler des tableaux multidimensionnels et d'appliquer des fonctions mathématiques complexes.

Matplotlib : est une bibliothèque de visualisation de données en Python. Elle permet de générer des graphiques, des courbes et des images de manière simple permettant une visualisation statique, animés et même interactifs. Il est possible d'obtenir n'importe quel graphique, comme les diagrammes à barres, les nuages de points, les histogrammes, voire les représentations 3D. Nous l'avons utilisée pour représenter visuellement les techniques appliquées pour le traitement des données.

Pandas : Pandas est une bibliothèque Python incontournable pour l'analyse et la manipulation de données structurées. Elle propose des structures puissantes, comme les Séries et les

DataFrames, facilitant le nettoyage, la transformation et l'analyse des données. Pandas permet d'importer/exporter depuis divers formats (CSV, Excel, SQL, etc.), de gérer les valeurs manquantes, de fusionner des tableaux, d'appliquer des opérations vectorisées et de regrouper ou agréger les données facilement. Elle s'intègre parfaitement avec NumPy et d'autres outils, rendant possible la création de pipelines pour le traitement de données complets et efficaces.

Sklearn : est une bibliothèque de machine learning en Python à la fois simple par son utilisation et très puissante. Elle couvre un large éventail d'algorithmes d'apprentissage supervisé et non supervisé, incluant la classification, la régression, le clustering et la réduction de dimensionnalité. Grâce à une API cohérente et intuitive, elle permet de faire du prototypage, tester et comparer facilement différents modèles. Elle propose aussi des outils pratiques pour le prétraitement des données, la gestion des valeurs manquantes, le rééchantillonnage, ainsi que des modules pour le traitement de texte et la sélection de caractéristiques.

2.6 Stratégie de développement

La figure 2.3 suivante présente de manière générale les différentes étapes de développement de notre solution pour optimiser la consommation d'énergie d'un système de convoyeurs à bande. Il couvre la préparation des données, la modélisation prédictive (générale et par type de cargo), l'évaluation des performances, et se conclut par des recommandations opérationnelles.

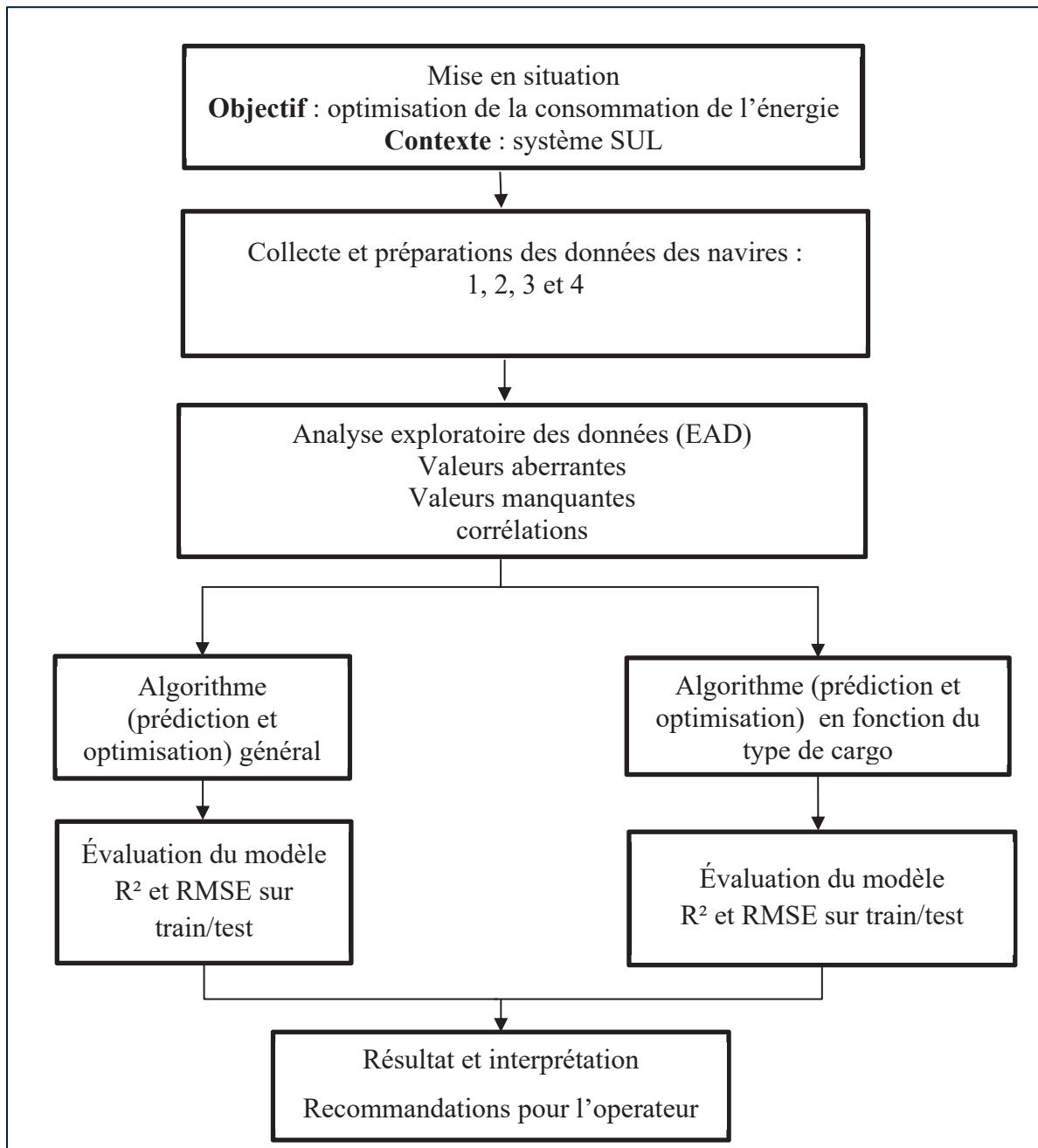


Figure 2.3 Étapes de développement de notre solution

2.7 Conclusion

Ce chapitre a décrit la démarche adoptée pour analyser, prédire et optimiser la consommation d'énergie des convoyeurs à bande. Les étapes, depuis la collecte des données jusqu'à la validation du modèle, ont été présentées en détail. Cette méthodologie fournit une base rigoureuse pour les résultats présentés dans les chapitres suivants.

CHAPITRE 3

CONCEPTION DU MODÈLE DE PRÉDICTION ET OPTIMISATION

3.1 Introduction

Dans ce travail, nous proposons un modèle pour la prédiction et l'optimisation de la consommation d'énergie des convoyeurs à bandes des navires SUL. Le modèle proposé est une combinaison de l'algorithme de forêt aléatoire pour la prédiction et l'algorithme L-BFGS-B (une méthode de gradient limitée avec contraintes) qui nous permettra de faire l'optimisation. Ce chapitre est consacré à la présentation détaillée de la conception du modèle proposé. Nous le commençons par présenter l'architecture générale du système proposé avec ces différentes phases. Puis on passe à une explication détaillée de différentes parties du système.

3.2 Développement du modèle de prédiction avec le modèle de forêt aléatoire

Selon Jayaramu et al. (2023), le modèle de forêt aléatoire est une méthode d'apprentissage automatique bien connue qui utilise l'apprentissage d'ensemble pour résoudre des problèmes de classification et de régression. Pour résoudre des problèmes difficiles, elle utilise une combinaison de classificateurs. Le nom 'Forêt' provient de l'algorithme, puisque celui-ci crée de nombreux arbres de décision. Un méta-algorithme qui aide à rendre les modèles d'apprentissage automatique plus précis est le bagging ou agrégation bootstrap, et il est utilisé pour entraîner cette forêt. Cette méthode est développée par Shibly, Sharma et Naleer (2021) et utilise les décisions collectives des arbres individuels pour produire des prédictions et trouver la moyenne ou la moyenne de toutes les projections des arbres donnant le résultat. La précision du modèle augmente proportionnellement à la densité des arbres dans la forêt. En améliorant la précision et en évitant les pièges du surapprentissage, les algorithmes de forêt aléatoire produisent des prédictions plus fiables que les algorithmes d'arbres de décision.

À la base, un arbre de décision est une structure de type organigramme où chaque nœud interne représente une décision basée sur une caractéristique, chaque branche représente le résultat de cette décision et chaque nœud terminal représente un résultat final

L'algorithme du modèle de la forêt aléatoire fonctionne en trois étapes principales. Tout d'abord, il utilise le bootstrapping, créant plusieurs sous-ensembles de données d'entraînement via un échantillonnage aléatoire avec remplacement, chaque sous-ensemble étant utilisé pour développer un arbre de décision individuel (Chen et al., 2015). Ensuite, l'algorithme intègre le caractère aléatoire des caractéristiques en sélectionnant un sous-ensemble aléatoire de caractéristiques lors de la division d'un nœud lors de la construction de l'arbre. Cette approche introduit de la diversité parmi les arbres, ce qui permet au modèle de capturer différents modèles dans les données (Chen et al., 2015). Enfin, après l'entraînement, les prédictions de tous les arbres de décision sont agrégées ; pour les tâches de régression, le résultat est la moyenne de toutes les prédictions d'arbre, tandis que, pour les tâches de classification, il est déterminé par un vote majoritaire.

Cet algorithme peut être défini mathématiquement à travers trois étapes principales. Tout d'abord, le bootstrap consiste à créer m Exemples bootstrap D_1, D_2, \dots, D_m à partir du jeu de données d'entraînement d'origine D de taille N par échantillonnage aléatoire avec remplacement. Ensuite, pour chaque échantillon d'amorçage D_{je} , un arbre de décision T_{je} est construit en partitionnant récursivement les données, la division à chaque nœud étant déterminée par l'évaluation des caractéristiques j et l'utilisation de critères d'impuretés, tels que l'impureté de Gini ou l'erreur quadratique moyenne (MSE). Enfin, pour une nouvelle observation x , la prédiction de chaque arbre T_{je} est obtenue, et la prédiction finale \hat{y} est calculé en faisant la moyenne des prédictions pour les tâches de régression ou en prenant le vote majoritaire pour les tâches de classification.

$$\hat{y} = \frac{1}{m} \sum_{i=1}^m \hat{y}_{je} \quad (11)$$

Il offre également plusieurs avantages. Il démontre une robustesse au surajustement ; en agrégeant plusieurs arbres, il réduit la variance et aide à atténuer le surapprentissage, ce qui est particulièrement important pour les ensembles de données complexes. Il permet également de

capturer la non-linéarité qui est un avantage clé, car elle ne suppose pas de relation linéaire entre les caractéristiques et la variable cible, ce qui la rend adaptée à la modélisation d'interactions complexes et non linéaires que l'on trouve couramment dans les données industrielles.

3.2.1 Conception de l'algorithme de prédiction

La figure ci-dessous représente les différentes étapes pour effectuer la prédiction de la consommation de l'énergie. Chaque étape comporte des actions précises qui permettront d'avoir de meilleurs résultats.

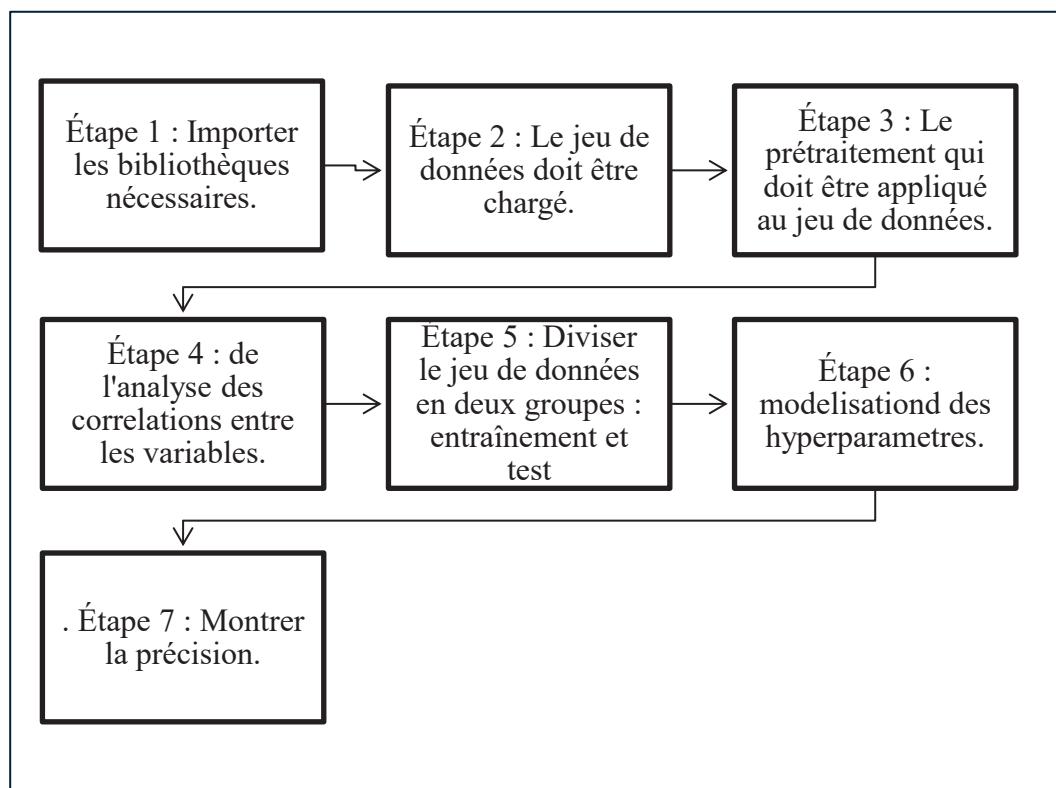


Figure 3.1 Étapes de prédiction de la consommation de l'énergie

Pour adapter le modèle d'apprentissage automatique de forêt aléatoire à différents problèmes, les paramètres du modèle doivent être ajustés. Le choix de la meilleure configuration de paramètres pour le modèle a un impact direct sur les performances du modèle.

Dans l'apprentissage supervisé, l'entrée x et la sortie y sont disponibles, et l'objectif est d'obtenir une fonction de modèle prédictif optimal f^* pour minimiser la fonction de coût $\mathcal{L}(f(x, y))$ qui modélise l'erreur entre la sortie estimée et les étiquettes de vérité terrain. La fonction f du modèle prédictif varie en fonction de sa structure de modèle. Avec des architectures de modèle limitées déterminées par différentes configurations d'hyperparamètres, le domaine de la fonction du modèle d'apprentissage automatique f est limité à un ensemble de fonctions F . Ainsi, le modèle prédictif optimal f^* peut être obtenu par :

$$f^* = \arg \min_{f \in F} \frac{1}{n} \sum_{i=1}^n \mathcal{L}(f(x_i), y_i) \quad (12)$$

où n est le nombre de points de données d'entraînement, x_i est le vecteur caractéristique de la i ème instance, y_i est la production réelle correspondante et \mathcal{L} est la valeur de la fonction de coût de chaque échantillon.

3.2.2 Hyperparamètres

Pour construire un modèle d'apprentissage automatique optimal, un éventail de possibilités doit être exploré. Le processus de conception de l'architecture de modèle idéale avec une configuration optimale des hyperparamètres est appelé réglage des hyperparamètres. Le réglage des hyperparamètres est considéré comme un élément clé de la construction d'un modèle d'apprentissage automatique efficace, en particulier pour les modèles basés sur des arbres, qui ont de nombreux hyperparamètres. Le processus de réglage des hyperparamètres est différent entre les différents algorithmes en raison de leurs différents types d'hyperparamètres, y compris les hyperparamètres catégoriels, discrets et continus. Afin d'assurer la robustesse, la performance et la généralisation du modèle de prédiction pour la consommation de l'énergie, une attention particulière sera portée à la sélection et à la gestion des hyperparamètres du modèle de forêt aléatoire. Ces paramètres influencent directement la qualité de la prédiction et la capacité du modèle à capturer les relations complexes entre les variables d'entrée et la consommation énergétique.

Nombre d'arbres (n_estimators) dans sklearn

Le paramètre `n_estimators` contrôle le nombre d'arbres de décision générés dans la forêt. Plus ce nombre est élevé, plus la variance des prédictions est réduite, ce qui renforce la stabilité du modèle. Toutefois, cela augmente le temps de calcul. Dans le cadre de cette étude, une valeur de 100 arbres a été retenue, car elle offrait un bon compromis entre précision et temps d'exécution. Des tests ont confirmé qu'augmenter au-delà de 100 n'apportait pas de gain significatif en performance.

Profondeur maximale des arbres (max_depth)

Il représente le nombre minimum d'échantillons requis pour diviser un nœud interne. Limiter la profondeur des arbres permet de réduire le risque de surapprentissage, en particulier sur des jeux de données contenant du bruit ou des corrélations. Dans notre cas, la profondeur a été laissée non contrainte, ce qui est fréquent dans l'utilisation du modèle de prédiction, car l'agrégation de nombreux arbres compense leur surapprentissage individuel. Toutefois, un suivi attentif des scores R² sur les jeux de test et d'entraînement permettra de vérifier l'absence de surajustement significatif.

Nombre de variables considérées à chaque division (max_features)

Le paramètre `max_features` détermine le nombre de variables explicatives testées à chaque division d'un nœud, ce qui introduit de la diversité entre les arbres. Par défaut, pour un modèle de régression, cette valeur est fixée à 1.0, ce qui s'avère généralement efficace. Cette valeur par défaut a été conservée après vérification, car elle assurait un bon équilibre entre variance et biais.

Graine aléatoire (random_state)

Afin de garantir la reproductibilité des résultats, une valeur fixe a été attribuée au paramètre `random_state`. Cela permet d'utiliser les mêmes échantillons aléatoires d'une exécution à l'autre, ce qui est essentiel pour comparer les performances et analyser les résultats de manière cohérente.

Stratégie de validation

Les performances du modèle ont été évaluées à l'aide d'une **séparation train/test** avec un ratio de 80/20, conservant ainsi un échantillon indépendant pour mesurer la capacité du modèle à généraliser. Les métriques R^2 et RMSE ont été utilisées pour évaluer la qualité de la prédiction à la fois sur les données d'apprentissage et de test.

Le Tableau 3.1 suivant résume l'ensemble des hyperparamètres utilisés pour construire le modèle de prédiction.

Tableau 3.1 Hyperparamètres du modèle

Hyperparamètres	Valeurs attribuées
n_estimators	100
max_depth	none
max_features	1.0
random_state	42
criterion	squared_error
bootstrap	True
min_samples_split	2
min_samples_leaf	1

3.2.3 Validation croisée

Plusieurs bibliothèques open source sont disponibles dans la littérature pour mettre en œuvre les théories et algorithmes d'apprentissage automatique. Dans le cadre de notre travail, nous avons utilisé *scikit-learn (sklearn)*, une bibliothèque libre en Python. Cette dernière permet notamment d'implémenter *GridSearchCV* pour identifier les hyperparamètres optimaux. Chaque combinaison d'hyperparamètres, définie dans un espace de recherche préalablement spécifié, est testée et évaluée à l'aide de la validation croisée. Une fois toutes les configurations explorées, la meilleure combinaison d'hyperparamètres, associée à son score de performance, est retournée.

Par ailleurs, sklearn propose également *RandomizedSearchCV*, une méthode alternative qui sélectionne aléatoirement un nombre prédéfini de combinaisons d'hyperparamètres dans l'espace de recherche. Chaque configuration est également évaluée par validation croisée, ce qui permet une exploration plus rapide lorsque l'espace des hyperparamètres est vaste.

3.2.4 Entrainement du modèle

Dans le cadre de cette étude, nous avons structuré l'entraînement du modèle prédictif de la consommation énergétique selon une approche à deux volets, combinant à la fois une modélisation globale et une modélisation spécifique à chaque type de cargaison. Le volet général consiste à entraîner un unique modèle de prédiction sur l'ensemble du jeu de données, sans distinction explicite du type de matière transportée. Cette démarche permet de capturer les tendances globales reliant les variables d'entrée (telles que le débit de déchargement ou la vitesse du moteur principal) à la consommation énergétique, en supposant que les relations mécaniques sous-jacentes sont relativement homogènes d'un cargo à l'autre. Toutefois, une analyse exploratoire préalable des données a révélé des écarts significatifs de consommation selon la nature des matériaux déchargés. Par exemple, des cargos transportant du sel ou du sucre présentent, à débit équivalent, des consommations sensiblement différentes de celles observées pour le minerai ou le gypse. Ces écarts laissent supposer que certaines caractéristiques physiques des matériaux (telles que la densité, l'humidité, la granularité ou le coefficient de frottement) influencent les pertes énergétiques au sein du système de convoyage.

En réponse à cette hétérogénéité, un volet spécifique a été introduit, consistant à entraîner des modèles indépendants pour chaque type de cargaison disposant d'un volume de données suffisant. Chaque sous-modèle est ainsi calibré exclusivement à partir des observations correspondant à une seule matière. Cette approche permet de mieux capter les effets propres aux interactions entre le matériau transporté et le système technique (résistance, usure, glissement), et donc de produire des prédictions plus précises. En pratique, ces modèles spécifiques sont ensuite utilisés pour guider l'optimisation du couple (débit, vitesse) dans une

logique de minimisation de la consommation énergétique, avec des résultats nettement plus pertinents que ceux obtenus avec le modèle global. Cette stratégie en deux temps s'inscrit donc dans une logique hybride alliant robustesse statistique et spécialisation opérationnelle, en vue d'une optimisation fine des paramètres de fonctionnement selon la nature des matières transportées.

3.3 Intégration de l'algorithme d'optimisation

3.3.1 Sélection des variables d'optimisation explicatives

Pour une solution opérationnelle réaliste dans un contexte industriel, il est impossible de faire une intervention sur l'ensemble des caractéristiques du convoyeur dans le but de minimiser la consommation d'énergie. Cependant, lorsqu'on observe le fonctionnement du convoyeur l'on peut remarquer qu'il existe des éléments clés sur lesquels l'opérateur peut interagir dans le but d'effectuer des opérations fiables et économiques. L'analyse exploratoire des données nous a permis de comprendre exactement quelles sont les variables qui influencent la consommation de l'énergie de façon proportionnelle. Parmi ces variables, nous avons retenu les trois variables suivantes :

Taux de décharge : mesurée en tonne par heure, elle constitue une variable directement contrôlable, reflétant le rythme auquel les matériaux sont transférés sur la bande du convoyeur. Il influence mécaniquement la charge imposée aux équipements, et donc la consommation énergétique du système. Un débit trop élevé peut engendrer des pics de consommation, tandis qu'un débit trop faible peut rallonger inutilement la durée d'opération, générant une consommation cumulative plus importante. De plus, après analyse de la matrice de corrélation entre les différentes variables, elle présente un coefficient de corrélation, noté r , le plus élevé avec la consommation de l'Énergie.

Vitesse des moteurs : bien qu'il existe plusieurs moteurs d'entraînement (maître et esclave), tous ces moteurs fonctionnent de manières identiques avec des caractéristiques communes. La variable vitesse a été choisie comme deuxième levier d'ajustement opérationnel accessible à l'opérateur, car elle conditionne le comportement dynamique de la machine, la puissance

délivrée, ainsi que les pertes énergétiques par frottement ou inertie. Ces deux variables sont donc des variables d'action, directement modifiables dans un contexte d'optimisation en temps réel, ce qui en fait des cibles naturelles pour un système d'aide à la décision.

Type de cargo : cette variable représente le type de matière déchargée. Bien qu'elle ne soit pas une variable technique, elle est intégrée comme variable explicative catégorielle afin de tenir compte des propriétés physico-mécaniques propres à chaque matériau transporté (densité, abrasivité, granularité), qui influencent la résistance au transport et la consommation énergétique. Contrairement aux deux premières, cette variable n'est pas un levier de contrôle, mais un contexte constraint de l'opération, car elle permet d'adapter dynamiquement les opérations à la situation réelle rencontrée.

3.3.2 Méthode L-BFGS-B utilisée pour l'optimisation.

Dans le cadre de ce projet, l'objectif est de minimiser la consommation énergétique prédictive par l'algorithme, en fonction des variables continues et catégorielle : le débit de déchargement (en tonnes par heure) et la vitesse de rotation du moteur (en tours par minute), le type de matériau déchargé, car celle-ci influence la consommation de l'énergie comme nous l'avons observée lors de l'EAD. Ces variables sont soumises à des contraintes de bornes physiques, correspondant aux limites opérationnelles des convoyeurs. Le modèle de prédiction n'étant pas exprimé analytiquement, mais issu d'un apprentissage supervisé, l'optimisation repose sur une fonction de type boîte noire, non dérivable formellement.

Pour résoudre ce problème, nous avons retenu l'algorithme L-BFGS-B, implémenté dans la fonction *scipy.optimize.minimize*. Cette méthode d'optimisation est spécifiquement conçue pour traiter des fonctions non linéaires sous contraintes de bornes sur les variables de décision. Elle repose sur une approche quasi-Newtonienne, qui permet d'approximer efficacement la matrice hessienne tout en limitant la consommation de mémoire. Cette caractéristique la rend particulièrement adaptée aux problèmes de petite dimension, comme celui étudié dans notre cas.

L-BFGS-B est réputée pour sa stabilité numérique, sa rapidité et sa capacité à respecter rigoureusement les contraintes imposées sur les variables. Bien qu'elle soit généralement utilisée pour des fonctions dérivables, elle a démontré empiriquement de bonnes performances dans le cadre d'optimisations sur des fonctions apprises par des modèles de machine learning, comme les forêts aléatoires. Ce choix a donc été motivé par des considérations à la fois théoriques et pratiques : robustesse, compatibilité avec une fonction non explicite, et respect strict des bornes imposées.

Cette méthode a été décrite notamment dans les travaux de Byrd et al. (1995) et Zhu et al. (1997) et constitue une référence classique dans le domaine de l'optimisation numérique des fonctions à variables continues contraintes.

3.3.3 Formulation mathématique du problème d'optimisation

Dans le cadre de l'optimisation des opérations de déchargement d'un convoyeur à bande, nous cherchons à minimiser la consommation d'énergie associée à ce processus. Cette consommation dépend notamment de deux variables opérationnelles : le taux de déchargement T exprimé en tonne par heure, et la vitesse de déchargement V. L'objectif est donc de déterminer les valeurs optimales de ces deux variables qui permettent de minimiser la consommation énergétique tout en respectant les contraintes opérationnelles.

Le problème peut être formulé dans le cadre standard d'un programme d'optimisation constraint. Soit $X = (V, T) \in \mathbb{R}^2$ le vecteur des variables de décision. La fonction objectif $E(X)$ représente la consommation d'électricité prédictive par le modèle de forêt aléatoire (en kW) en fonction des valeurs choisies de V et T. On cherche donc à résoudre le problème suivant : $\min E(X) = \text{Consommation d'électricité prédictive en kW}$

$$E(X) = \hat{y}(V, T) \quad (3-1)$$

$$\min_{X \in \mathbb{R}^2} E(X) \quad (3-2)$$

Sous contraintes : $V_{min} \leq V \leq V_{max}$ (3-3)

$$T_{min} \leq T \leq T_{max} \quad (3-4)$$

Ces bornes représentent les plages minimale et maximale observées dans le jeu de donnée pendant le fonctionnement. Le modèle $E(X)$ n'a pas d'expression mathématique, car elle est estimée par l'algorithme de prédiction. Le problème d'optimisation est résolu de manière numérique par l'algorithme L-BFGS-B. Cette approche permet d'adapter dynamiquement les recommandations selon les types de cargaisons (sel, minerai, céréales, etc.), tout en tenant compte de la variabilité opérationnelle et des spécificités physiques des matériaux. La démarche suivie est la suivante :

1. **Définition des bornes** : Les plages de variation de V et T sont déterminées à partir des valeurs minimales et maximales observées dans le jeu de données réel. Ces bornes sont cruciales pour garantir que les solutions proposées restent dans un domaine physiquement réalisable.
2. **Initialisation du problème** : Pour chaque cargaison et chaque navire, une solution initiale est fixée dans l'intervalle de recherche, qui est généralement le centre des bornes.
3. **Évaluation de la fonction objective** : À chaque itération, l'algorithme propose une nouvelle combinaison (V,T) qui est évaluée par le modèle de prédiction pour obtenir la consommation d'énergie associée.
4. **Mise à jour itérative** : L-BFGS-B utilise des approximations des dérivées partielles pour construire une direction de descente, tout en respectant les contraintes de bornes. Il poursuit jusqu'à la convergence vers un minimum local.
5. **Vérification de validité** : Enfin, les solutions sont filtrées pour ne retenir que celles qui respectent une contrainte supplémentaire liée au temps maximal de déchargement ($\leq 3h$), afin de garantir la faisabilité opérationnelle.

Nous avons donc une approche qui permet d'obtenir rapidement, pour chaque cargo, des paramètres de fonctionnement optimal adapté, avec une consommation énergétique minimisée selon la contrainte de temps. Elle est flexible et reproductible et peut être réappliquée à d'autres navires ou d'autres jeux de données en modifiant simplement les bornes et les modèles prédictifs.

3.4 Conclusion

Dans ce chapitre nous avons présenté de manière détaillée l'ensemble des algorithmes utilisés pour réaliser l'optimisation de la consommation de l'énergie. Notre modèle basé sur l'apprentissage automatique effectue cette optimisation combinant un modèle de prédiction via l'algorithme de forêt aléatoire et une optimisation numérique via la méthode L-BFGS-B, notre approche permet de proposer des configurations opérationnelles efficaces et adaptées à chaque type de cargaison. Elle constitue ainsi une base solide pour le développement d'un système d'aide à la décision en exploitation portuaire.

CHAPITRE 4

RÉSULTATS ET DISCUSSION

4.1 Introduction

Ce chapitre présente et analyse les résultats issus des différentes étapes méthodologiques développées dans le cadre de ce mémoire. L'objectif principal est d'évaluer la pertinence des modèles élaborés, tant pour la prédiction de la consommation énergétique des convoyeurs à bande que pour l'optimisation de leurs paramètres de fonctionnement, dans une perspective d'aide à la décision opérationnelle.

Dans un premier temps, les résultats de l'analyse exploratoire des données (EAD) sont exposés, permettant de mettre en évidence les principales tendances et de justifier les choix de modélisation adoptés. Ensuite, les performances du modèle de prédiction, basé sur l'algorithme de la forêt aléatoire, sont évaluées de manière globale, puis affinées en tenant compte des spécificités liées aux différents types de cargos. Cette analyse différenciée permet de révéler l'hétérogénéité des comportements énergétiques selon la nature du matériau déchargé.

Par la suite, les résultats de l'optimisation à l'aide de l'algorithme L-BFGS-B sont présentés selon deux volets complémentaires : d'une part, une optimisation globale indépendante du type de cargo ; d'autre part, une optimisation spécifique intégrant les particularités propres à chaque matière.

Enfin, une discussion critique est proposée afin de mettre en perspective l'ensemble des résultats obtenus, d'en identifier les implications pratiques et les limites, et de formuler des recommandations opérationnelles ainsi que des pistes pour de futurs travaux.

4.2 Résultat des analyses exploratoires des données

4.2.1 Gestion des valeurs manquantes

La base de données mise à notre disposition provient des informations collectées pour 4 navires : Navire 1, Navire 2, Navire 3, Navire 4. Ces données ont été chargées de manière indépendante dans notre environnement d'exécution. Les Figures 4.1 et 4.2 suivantes nous permettent de visualiser les données téléversées et d'observer la présence des valeurs manquantes pour chaque variable en fonction des différents navires.

Transfer master motor temp	0
Transfer slave motor temp	0
Transfer master motor speed	0
Transfer slave motor speed	0
Tunnel master motor power	0
Tunnel slave motor power	0
Tunnel master motor temp	0
Tunnel slave motor temp	0
Tunnel master motor speed	0
Tunnel slave motor speed	0
AE1	0
AE2	0
AE3	0
House usage	0
Discharge rate	0
Electricity consumption	626
Voyage	1
Cargo	1
Port	1
dtype: int64	
	date Electricity consumption Voyage
count	39262 38637.000000 39262.000000
mean	2021-09-29 13:29:17.294585088 51.575754 21044.499032
min	2020-01-02 21:24:00 -1655.420388 19036.000000
25%	2020-12-03 17:11:15 -0.041542 20036.000000
50%	2021-10-18 19:05:30 12.372502 21037.000000
75%	2022-08-09 07:01:45 149.941932 22016.000000
max	2023-06-01 11:53:00 1388.152293 23008.000000
std	Nan 328.002993 1126.280123

Figure 4.1 Aperçu global du jeu de données du navire 1

boom master motor power	0
boom slave motor power	0
boom master motor temp	0
boom slave motor temp	0
boom master motor speed	0
boom slave motor speed	0
boom inclination angle	0
Loop master motor power	0
Loop slave motor power	0
Loop master motor temp	0
Loop slave motor temp	0
Loop master motor speed	0
Loop slave motor speed	0
Transfer master motor power	0
Trasfer slave motor power	0
Transfer master motor temp	0
Transfer slave motor temp	0
Transfer master motor speed	0
Transfer slave motor speed	0
Tunnel master motor power	0
Tunnel slave motor power	0
Tunnel master motor temp	0
Tunnel slave motor temp	0
Tunnel master motor speed	0
Tunnel slave motor speed	0
AE1	0
AE2	4036
AE3	4036
House usage	4036
Discharge rate	4036
Electricity consumption	0
Voyage	0
Cargo	0
Port	0

Figure 4.2 Aperçu global du jeu de données du navire 3

L'analyse inter-fichiers a révélé une distribution hétérogène des valeurs manquantes : le navire 3 présente des absences importantes sur le débit de déchargement, alors que le navire 1 et le navire 2 en sont exempts. À l'inverse, les fichiers de ces derniers présentent des lacunes sur la

variable cible « **Consommation électrique** ». Cela pourrait s'expliquer par l'absence d'activité, l'erreur de lecture du capteur, des données non enregistrées par des capteurs ou procédures différentes selon les sites ou convoyeurs concernés. En conséquence, un filtrage adapté à chaque source a été appliqué pour garantir la cohérence des données utilisées dans les modèles de prédiction.

La variable « **Discharge rate** (débit de déchargement) » est un facteur explicatif central influant directement sur la consommation électrique. Or, la variable « **Consommation électrique** » est la cible de notre modèle de prédiction. Afin de garantir la qualité de l'apprentissage, toutes les lignes présentant une valeur manquante sur l'une ou l'autre de ces deux variables ont été supprimées. Cette décision est justifiée par la forte interdépendance entre ces deux paramètres, comme le montre la Figure 4.3 ci-dessous, rendant leur séparation incohérente dans le cadre d'une modélisation supervisée. De plus ces valeurs manquantes représentent moins de 10% sur l'ensemble des données, donc, malgré cette suppression, le jeu de données reste fiable et conséquent pour l'apprentissage du modèle.

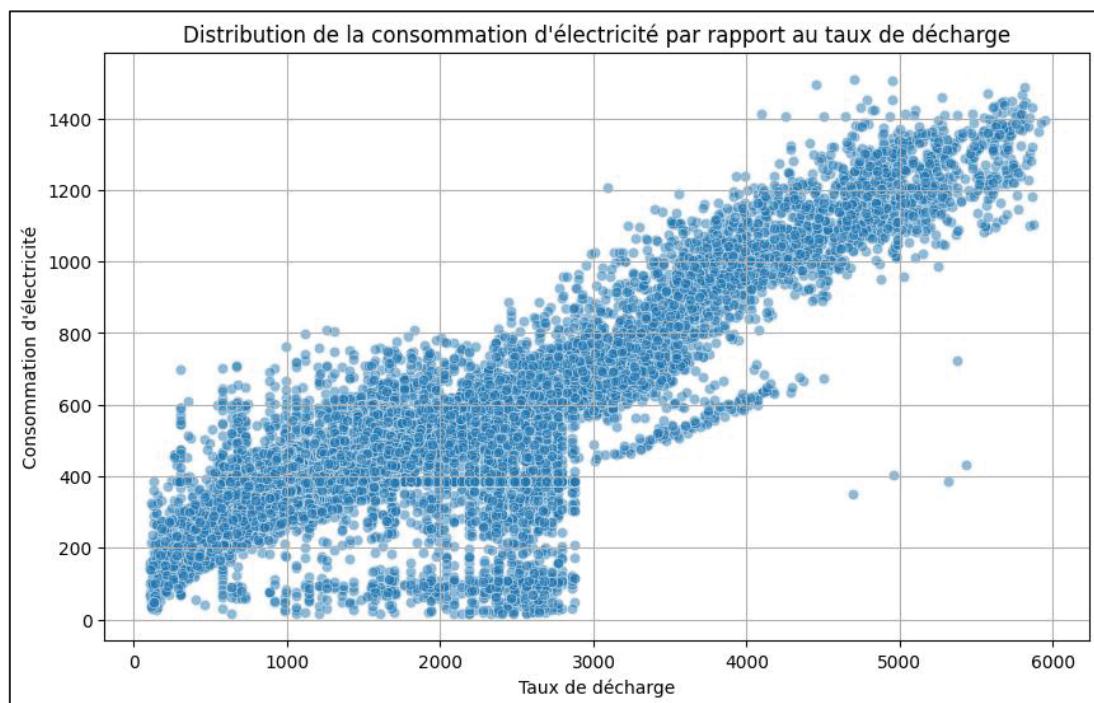


Figure 4.3 Distribution de la consommation d'électricité par rapport au taux de décharge du navire 3

Le navire 4 sera exclu de l'analyse, car après explorations de son jeu de données, nous avons constaté que celui-ci ne possède aucune information sur la variable cible qui est « **Consommation électrique** » comme nous pouvons observer sur la Figure 4.4 suivante :

```
print(df.columns)
print(df.shape) # Renvoie le tuple (nombre de lignes, nombre de colonnes)

Index(['date', 'boom master motor power ', 'boom slave motor power',
       'boom master motor temp', 'boom slave motor temp',
       cellules de code master motor speed', 'boom slave motor speed',
       'boom inclination angle', 'Loop master motor power ',
       'Loop slave motor power ', 'Loop master motor temp',
       'Loop slave motor temp', 'Loop master motor speed',
       'Loop slave motor speed', 'Transfer master motor power ',
       'Transfer slave motor power ', 'Transfer master motor temp',
       'Transfer slave motor temp', 'Transfer master motor speed',
       'Transfer slave motor speed', 'Tunnel master motor power',
       'Tunnel slave motor power ', 'Tunnel master motor temp',
       'Tunnel slave motor temp', 'Tunnel master motor speed',
       'Tunnel slave motor speed', 'Unnamed: 26', 'Unnamed: 27', 'Unnamed: 28',
       'Discharge rate '],
      dtype='object')
(41504, 30)
```

Figure 4.4 Variables du navire 4

D'après cette figure, la variable « **Consommation électrique** » est absence dans le jeu de donnée, ce qui pourrait rendre impossible la modélisation de l'algorithme de prédiction selon le principe d'apprentissage. De plus nous avons également le manque d'information sur le type de cargaison qui est déchargée rendant la modélisation contextuelle impossible au niveau de l'optimisation de la consommation de l'énergie. Par ailleurs nous observons que le fichier comporte 30 colonnes au lieu de 35. Ce manque d'informations critique rend impossible l'analyse de ce navire.

4.2.2 Traitement des valeurs aberrantes

Lors de l'analyse descriptive de chaque variable, nous avons observé des valeurs extrêmement négatives comprises entre [-750, 0] pour certains navires et d'autres allant jusqu'à -1500 sur la variable « **Consommation électrique** ». Ces valeurs représentent des consommations

énergétiques inférieures à zéro, ce qui est irréaliste. Ces valeurs aberrantes peuvent fausser et induire en erreur le processus de formation de l'apprentissage automatique, ce qui se traduit par des temps de formation moins précis et plus longs et des résultats plus médiocres.

La Figure 4.5 ci-dessous illustre la distribution de la variable « **Consommation électrique** » et la détection des valeurs aberrantes à l'aide de la méthode de l'écart interquartile (l'IQR). Les lignes pointillées de couleur orange indiquent ces bornes limites : toute valeur située en dehors de cet intervalle est considérée comme potentiellement aberrante. En complément, une ligne violette verticale marque le seuil zéro, car les valeurs de consommation inférieures à zéro sont physiquement impossibles dans ce contexte et ont donc été automatiquement classées comme aberrantes. La zone bleue représente les observations considérées comme normales, tandis que la zone rouge correspond aux valeurs aberrantes détectées selon la méthode.

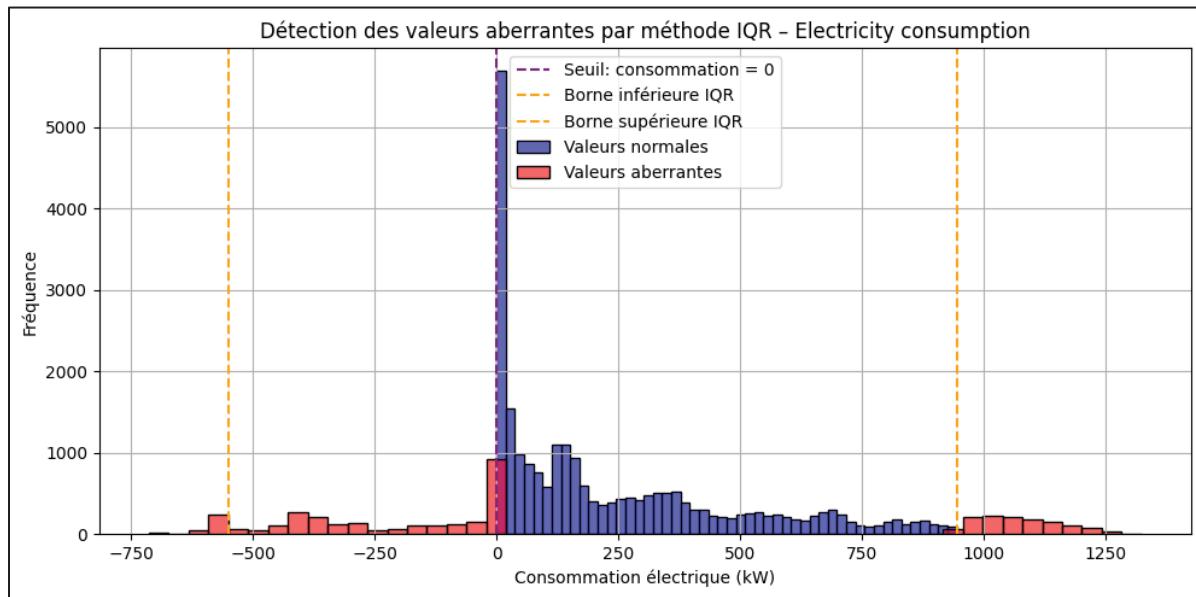


Figure 4.5 Détection des valeurs aberrantes par la méthode de l'écart interquartile (IQR)

Un nettoyage spécifique est appliqué à cette variable, notamment la suppression des valeurs négatives et des valeurs nulles. En revanche nous avons décidé de garder le seuil de consommations extrêmes positives, car il pourrait contenir des informations. Le choix des

seuils de nettoyage des données est justifié par les réalités physiques et opérationnelles du système étudié.

Les observations dont la consommation électrique est inférieure à 15 kW ont été supprimées, car elles correspondent à des situations non représentatives d'un fonctionnement actif du système. En pratique, une consommation aussi faible peut refléter des phases de veille ou d'arrêt, des lectures erronées provenant de capteurs défectueux, ou encore des artefacts liés à la transmission ou l'enregistrement des données. Ces données sont donc peu utiles pour entraîner un modèle destiné à prédire la consommation en fonctionnement normal, et leur maintien risquerait de biaiser le modèle vers des comportements énergétiques non pertinents pour les opérateurs.

Appliquer ces 2 actions à la fois, c'est-à-dire supprimer toutes les valeurs négatives et celle inférieure à 15kW, revenait à perdre près de 50% de la quantité globale du jeu de données, ce qui est énorme et pourrait affecter la précision du modèle d'apprentissage plus tard. Nous avons donc dans un premier temps remplacé les valeurs négatives par la médiane positive et ensuite supprimer les valeurs inférieures à 15kW. Cette action nous a permis de garder une quantité de données considérable pour notre modèle de prédiction. La Figure 4.6 permet une visualisation avant après du traitement de ces valeurs aberrantes.

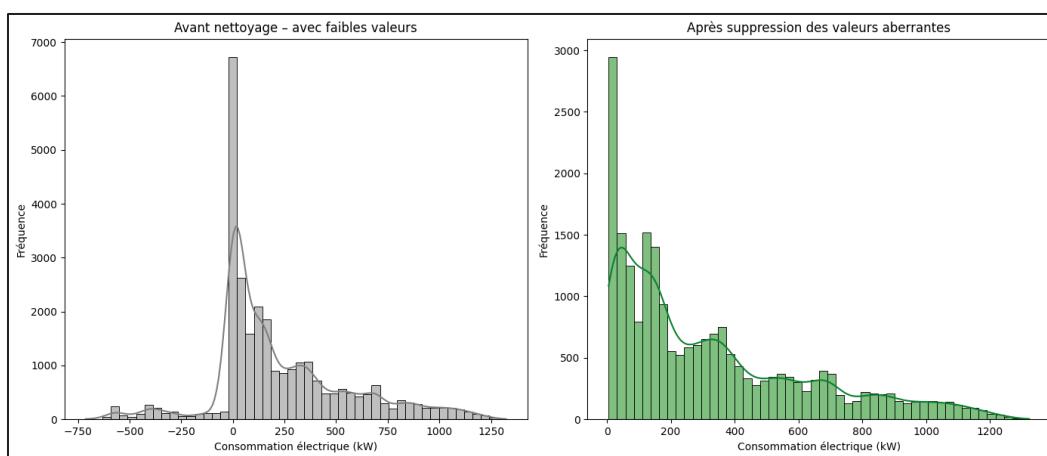


Figure 4.6 Comparaison des données après nettoyage

La figure ci-dessus représente une comparaison entre le jeu de données initiales et celui après nettoyage. Cependant, les valeurs extrêmes positives ont été volontairement conservées, car elles peuvent correspondre à des pics de consommation normaux correspondant à des charges exceptionnellement lourdes ou à des accélérations mécaniques plus importantes. Elles peuvent révéler des situations critiques ou des scénarios d'usage intensif importants à prendre en compte dans l'optimisation énergétique. Elles sont physiquement possibles dans le contexte de déchargement de matières lourdes, comme le minerai ou le sel. La présence de ces valeurs extrêmes permet donc de modéliser avec précision la variabilité opérationnelle réelle et de proposer des recommandations robustes, même en cas de fonctionnement à haute intensité.

Le même principe a été appliqué par la suite sur chaque fichier de données et en s'appuyant sur un recalculation des valeurs aberrantes, nous avons pu vérifier que celles-ci ont été correctement traitées. La Figure 4.7 ci-dessous confirme donc l'efficacité de la méthode appliquée sur l'ensemble du jeu de données du navire 2.

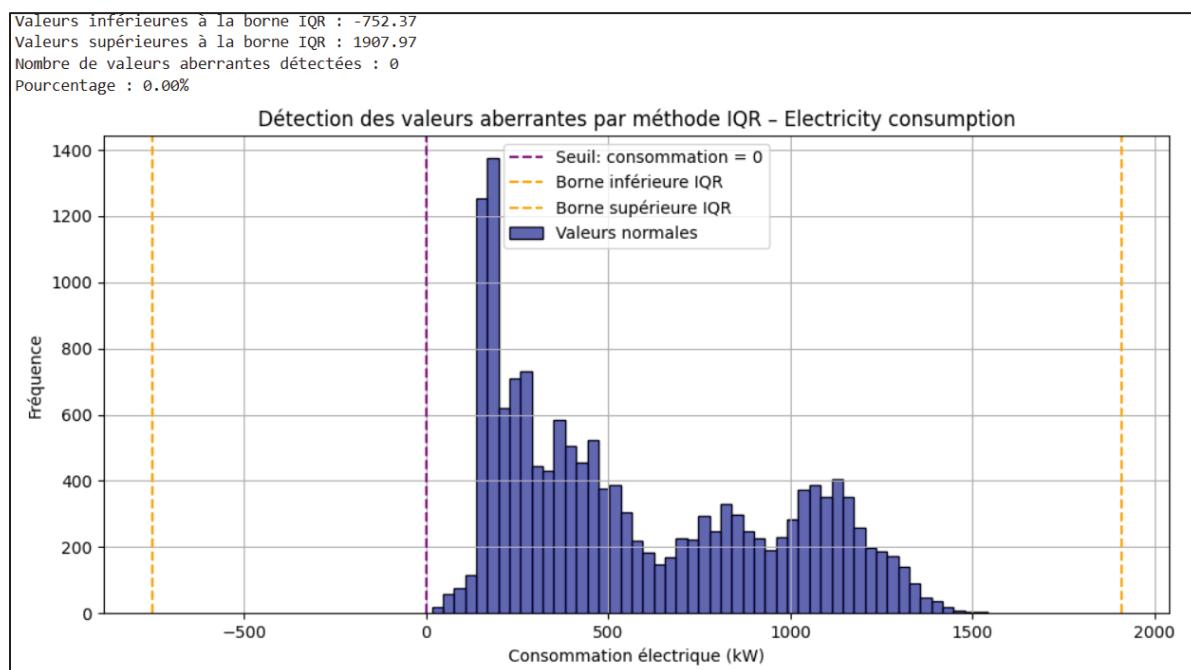


Figure 4.7 Vérification des données après traitement des valeurs aberrantes

Cette même opération a été effectuée sur l'ensemble des variables comportant le jeu de données, notamment le taux de décharge et les différentes vitesses et puissances d'entraînement.

4.2.3 Analyse des relations entre les variables

Cette analyse consiste à évaluer la relation qui existe entre les variables afin d'identifier celles qui influencent directement ou indirectement la variable « **Consommation de l'énergie** ». Pour cela, nous avons calculé la matrice de corrélation représentée par la Figure 4.8. Cette matrice nous permet de visualiser une corrélation très forte entre la variable « **Consommation de l'énergie** » et « **discharge rate** » avec une valeur de $r = 0.97$ et la variable « **Boom master motor speed** » avec une valeur de $r = 0.81$. Sachant que r représente le coefficient de corrélation entre 2 variables, plus cette valeur se rapproche de 1, plus elle démontre la forte dépendance qui existe entre les deux variables. Cette forte corrélation indique que l'augmentation du débit de déchargement ou de la vitesse des moteurs est généralement associée à une élévation de la consommation d'énergie. Cela est intuitivement cohérent avec le fonctionnement mécanique des convoyeurs dans la mesure où plus le système travaille vite ou fort (en puissance), plus il consomme de l'électricité.

Un coefficient de corrélation égal à 0 signifie qu'il n'existe aucune relation entre les 2 variables.

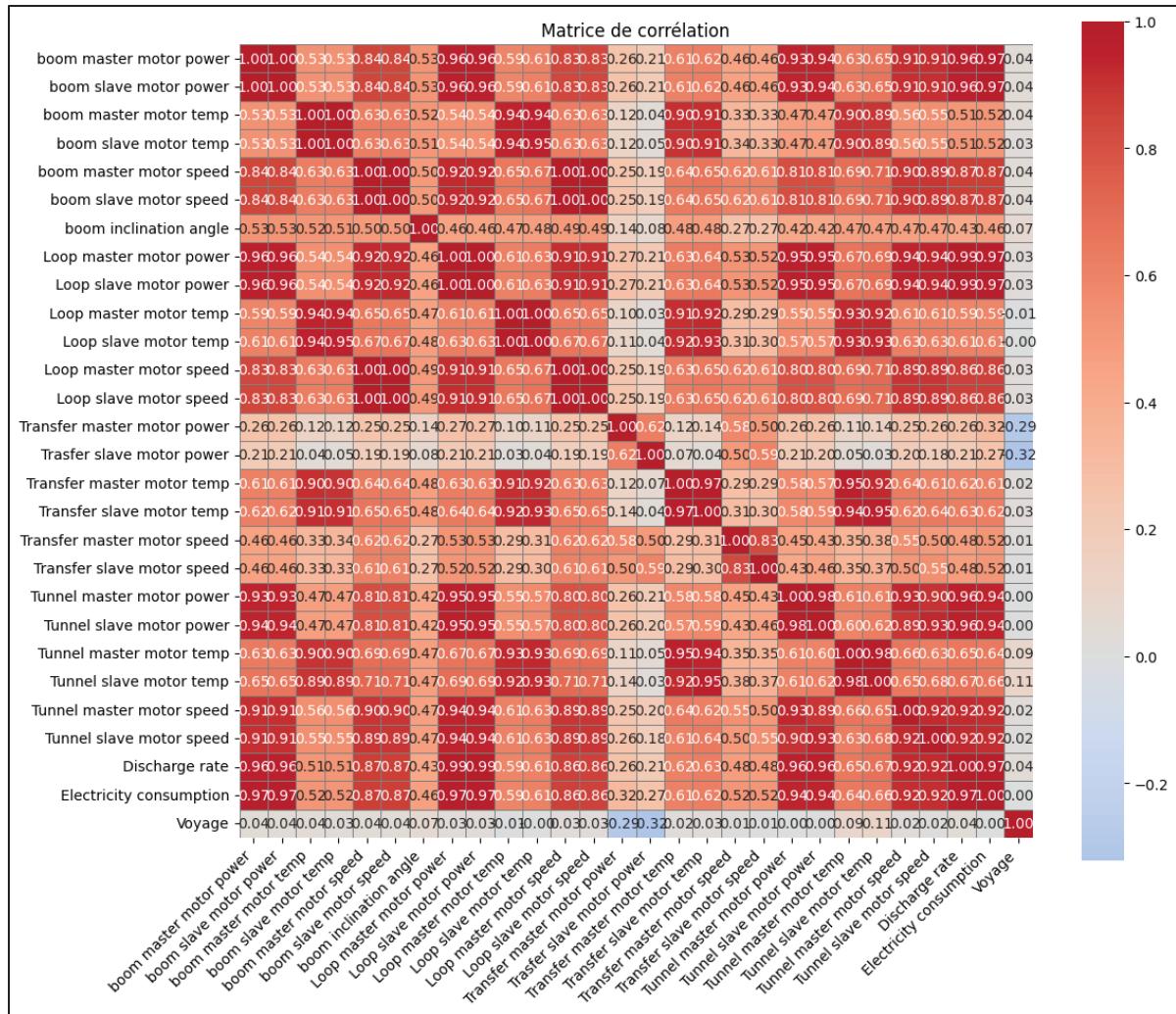


Figure 4.8 Matrice de corrélation du navire 2

De même, nous pouvons également observer une forte corrélation $r = 1$ entre « **boom master motor power** » et « **boom slave motor power** » suggérant une relation directe dans leur comportement. Cette corrélation est observée pour chaque moteur et son esclave (BOOM, TUNNEL, LOOP, TRANSFER). Dans ce cas, il sera préférable pour notre modèle d'utiliser qu'une seule dans le modèle pour réduire la redondance.

La Figure 4.9 suivante permet de visualiser la relation linéaire qui existe entre le **taux de décharge** et la **consommation d'électricité** après la suppression des valeurs aberrantes.

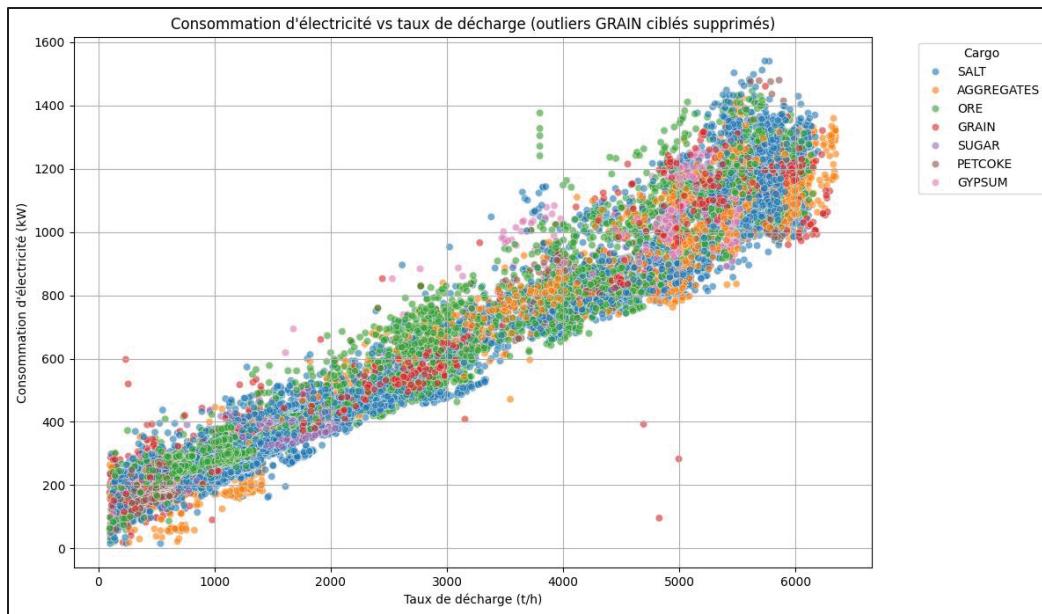


Figure 4.9 Diagramme de dispersion nettoyé de la consommation d'électricité en fonction du taux de décharge du navire 2

- Analyse de la consommation de l'énergie selon le type de cargo**

Le diagramme en boîte présenté par la Figure 4.10 ci-dessous illustre la répartition statistique de la consommation électrique (en kW) selon le type de cargo transporté. Cette visualisation permet d'évaluer la dispersion, la médiane, et la présence de valeurs extrêmes pour chacun des matériaux, fournissant ainsi des indications sur les comportements énergétiques spécifiques.

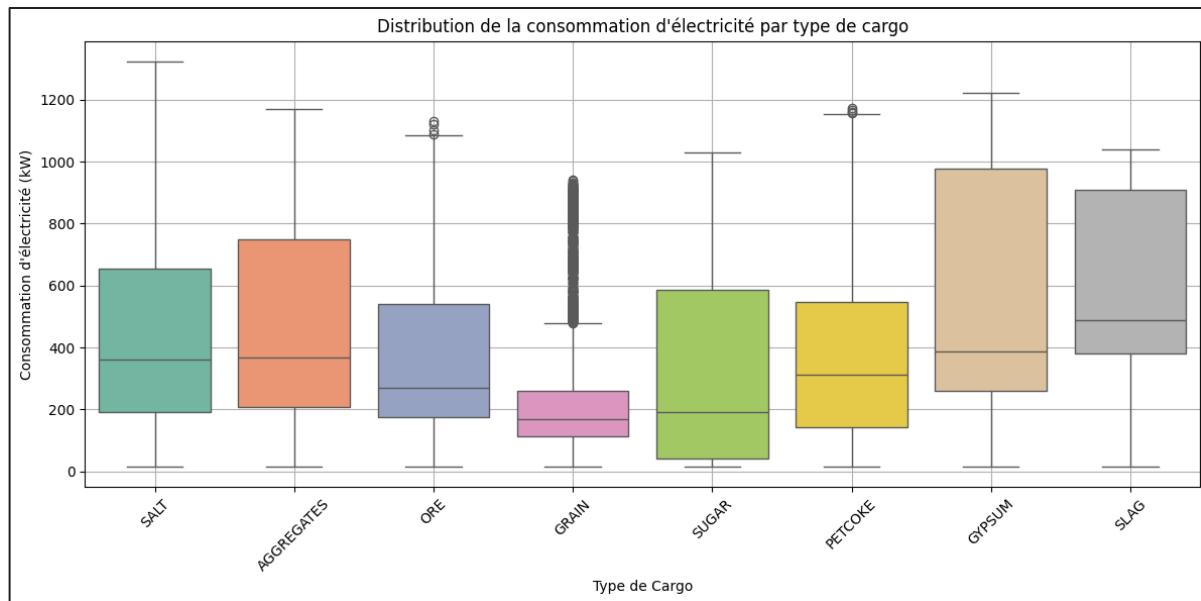


Figure 4.10 Distribution de la consommation d'électricité

On constate que le type de cargaison a clairement un impact significatif sur la consommation énergétique, car chaque cargaison possède des propriétés physiques uniques, telles que le poids, la taille et la composition des matériaux qui influencent directement leur densité, ce qui doit être pris en compte dans les stratégies d'optimisation.

Nous observons que certaines cargaisons, comme GYPSUM et SLAG possèdent une médiane plus élevée que les autres, par conséquent consomment plus d'énergie, ce qui peut être lié aux conditions de manipulation ou à la densité du matériau. En particulier, le déchargement du gypsum présente la plus grosse dispersion. À l'inverse, les cargos GRAIN et SUGAR présentent des médianes plus faibles et une dispersion modérée, ce qui laisse entendre une consommation plus stable et relativement modérée lors de leurs déchargements. Cette même variation est observable sur le navire 3 dont la distribution de la consommation d'électricité est présentée à l'annexe II.

- Analyse de la consommation de l'énergie selon le port**

Elle permet de visualiser la consommation de l'énergie sur les différents ports où les navires effectuent des déchargements. Chaque port possède des caractéristiques propres

(infrastructures, conditions d'accès, temps d'attente, types de quais, etc.) qui peuvent influencer la consommation énergétique lors du déchargement. La Figure 4.11 ci-dessous illustre les différentes consommations selon le type de port.

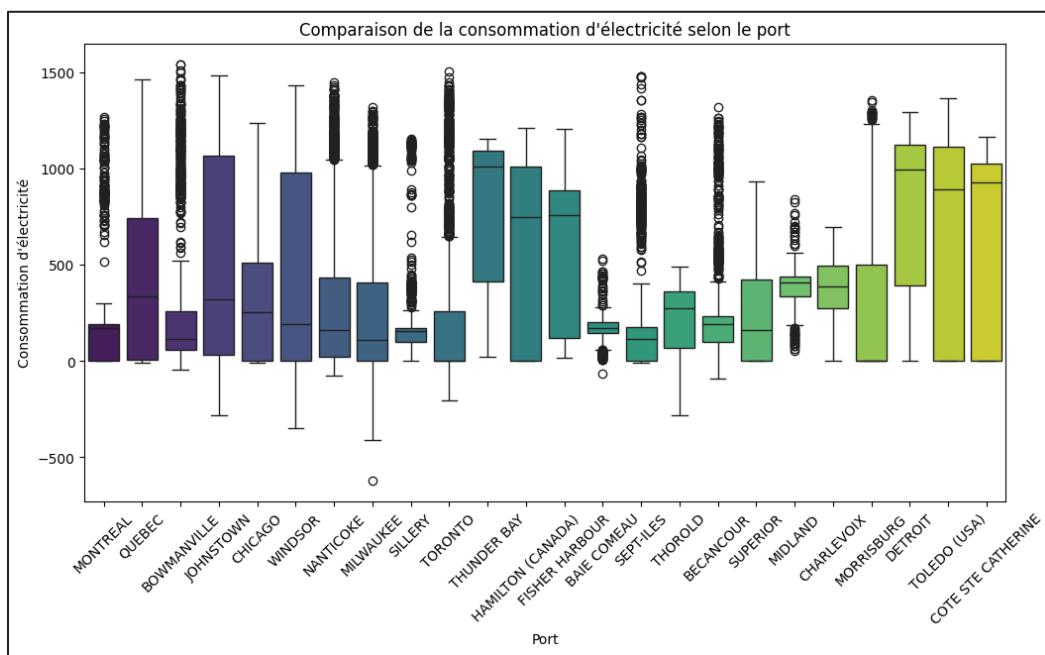


Figure 4.11 Comparaison de la consommation d'énergie selon le port

Nous observons que les ports avec une consommation énergétique élevée et variable, comme **THUNDER BAY**, **TOLEDO (É.-U.)** qui nécessitent une attention particulière pour comprendre les facteurs responsables (type de cargaison, conditions logistiques, infrastructures énergétiques, etc.).

4.3 Résultat du modèle de prédiction forêt aléatoire

Après l'étape de prétraitement des données, les jeux de données correspondant aux différents navires ont été intégrés individuellement dans l'algorithme de prédiction. Cette section présente les résultats de prédiction obtenus après l'entraînement et la validation du modèle de forêts aléatoires.

Conformément à la méthodologie, les résultats sont exposés selon deux approches distinctes. Dans un premier temps, le modèle a été entraîné sur l'ensemble des données sans distinction de type de cargo, ce qui constitue une analyse globale. Ensuite, une seconde série de prédictions a été réalisée en tenant compte du type de matériau à décharger, dans le but de mieux capturer les spécificités opérationnelles associées à la nature des matériaux manipulés. Cette démarche est justifiée par l'analyse exploratoire des données, qui a mis en évidence l'influence significative de ce facteur sur la consommation énergétique.

4.3.1 Prédiction globale

Suivant notre méthodologie, les résultats de l'évaluation du modèle de prédiction global, entraîné sur l'ensemble des observations sans distinction explicite du type de cargo, sont représentés dans le Tableau 4.1 avec les métriques de performances suivantes.

Tableau 4.1 Métriques de performance pour les prédictions globales

Navires	R ² train	R ² test	Train RMSE	Test RMSE
Navire 1	0.9963	0.9807	14.05	32.51
Navire 2	0.999	0.9930	12.27	32.72
Navire 3	1	1	0.05	0.11

De manière générale, sur l'ensemble des résultats pour les différents navires, nous pouvons observer que le modèle de prédiction a une très bonne capacité de généralisation. Il a su capter les relations non linéaires entre les différentes variables explicatives, comme le taux de décharge, la puissance des moteurs, leur vitesse sans oublier l'angle d'inclinaison de la flèche afin de prédire correctement la consommation de l'énergie. Cette performance est observée sur la valeur du R² qui est supérieur à 0.98 que ce soit sur les données d'entraînement que sur les données test.

Nous observons également un RMSE faible, particulièrement pour le navire 3 qui se rapproche de zéro. Ces résultats presque parfaits ont été réévalués avec une validation croisée afin de limiter le surapprentissage. Les résultats de cette manœuvre présentée dans l'Annexe III

confirme l'exactitude des performances du modèle. Ces métriques permettent donc de confirmer que le modèle de forêt aléatoire est adapté à ce type de données pour prédire la consommation de l'énergie.

Pour mieux visualiser ces performances, nous avons fait une représentation 3D à travers la Figure 4.12 de la consommation de l'énergie réelle et celle prédictive afin de comparer les différentes distributions. Les figures sont très semblables, traduisant une bonne représentation des dynamiques globales du système.

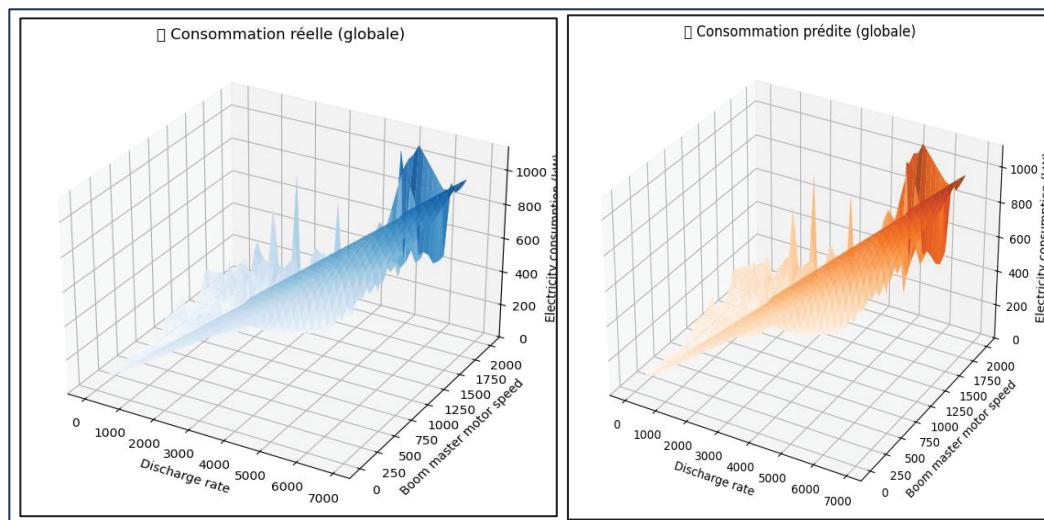


Figure 4.12 Comparaison des diagrammes de la consommations globale réelle vs celle prédictive

- *Discharge rate* représente le taux de déchargement
- *Boom master motor speed* est la vitesse du moteur principal de la flèche

On constate que la distribution de la consommation globale prédictive est quasiment identique à celle de la consommation réelle, ce qui démontre que l'algorithme de prédiction fonctionne normalement.

4.3.2 Prédition selon le type de cargo

Une deuxième série d'évaluations a été conduite selon le type de cargaison, permettant ainsi d'analyser plus finement les performances du modèle dans des contextes spécifiques. Les

Tableaux 4.2 à 4.4 présentent les métriques de performance du modèle de forêt aléatoire pour chaque type de matériau transporté, respectivement pour les navires 1, 2 et 3.

Tableau 4.2 Résultat du modèle de forêt aléatoire pour le navire 1

Cargo	N_train échantillon	N_test échantillon	R ² train	R ² test	Train RMSE	Test RMSE
SALT	8801	2201	0.9951	0.9778	21.84	46.63
AGGREGATES	4964	1242	0.9988	0.9922	4.00	10.15
PETCOKE	4283	1071	0.9950	0.9727	6.67	16.48
ORE	1912	479	0.9981	0.9906	4.04	9.12
GYPSUM	2592	648	0.9978	0.9837	8.05	21.67
GRAIN	1267	317	0.9988	0.9944	3.70	7.80

Tableau 4.3 Résultat du modèle de forêt aléatoire pour le navire 2

Cargo	N_train échantillon	N_test échantillon	R ² train	R ² test	Train RMSE	Test RMSE
SALT	7467	1867	0.9988	0.9916	13.78	35.57
AGGREGATES	1458	365	0.9989	0.9940	14.47	34.88
ORE	4416	1104	0.9993	0.9937	10.18	29.84
GRAIN	2955	739	0.9984	0.9874	12.34	34.02
SUGAR	867	217	0.9977	0.9898	11.07	21.83
PETCOKE	977	245	0.9992	0.9869	8.49	34.88
GYPSUM	230	58	0.9979	0.9855	17.45	42.39

Tableau 4.4 Résultat du modèle de forêt aléatoire pour le navire 3

Cargo	N_train échantillon	N_test échantillon	R ² train	R ² test	Train RMSE	Test RMSE
SALT	1081	271	0.9986	0.9923	14.90	34.39
GYPSUM	1261	316	0.9980	0.9847	15.19	40.57
PETCOKE	2490	623	0.9965	0.9785	8.67	23.42
GRAIN	3225	807	0.9939	0.9751	8.92	17.79
AGGREGATES	4802	1201	0.9950	0.9650	12.60	32.81
SUGAR	91	23	0.9310	0.9539	19.24	13.71
ORE	184	46	0.9831	0.7262	16.48	65.92

Ces différents résultats confirment une fois de plus les bonnes performances du modèle, malgré les ajustements apportés en fonction du type de cargo. En fonction du nombre d'échantillons disponibles, tant pour l'entraînement que pour les tests, la valeur du R² varie entre 0,97 et 0,99. Par exemple, pour le navire 2, le modèle atteint un R² de 0,9937 sur les données de test pour le cargo de type ORE, et de 0,9916 pour le cargo SALT. Par ailleurs, on observe une différence notable entre les valeurs de RMSE sur les jeux d'entraînement et de test. Cette divergence suggère que le modèle ne souffre pas de surapprentissage, ce qui renforce la validité de ses performances.

La Figure 4.13 illustre la similarité entre la consommation d'énergie prédictive pour le cargo GYPSUM du navire 2 et sa consommation réelle. Cette concordance visuelle confirme la performance du modèle de prédiction.

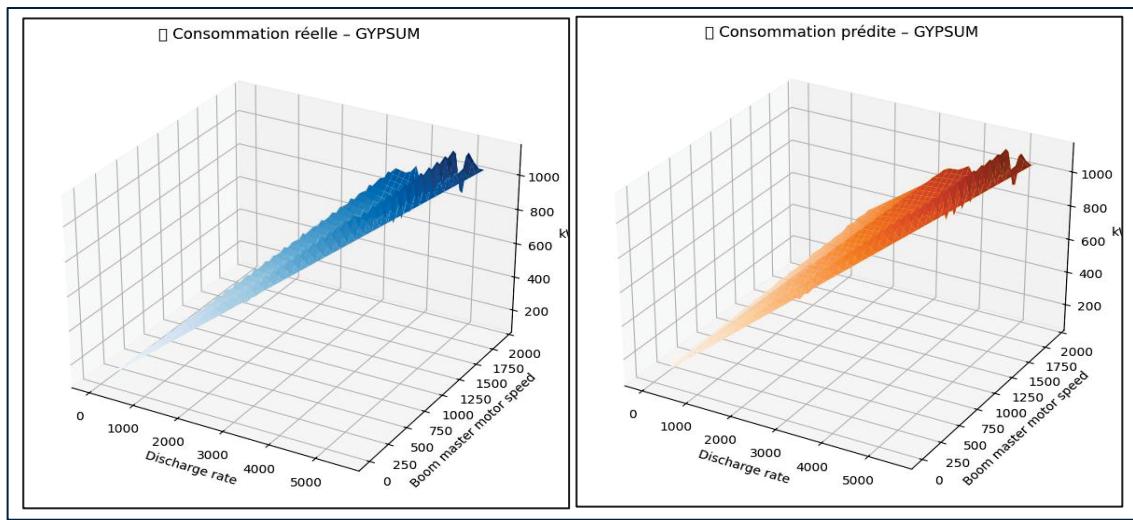


Figure 4.13 Visualisation comparative de la prédition pour le cargo GYPSUM du navire 2

De même, la Figure 4.14 confirme l'efficacité du modèle de prédition en illustrant que la consommation d'énergie prédicté pour le cargo SALT à bord du navire 3 est très proche de sa consommation réelle.

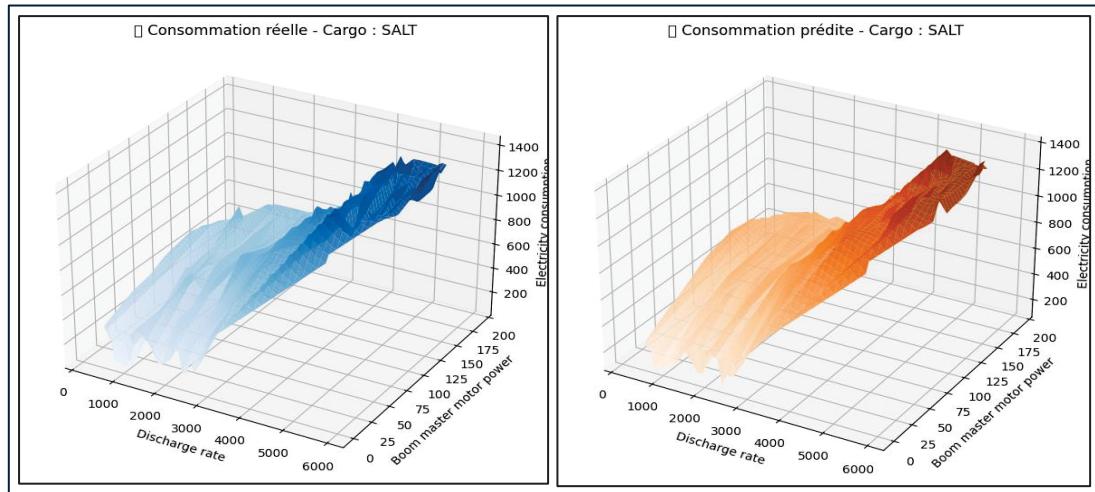


Figure 4.14 Visualisation comparative de la prédition pour le cargo SALT du navire 3

4.3.3 Limite observée Cas du cargo ORE du navire 3

Une baisse de précision est observée dans le cas du cargo ORE du navire 3. Comme l'illustre la Figure 4.15, la courbe de consommation d'énergie prédictive s'écarte de manière significative de la courbe réelle. Le coefficient de détermination R^2 sur les données de test chute à 0,7262, tandis que le RMSE atteint 65,92, ce qui témoigne d'une diminution notable de la performance du modèle dans ce cas. Les deux figures suivantes permettent de comparer visuellement la consommation d'énergie réelle et celle prédictive pour le cargo ORE du navire 3.

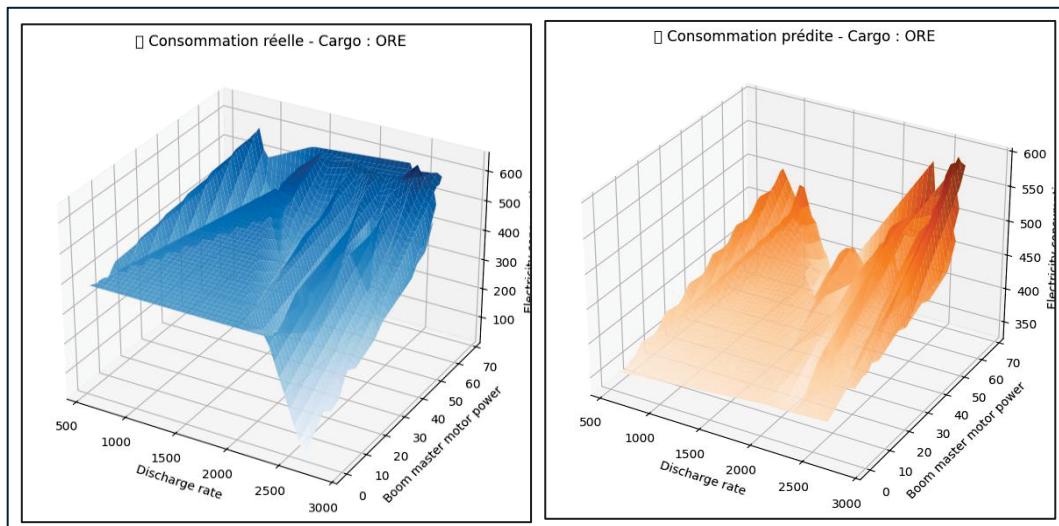


Figure 4.15 Visualisation comparative de la prédition du cargo ORE du navire 3

Il est clair de constater de manière visuelle que les courbes ne sont pas semblables. Cette différence pourrait s'expliquer par la dispersion énorme de l'échantillon de données traité. Cette disparité est renforcée par l'insuffisance du volume des variables qui ne suffit pas à expliquer à elles seules la totalité de la consommation pour ce type de cargo. Cela souligne l'intérêt d'enrichir le modèle avec d'autres variables pertinentes.

4.4 Résultat de l'algorithme d'optimisation

Après avoir validé la première étape qui consistait à prédire la consommation de l'énergie, nous intégrons l'algorithme d'optimisation pour identifier les configurations de

fonctionnement qui minimise la consommation énergétique. Deux approches sont mises en œuvre et les résultats sont présentés dans les sections suivantes.

4.4.1 Recherche exhaustive

Elle consiste à identifier pour chaque type de cargaison des différents navires les meilleures combinaisons possibles entre les variables opérationnelles « vitesse, débit » qui conduisent à une consommation d'énergie minimale du système. Les résultats de cette recherche sont présents dans les Tableaux 4.5 à 4.7 respectivement pour les navires : 1, 2 et 3.

Tableau 4.5 Résultat de la recherche exhaustive pour le navire 1

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)
SALT	2000.0	2071.35	188.43	4.83
AGGREGATES	1400.0	908.25	83.15	11.01
PETCOKE	100.0	475.26	47.40	21.04
ORE	2000.0	333.33	98.21	30.00
GYPSUM	1900.0	3563.06	28.39	2.81
GRAIN	1400.0	576.04	86.38	17.36

Tableau 4.6 Résultat de la recherche exhaustive pour le navire 2

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)
SALT	1000.0	950.10	396.81	10.53
AGGREGATES	1400.0	333.33	574.10	30.00
ORE	100.0	1251.50	372.66	7.99
GRAIN	900.0	647.03	275.27	15.46
SUGAR	100.0	333.33	220.78	30.00
PETCOKE	400.0	333.33	220.71	30.00
GYPSUM	2000.0	1969.45	750.18	5.08

Tableau 4.7 Résultat de la recherche exhaustive pour le navire 3

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)
SALT	100.00	1581.09	455.48	6.32
PETCOKE	100.00	2830.22	377.92	3.53
AGGREGATES	2000.00	333.33	491.95	30.00
GYPSUM	2000.00	927.83	422.95	10.78
GRAIN	1577.78	333.33	296.93	30.00
SUGAR	100.00	1181.11	87.42	8.47
ORE	100.00	333.33	425.05	30.00

De manière générale, cette approche permet d'établir une référence de performance énergétique théorique. L'exploitation des résultats de ces tableaux permet d'établir que, contrairement à d'autres navires, les plages de consommation minimales du Navire 1 sont très faibles. Cette observation est aussi faite sur son diagramme Boxplot, qui représentait la consommation d'énergie en fonction des différents cargos (confére annexe II) où la médiane de presque toutes les cargaisons se trouve inférieure à 50 KW.

On observe également qu'il n'y a aucune corrélation clairement définie entre le temps de déchargement de chaque matériau et des différents navires. Chaque navire par sa conception ou en fonction des propriétés physiques des matériaux un temps de déchargement peut différer de l'un à l'autre. Cependant cette méthode présente un temps très élevé de calcul, surtout pour ces jeux de données très volumineux.

4.4.2 Optimisation avec L-BFGS-B

Cette deuxième approche d'optimisation de la consommation de l'énergie, qui intègre des contraintes non linéaires ainsi que des bornes sur les variables explicatives (comme précisé dans la phase de conception), vise à déterminer les valeurs optimales minimisant la consommation d'énergie. Une contrainte de durée maximale de 3 heures a été imposée dans l'algorithme afin de refléter les contraintes temporelles réalistes liées à l'exploitation portuaire.

Les résultats obtenus pour les différents navires sont synthétisés dans les Tableaux 4.8, 4.9 et 4.10.

Tableau 4.8 Optimisation du navire 2 avec L-BFGS-B

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)	Valide ($\leq 3h$)
SALT	136.50	3664.43	402.43	2.73	✓
AGGREGATES	1784.72	4660.64	648.04	2.15	✓
ORE	166.77	4482.26	418.12	2.23	✓
GRAIN	1013.46	3761.70	454.98	2.66	✓
SUGAR	1160.51	3002.97	629.42	3.33	✗
PETCOKE	1234.48	3487.95	328.13	2.87	✓
GYPSUM	296.69	4140.07	812.55	2.42	✓

Tableau 4.9 Optimisation du navire 3 avec L-BFGS-B

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)	Valide ($\leq 3h$)
SALT	901.12	3307.46	941.59	3.02	✗
AGGREGATES	1199.23	3047.81	523.42	3.28	✗
ORE	900.52	1771.67	427.07	5.64	✗
GRAIN	900.14	1771.67	313.17	5.64	✗
SUGAR	0.00	1771.67	87.70	5.64	✗
PETCOKE	1506.81	2372.66	407.37	4.21	✗
GYPSUM	900.91	3175.24	864.17	3.15	✗

Tableau 4.10 Optimisation du navire 1 avec L-BFGS-B

Cargo	Vitesse optimale (tr/min)	Débit optimal (t/h)	Consommation minimale (kW)	Temps estimé (h)	Valide ($\leq 3\text{h}$)
SALT	1050.0	3802.22	171.36	2.63	✓
AGGREGATES	1050.0	4264.19	91.26	2.35	✓
PETCOKE	1050.0	3030.04	42.82	3.30	✗
ORE	1050.0	2961.88	89.53	3.38	✗
GYPSUM	1050.0	4367.54	40.91	2.29	✓
GRAIN	1050.0	2639.07	90.48	3.79	✗

En analysant les tendances générales des résultats obtenus pour chaque navire, on observe que la méthode L-BFGS-B tend à privilégier des débits de déchargement très élevés. Cette stratégie permet un déchargement rapide — souvent inférieur ou égal à 3 heures — notamment dans le cas du navire 2, pour lequel la majorité des types de cargo respectent la contrainte temporelle fixée.

Cependant, cette accélération du processus peut entraîner, pour certaines cargaisons, une augmentation significative de la consommation énergétique. Néanmoins, dans d'autres cas, la consommation demeure modérée, voire inférieure à celle observée avec une recherche exhaustive, tout en réduisant considérablement le temps de déchargement. Ces résultats suggèrent une optimisation énergétique globale satisfaisante, témoignant d'un bon compromis entre rapidité d'exécution et efficacité énergétique.

Dans certains cas particuliers, cette méthode présente des performances efficaces comme PETCOKE sur le navire 1 où **3030 t/h** sont déchargés pour seulement **42.82 kW**, ce qui signifie un très bon rendement malgré un temps légèrement supérieur à 3h. Concernant les navires 1 et 3, aucune configuration optimale ne permet de respecter le seuil de 3 heures malgré l'augmentation considérable de la quantité de matière à décharger. Cela suggère que ces

navires présentent des limitations structurelles ou mécaniques ne permettant pas de combiner haut débit et faible consommation dans un temps restreint.

4.4.3 Analyse comparative des 2 méthodes

Bien que les 2 méthodes nous aient permis d'avoir des combinaisons optimales pour les variables explicatives nous permettant d'avoir des consommations réduites, chaque méthode présente des performances différentes en fonction des objectifs d'optimisation. Nous allons les évaluer selon les différentes situations.

- **Cas de la consommation de l'énergie**

Les graphiques suivants illustrent de manière comparative la consommation énergétique (en kWh) pour six types de cargaisons (SALT, AGGREGATES, PETCOKE, ORE, GYPSUM, GRAIN), en distinguant les résultats obtenus par la méthode de recherche exhaustive (courbe bleue) et l'optimisation par L-BFGS-B (courbe orange) chaque navire. La Figure 4.16 représente la comparaison de la consommation de l'énergie du navire 1.

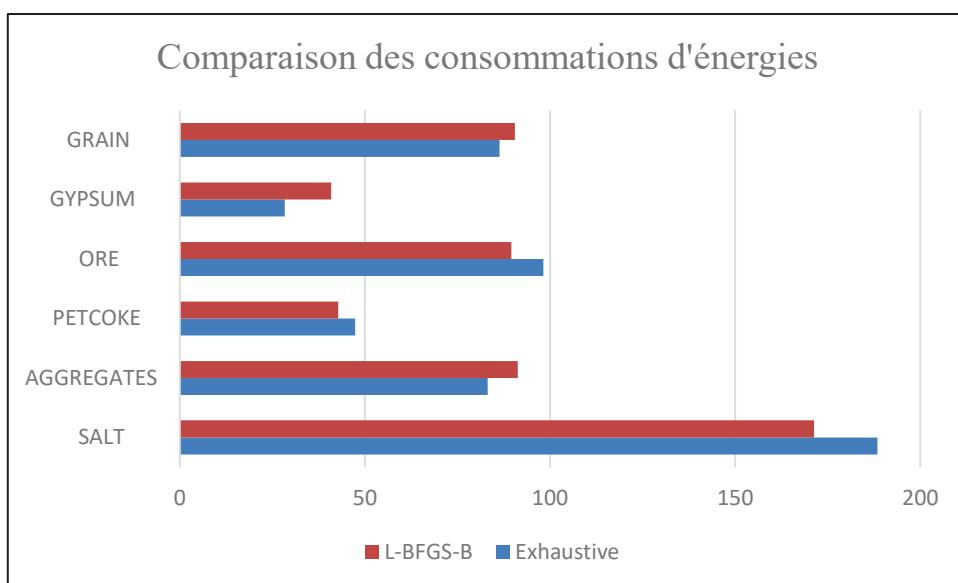


Figure 4.16 Consommation optimale du navire 1

De manière générale, on observe que les deux courbes suivent une tendance similaire, ce qui indique que les deux méthodes réagissent de manière cohérente aux propriétés des matériaux déchargés. La méthode L-BFGS-B parvient globalement à réduire légèrement la consommation énergétique, notamment pour les cargaisons SALT, ORE et surtout GYPSUM. Pour les autres cargaisons, les performances énergétiques des deux méthodes sont quasi équivalentes. Cela tend à confirmer que L-BFGS-B est capable d'atteindre des configurations proches ou meilleures que celles obtenues par recherche exhaustive, mais avec un temps de calcul bien inférieur, ce qui constitue un avantage opérationnel considérable.

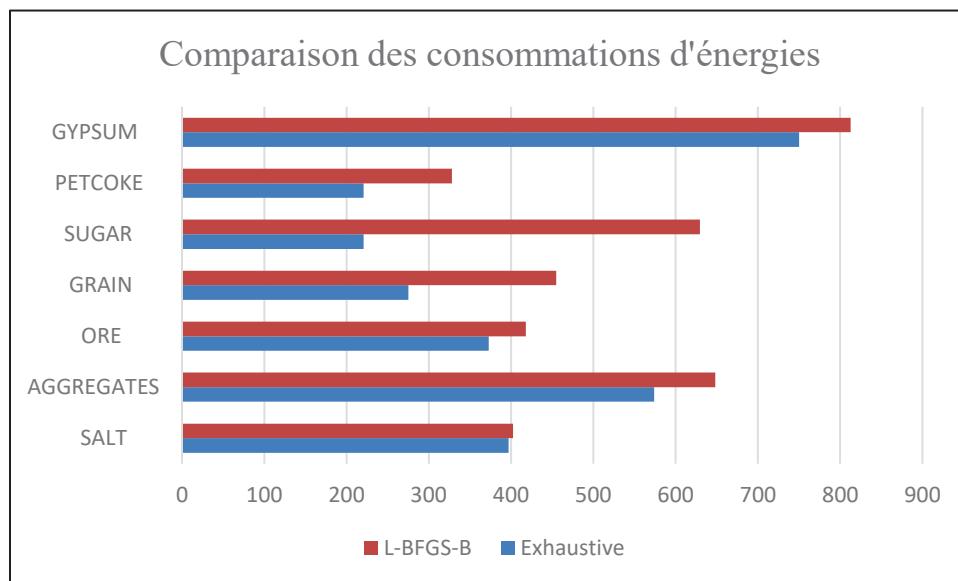


Figure 4.17 Consommation optimale du navire 2

Sur le navire 2, d'après la Figure 4.17 ci-dessus, la méthode de recherche exhaustive démontre une meilleure maîtrise énergétique globale. Si la méthode L-BFGS-B permet sans doute de réduire les temps d'opération, elle entraîne ici une hausse systématique de la consommation d'énergie, parfois marquée, notamment ceux des cargos AGGREGATES et SUGAR où nous observons un pic de consommation dépassant nettement celle de la méthode exhaustive, avec une hausse supérieure à 100 kWh dans chaque cas qui pourrait refléter une sollicitation excessive des moteurs pour maintenir un haut débit. Il serait donc intéressant, pour ce navire,

d'intégrer un plafond énergétique nécessaire pour tirer pleinement profit des capacités d'optimisation rapide de L-BFGS-B sans détériorer la performance énergétique.

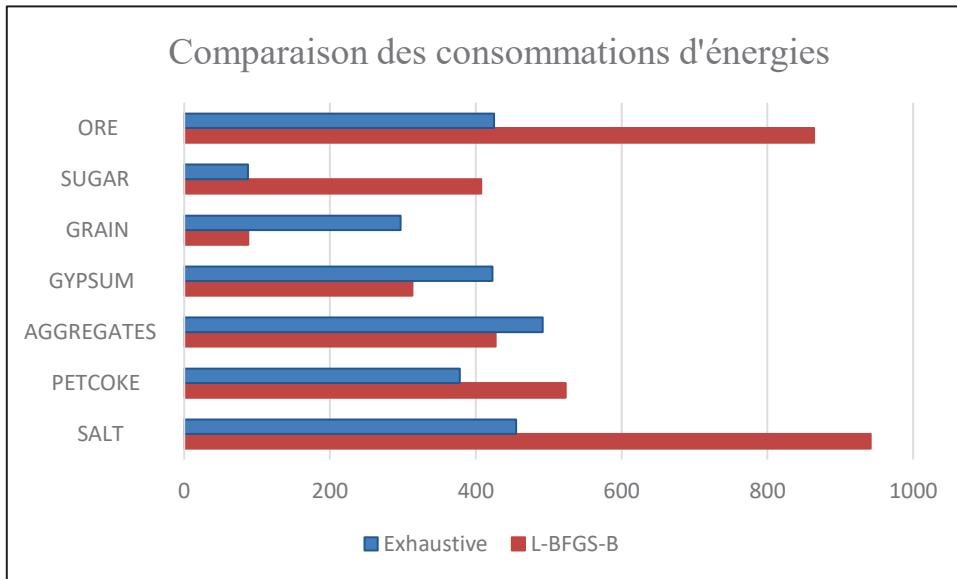


Figure 4.18 Consommation optimale du navire 3

Dans le cas du navire 3, la méthode L-BFGS-B se révèle moins efficace, entraînant une consommation d'énergie nettement plus élevée que celle obtenue par la recherche exhaustive pour la majorité des cargaisons. Les écarts sont particulièrement marqués pour les cargaisons de type SALT et ORE. Toutefois, cette méthode permet une consommation légèrement inférieure pour certains types de cargos, tels que AGGREGATES, GYPSUM et GRAIN. Ces résultats indiquent que, pour ce navire, l'optimisation de la vitesse par L-BFGS-B se fait au détriment de l'efficacité énergétique, engendrant une surconsommation notable.

- Cas de la productivité opérationnelle**

Les graphiques suivants illustrent de manière comparative la productivité (en t/h) pour six types de cargaisons (SALT, AGGREGATES, PETCOKE, ORE, GYPSUM, GRAIN), en distinguant les résultats obtenus par la méthode de recherche exhaustive (courbe bleue) et l'optimisation par L-BFGS-B (courbe orange) de chaque navire.

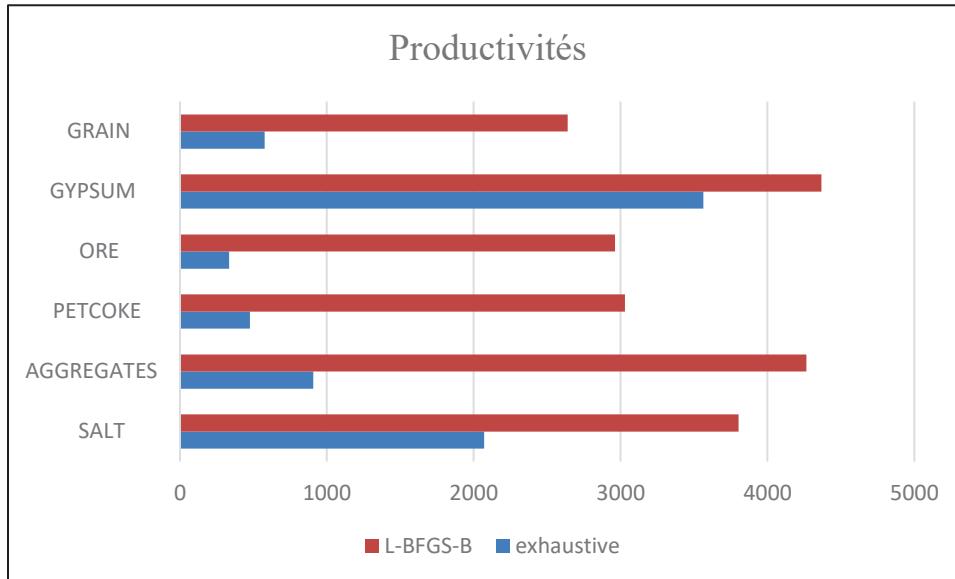


Figure 4.19 Productivités optimales du navire 1

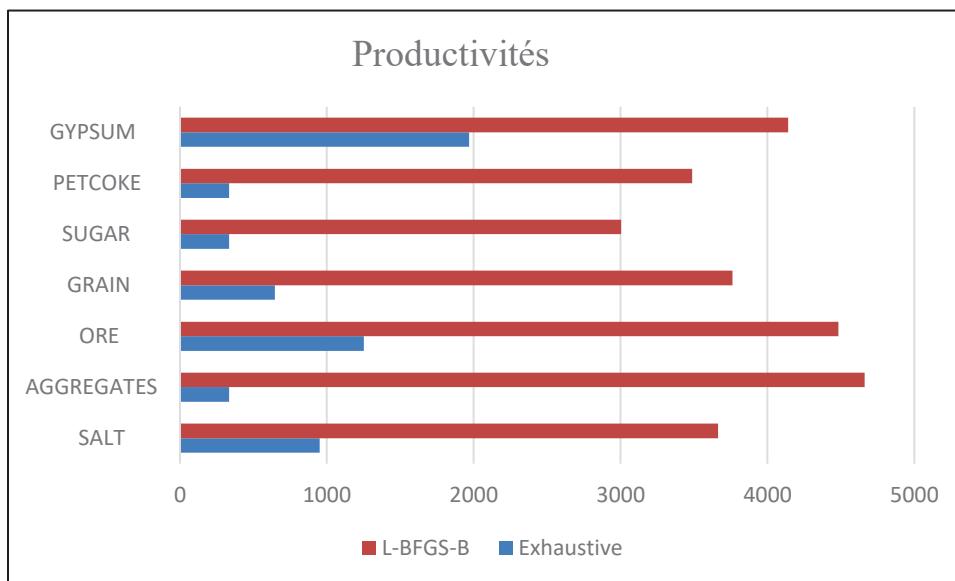


Figure 4.20 Productivités optimales du navire 2

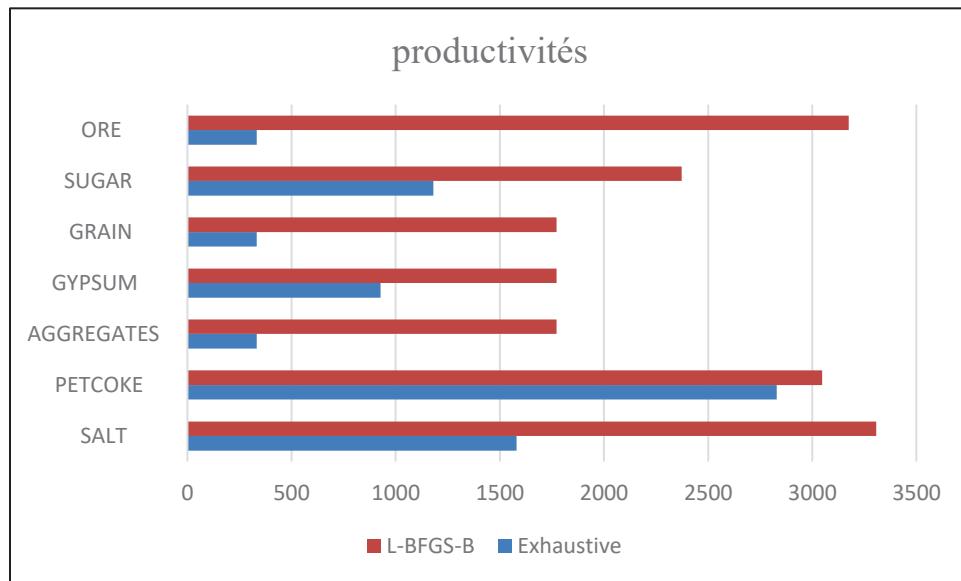


Figure 4.21 Productivités optimales du navire 3

L’analyse des Figures 4.19, 4.20 et 4.21, qui illustrent les productivités optimales atteintes respectivement pour les navires 1, 2 et 3, met en évidence une tendance claire et significative : la méthode d’optimisation L-BFGS-B surpassé systématiquement la recherche exhaustive en termes de débit de déchargement et de temps d’exécution, indépendamment du type de cargaison. De manière générale, les débits obtenus avec L-BFGS-B varient entre 3 000 et 4 700 tonnes par heure (t/h), tandis que ceux de la méthode exhaustive restent souvent limités à 1 000 – 1 500 t/h. Dans certains cas extrêmes, la méthode exhaustive nécessite jusqu’à 30 heures pour décharger une petite quantité de matière à un débit de seulement 333,33 t/h. Cette supériorité s’explique par la capacité de l’algorithme L-BFGS-B à explorer plus efficacement l’espace des paramètres, en exploitant les gradients de la fonction objectif. Cela lui permet d’identifier rapidement des configurations mécaniques qui optimisent le fonctionnement des convoyeurs, atteignant ainsi des régimes de performance bien plus élevés.

Le navire 2 se distingue par sa forte réactivité à l’optimisation via la méthode L-BFGS-B, affichant une productivité élevée et stable pour l’ensemble des cargaisons. En revanche, les navires 1 et 3 présentent des profils plus hétérogènes : bien que l’augmentation de productivité soit systématiquement observée sous L-BFGS-B, elle demeure parfois plus modérée,

notamment pour des cargaisons comme le PETCOKE ou le SALT. Ces variations peuvent s'expliquer par des contraintes mécaniques spécifiques à chaque navire, ou par des limites physiques liées aux caractéristiques des matériaux traités.

4.5 Outil de recommandation

Afin de faciliter l'exploitation des résultats obtenus dans le cadre de ce mémoire, un tableau de bord décisionnel Power Bi a été conçu. Intitulé OPTI-CONVOY, cet outil visuel permet aux opérateurs portuaires et aux gestionnaires de flotte de comparer rapidement les performances opérationnelles des navires selon les différentes stratégies d'optimisation testées.

Le tableau de bord se concentre sur deux variables stratégiques : la vitesse de convoyage (en m/s) et le taux de décharge (en tonnes/heure), qui sont au centre des décisions opérationnelles relatives à l'efficacité énergétique. À travers des graphiques comparatifs pour chacun des trois navires étudiés (Navire 1, Navire 2, Navire 3), l'utilisateur peut visualiser l'impact de la méthode exhaustive par rapport à l'optimisation L-BFGS-B, pour chaque type de cargaison analysée (sel, mineraï, céréales). Les fonctionnalités clés qu'on pourra observer sur tableau de bord sont :

Graphiques comparatifs : qui permet d'afficher clairement les valeurs (vitesse et débit) optimales obtenues par chaque méthode, pour chaque cargaison.

Filtres interactifs : ici l'opérateur pourra sélectionner le type de navire ou de cargaison et la méthode d'optimisation pour un affichage personnalisé.

Recommandation automatique : une synthèse visuelle suggère la méthode optimale selon un objectif donné (temps, énergie ou compromis). Une légende intuitive présentée sur la figure 4.22 ci-dessous, permet de distinguer en un coup d'œil les gains de productivité ou les hausses de consommation.

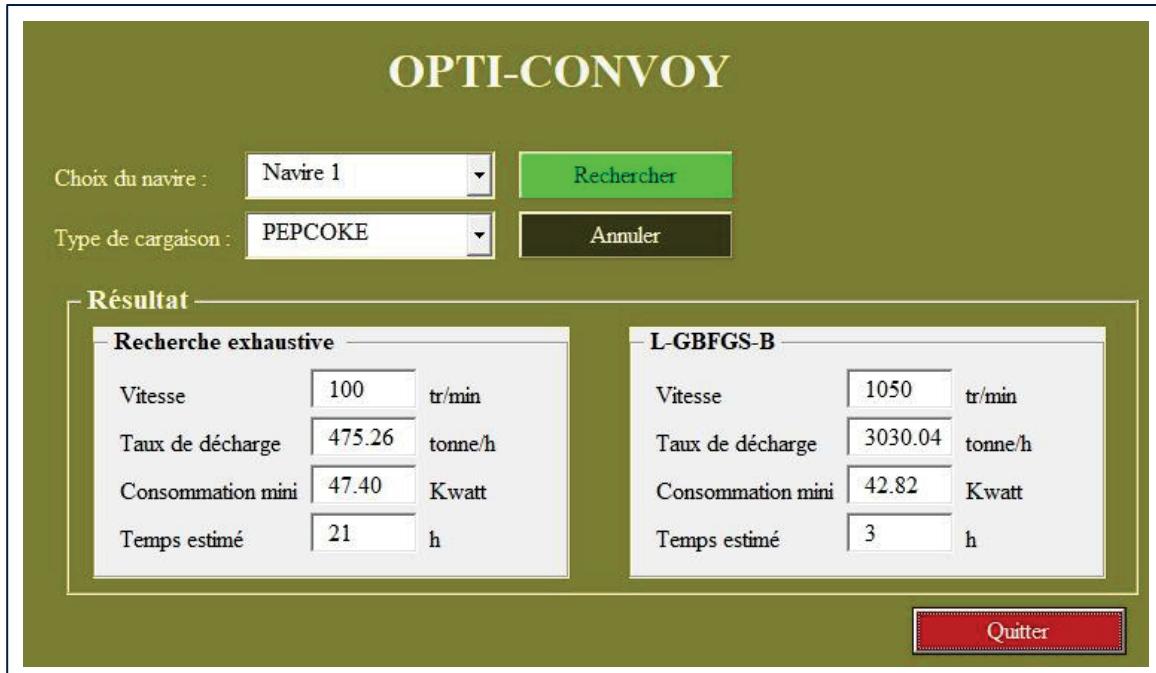


Figure 4.22 Tableau de bord itératif

4.6 Discussion

Les résultats présentés dans les sections précédentes permettent de mettre en évidence plusieurs tendances majeures concernant la prédiction de la consommation énergétique et l'optimisation du fonctionnement des convoyeurs à bande sur les navires auto-déchargeurs. Cette discussion propose une mise en perspective critique de ces résultats, en évaluant la pertinence des méthodes appliquées et en identifiant les implications opérationnelles qui en découlent.

Tout d'abord, les performances du modèle de prédiction de l'algorithme forêt aléatoire s'avèrent très satisfaisantes. Que ce soit pour l'analyse globale ou par type de cargo, les coefficients de détermination R^2 obtenus sont supérieurs à **0.97** dans la majorité des cas, et les erreurs quadratiques moyennes RMSE restent faibles. Ces résultats confirment la capacité du modèle à capturer les relations non linéaires complexes entre les variables opérationnelles (débit, vitesse, puissance) et la consommation d'énergie. Toutefois, certains cas limites,

comme celui du cargo ORE sur le navire 3 illustrent les limites du modèle lorsque la taille de l'échantillon est trop restreinte ou lorsque les variables explicatives ne suffisent pas à représenter toute la variabilité du système. Cela met en évidence la nécessité d'enrichir les jeux de données et d'intégrer des variables contextuelles tels que, humidité ou toute autre information pouvant enrichir le modèle.

Sur le plan de l'optimisation, les deux méthodes ont été testées, la recherche exhaustive et L-BFGS-B, ont permis d'identifier des configurations efficaces entre le taux de charge et la vitesse du convoyeur, mais avec des résultats contrastés selon les objectifs fixés. D'une part, la recherche exhaustive fournit des configurations énergétiquement économiques, au prix de temps de déchargement souvent très élevés, allant jusqu'à 30 heures pour certains cargos. D'autre part, l'algorithme L-BFGS-B permet d'atteindre des débits extrêmement élevés, réduisant systématiquement la durée des opérations à moins de 3 heures pour la plupart des cargos à bord du navire 2. Ce gain de temps est toutefois parfois contrebalancé par une consommation d'énergie plus importante, notamment sur les navires 2 et 3 pour certains cargos.

Les analyses comparatives ont révélé que le navire 2 présente une très bonne compatibilité avec l'optimisation L-BFGS-B, avec des performances élevées tant en productivité qu'en consommation. Le navire 3, à l'inverse, réagit mal à l'accélération du débit imposé par l'algorithme, engendrant une surconsommation énergétique marquée. Le navire 1, offre un profil intermédiaire, avec des gains de productivité nets sous L-BFGS-B et une consommation souvent modérée.

L'analyse croisant les critères de consommation et de productivité montre que l'optimisation ne peut être abordée sous un angle unique. Si l'objectif est la réduction du temps de déchargement, L-BFGS-B est la méthode à privilégier. En revanche, pour minimiser la consommation d'énergie ou pour répondre à des objectifs environnementaux ou économiques, la recherche exhaustive conserve tout son intérêt. Une voie d'amélioration future consisterait

à développer une méthode hybride, capable d'intégrer un critère multicritère (temps + énergie) et de pondérer dynamiquement les priorités selon les contextes logistiques.

Malgré l'homogénéité apparente des matériaux transportés (par exemple, le GYPSUM), on observe des différences marquées de consommation énergétique selon les navires. Le même type de cargaison peut générer des niveaux de consommation très différents selon le navire. Ces écarts pourraient s'expliquer par la variabilité structurelle des systèmes de convoyeurs entre navires (nombre d'étages, type de moteurs, inclinaison), mais aussi par les conditions opérationnelles spécifiques (ports, conditions d'humidité, distance de déchargement). De plus, chaque modèle de prédiction est ajusté sur un historique propre à un navire donné, ce qui reflète les comportements énergétiques appris localement plutôt qu'un standard internavire. Ces différences soulignent l'importance de calibrer les paramètres d'optimisation et les modèles en fonction du profil technique de chaque navire, et non uniquement du type de cargaison.

Enfin, cette étude confirme que les systèmes de convoyeurs peuvent bénéficier pleinement des outils d'apprentissage automatique et d'optimisation pour améliorer leur efficience opérationnelle. L'intégration de ces approches dans les chaînes de décision portuaires pourrait permettre de réduire les coûts d'exploitation, d'améliorer la planification des opérations et de participer aux objectifs globaux de transition énergétique du transport maritime.

4.7 Conclusion

Ce chapitre a permis de valider, par une analyse approfondie, la pertinence de l'approche méthodologique adoptée dans ce mémoire. En combinant l'entraînement de modèles de prédiction avec la mise en œuvre de deux stratégies d'optimisation distinctes, des tendances claires ont pu être dégagées concernant les performances énergétiques et opérationnelles des convoyeurs à bande.

L'approche fondée sur l'algorithme de la forêt aléatoire s'est révélée particulièrement efficace pour prédire la consommation d'énergie avec un haut degré de précision. Parallèlement,

l'optimisation à l'aide de la méthode L-BFGS-B a permis, dans la majorité des cas, d'accélérer de manière significative le processus de déchargement, parfois au prix d'une légère augmentation de la consommation énergétique, mais offrant globalement un bon compromis entre rapidité et efficience. Ainsi, l'intégration d'outils avancés d'intelligence artificielle dans une stratégie globale de transition numérique et écologique apparaît comme un levier prometteur pour améliorer à la fois la performance et la durabilité des opérations de transport maritime.

CONCLUSION ET PERSPECTIVE

Conclusion générale

Ce mémoire a exploré une méthode d'optimisation des convoyeurs à bande embarqués sur des navires auto-déchargeurs à travers une approche combinant les méthodes d'apprentissage automatique et les algorithmes d'optimisation. L'étude s'est limitée à 3 navires dont les données ont été collectées pendant toutes les phases de fonctionnements. L'objectif étant d'optimiser la consommation de l'énergie des convoyeurs à bandes en proposant des recommandations aux opérateurs pour leur permettre d'effectuer des opérations de déchargements efficaces. L'ensemble de ce travail c'est articulé sur 4 chapitres chacun apportant des éclairages précis et complémentaire sur les différentes possibilités d'amélioration des performances opérationnelles dans un contexte portuaire.

Dans un premier temps, une analyse exploratoire des données issues de capteurs réels a permis d'identifier les variables influentes sur la consommation énergétique, telles que la vitesse de convoyage, le débit de décharge, ou encore le type de cargaison transportée. Cette étape a mis en évidence la variabilité importante des comportements selon les navires, soulignant l'importance d'une modélisation selon le type de cargaison.

Par la suite un modèle de prédiction basé sur l'algorithme de forêt aléatoire a été entraîné afin d'estimer la consommation d'électricité à partir des paramètres opérationnels. Ce modèle a démontré une excellente capacité prédictive avec des coefficients de détermination R^2 supérieurs à 0,97 dans la plupart des navires que ce soit sur les données d'entraînement ou les données test traduisant une forte adéquation entre les prédictions et les valeurs réelles. Toutefois, certains cas marginaux ont mis en lumière les limites du modèle lorsqu'il est appliqué à des cargos sous-représentés ou à des configurations atypiques.

La troisième étape du mémoire a consisté à formuler le problème d'optimisation de manière mathématique, en cherchant à minimiser la consommation d'énergie $E(X)$ sous des contraintes现实的 sur les variables de décision (vitesse et débit). Deux méthodes ont été comparées : la

recherche exhaustive et l'algorithme L-BFGS-B. Cette comparaison a révélé des profils d'optimisation contrastés. La recherche exhaustive, bien que précise, génère des configurations très économies, mais peu productives. L-BFGS-B, au contraire, permet d'obtenir des gains de productivité significatifs, parfois au détriment de la réduction énergétique, selon le navire et le type de cargaison. Ces résultats confirment la nécessité d'un arbitrage dynamique de l'opérateur entre vitesse d'exécution, consommation énergétique et contraintes techniques.

Enfin, une évaluation comparative par navire a mis en évidence la sensibilité des résultats aux spécificités techniques et aux comportements propres à chaque unité maritime. Le navire 2 s'est avéré particulièrement adapté à une optimisation agressive via L-BFGS-B, tandis que le navire 3 a montré une surconsommation notable sous contrainte d'accélération du débit. Le navire 1 quant à lui se situe dans un équilibre intermédiaire.

Néanmoins certaines limites sont observées dans ce travail notamment le déséquilibre partiel de l'ensemble du jeu de donnée, car certains cargos ne sont pas vraiment représentés. La distribution de la consommation de l'énergie n'est pas uniforme selon les différents navires bien qu'ils effectuent les opérations dans des contexte similaire. Nous observons également l'inadéquation de l'algorithme L-BFGS-B dans certains cas où le gradient est difficilement interprétable en raison du caractère non différentiable des forêts aléatoires.

Perspectives

Ces recommandations ont pour but d'améliorer les résultats, de renforcer la robustesse, la précision et la portée des analyses futures.

Enrichissement du jeu de données avec intégration de nouvelles variables explicatives

En raison de la diversité des états opérationnels enregistrés (navigation, chargement, pause, déchargement). Il est recommandé de mettre en œuvre une stratégie de collecte plus ciblée, avec des intervalles de temps plus courts, concentrée sur les phases actives de déchargement. Cela permettrait de mieux capter les fluctuations rapides des variables critiques, facteurs majeurs influençant la consommation énergétique.

De plus le modèle pourrait être améliorée de manière significative en intégrant les variables contextuelles supplémentaire comme le poids total des cargaisons, la densité des matériaux manipulés, les conditions météorologiques et même les caractéristiques techniques de chaque navire.

Exploitation des données pour la maintenance prédictive

Les données opérationnelles collectées dans ce projet peuvent également être mobilisées pour détecter des signes précoce d'usure ou de dégradation des équipements. Une dérive inhabituelle de la consommation d'énergie à débit constant, une chute de performance à vitesse stable, ou des pics anormaux de puissance peuvent signaler soit usure des moteurs ou des courroies, une mauvaise lubrification ou simplement une défaillance mécanique. Il serait donc pertinent de développer, en parallèle de l'optimisation énergétique, un module de détection d'anomalies basé sur l'apprentissage automatique, capable d'alerter les opérateurs sur la nécessité de maintenance préventive.

ANNEXE I

MATRICE DE CORRÉLATION

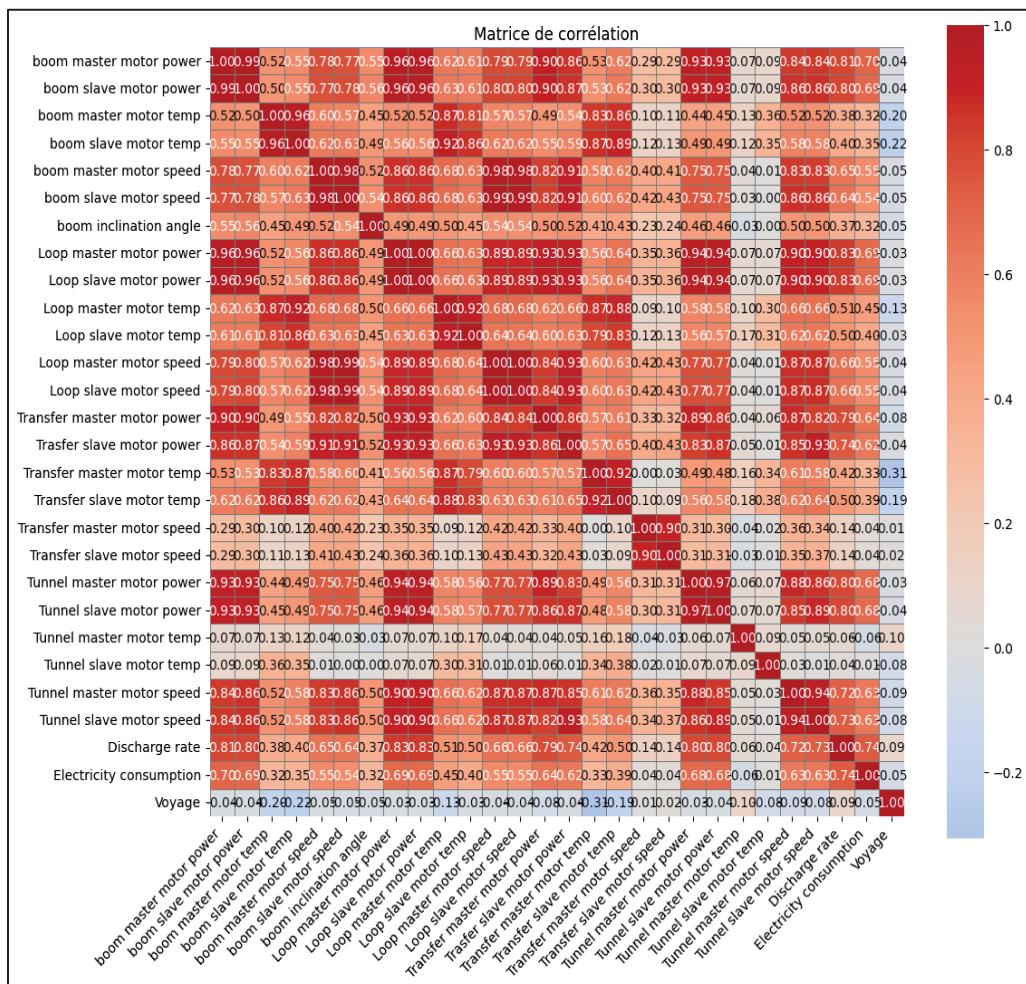


Figure-A I-1 Matrice de corrélation du navire 1

ANNEXE II

DISTRIBUTION DE LA CONSOMMATION D'ÉLECTRICITÉ

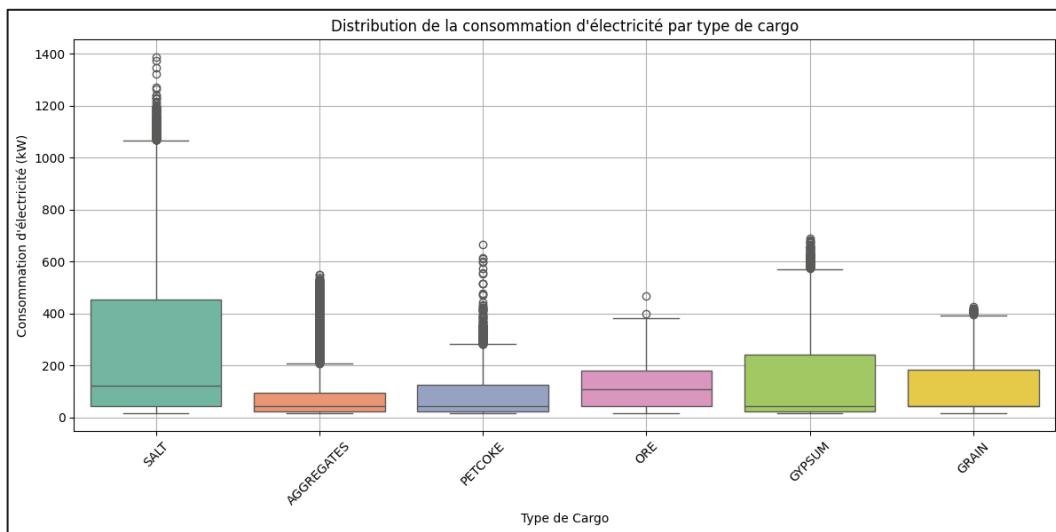


Figure-A II-1 Distribution de la consommation d'électricité (cas du navire 1)

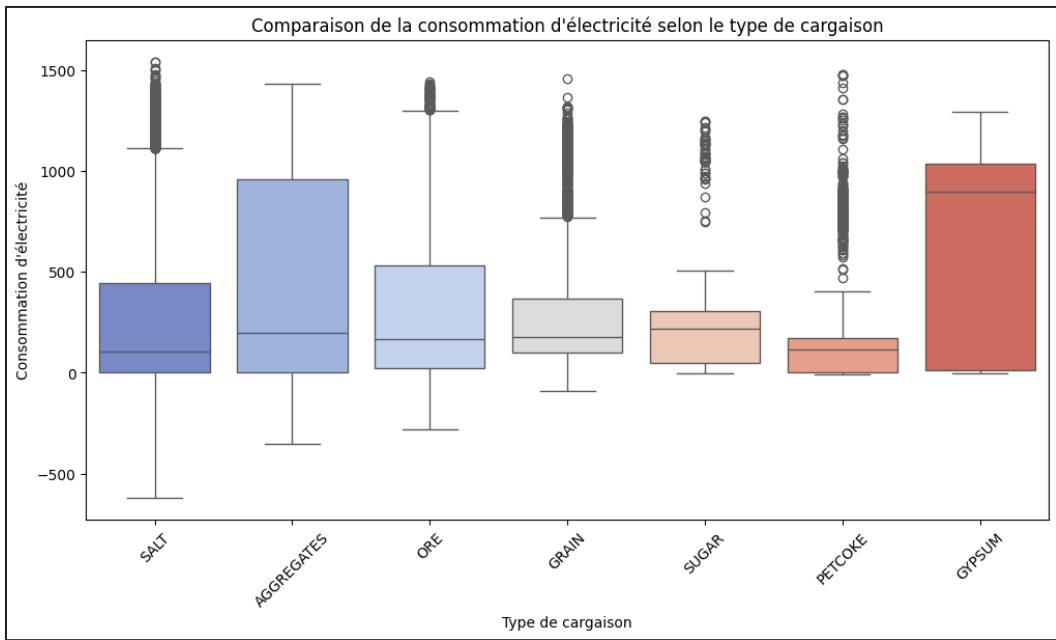


Figure-A II-2 Distribution de la consommation d'électricité (cas du navire 2)

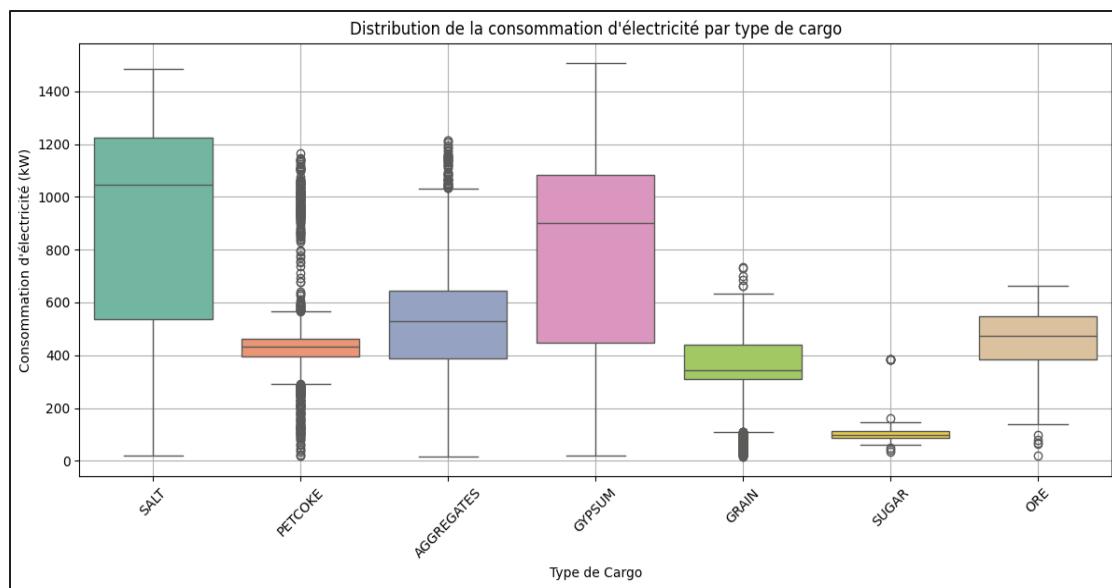


Figure-A II-3 Distribution de la consommation d'électricité (cas du navire 3)

ANNEXE III

VALIDATION CROISÉE

```
[1] from sklearn.model_selection import cross_val_score
    import numpy as np

    # Cross-validation sur 5 folds (peut être changé à 10 pour plus de robustesse)
    scores = cross_val_score(pipeline, X_train, y_train, cv=5, scoring='r2')

    # Affichage des résultats
    print("⌚ Scores R2 par fold : ", scores)
    print(f"✓ Moyenne R2      : {scores.mean():.4f}")
    print(f"☒ Écart-type R2   : {scores.std():.4f}")

→⌚ Scores R2 par fold : [0.99999834 0.99996698 0.99999584 0.99999585 0.99996318]
  ✓ Moyenne R2      : 1.0000
  ☒ Écart-type R2   : 0.0000
```

Figure-A III-1 Extrait de code

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Agamy, K. (2023). Ships and Offshore Structures. *18*(8), 1104-1115. Repéré à <https://etsmtl.on.worldcat.org:443/atoztitles/link/?sid=Elsevier:Scopus>
- \Alamu, O., Olwal, T. O., & Migabo, E. M. (2025). Machine Learning Applications in Energy Harvesting Internet of Things Networks: A Review. *IEEE Access*, *13*, 4235-4266. doi: 10.1109/ACCESS.2024.3525263. Repéré à <http://dx.doi.org/10.1109/ACCESS.2024.3525263>
- Chen, H., Zeng, F., Du, J., & Zhou, S. (2015). An improved optimal model for energy-saving of belt conveyors based on genetic algorithm. Dans *2015 IEEE International Conference on Information and Automation* (pp. 2784-2788). doi: 10.1109/ICInfa.2015.7279759
- Chiandussi, G., Codegone, M., Ferrero, S., & Varesio, F. E. (2012). Comparison of multi objective optimization methodologies for engineering applications. *Computers & Mathematics with Applications*, *63*(5), 912-942. doi: <https://doi.org/10.1016/j.camwa.2011.11.057>. Repéré à <https://www.sciencedirect.com/science/article/pii/S0898122111010406>
- Daras, G. (2017). Conception et réalisation d'un outil de traitement et analyse des données spatiales pour l'aide à la décision : application au secteur de la distribution (Ph.D., Ecole Polytechnique, Montreal (Canada), Canada -- Quebec, CA). Repéré à ProQuest Dissertations & Theses Global; ProQuest Dissertations & Theses Global Closed Collection. (27537015). Repéré à <https://www.proquest.com/docview/2291445971?accountid=27231&sourcetype=Dissertations%20&%20Theses>
- <https://etsmtl.on.worldcat.org/atoztitles/link?sid=ProQ:&issn=&volume=&issue=&title=Conception+et+r%C3%A9alisation+d%E2%80%99un+outil+de+traitement+et+analyse+des+donn%C3%A9es+spatiales+pour+l%E2%80%99aide+%C3%A0+la+d%C3%A9cision+%3A+application+au+secteur+de+la+distribution&spage=&date=2017-01-01&atitle=Conception+et+r%C3%A9alisation+d%E2%80%99un+outil+de+traitement+et+analyse+des+donn%C3%A9es+spatiales+pour+l%E2%80%99aide+%C3%A0+la+d%C3%A9cision+%3A+application+au+secteur+de+la+distribution&au=Daras%2C+Gautier&id=doi:>
- Feriani, A., & Hossain, E. (2021). Single and Multi-Agent Deep Reinforcement Learning for AI-Enabled Wireless Networks: A Tutorial. *IEEE Communications Surveys & Tutorials*, *23*(2), 1226-1252. doi: 10.1109/COMST.2021.3063822
- Group, T. F. (2025). Identification des dangers pour les vraquiers autodéchargeurs. Repéré à <https://www.tandfonline.com/doi/full/10.1080/17445302.2022.2107307#d1e143>
- Guide, B. C. (2010). Repéré à <https://bulkcarrierguide.com/self-unloaders.html>

HADDOUCHE, M. (2023). *L'efficacité énergétique et l'industrie 4.0 dans l'industrie du bois* (UNIVERSITÉ DU QUÉBEC À RIMOUSKI).

Hermann, M., Bücker, I., & Otto, B. (2020). Industrie 4.0 process transformation: findings from a case study in automotive logistics. *Journal of Manufacturing Technology Management*, 31(5), 935-953. doi: 10.1108/JMTM-08-2018-0274. Repéré à <https://doi.org/10.1108/JMTM-08-2018-0274>

Hiltermann, J., G., L., L., S. D., C., R. J., M., D. J. A. J., & and Pang, Y. (2011). A Methodology to Predict Power Savings of Troughed Belt Conveyors by Speed Control. *Particulate Science and Technology*, 29(1), 14-27. doi: 10.1080/02726351.2010.491105. Repéré à <https://doi.org/10.1080/02726351.2010.491105>

Huang, J., Li, C., Long, H., Wang, J., Wen, B., & Yin, L. (2023). Multi-objective optimization design of a heavy-duty folding mechanism and Self-discharging equipment development. *Mechanics & Industry*, 24, 19. doi: 10.1051/meca/2023015. Repéré à https://www.mechanicsindustry.org/articles/meca/full_html/2023/01/mi220032/mi220032.html

Ilyushin, B. B. (2024). On Applicability of IQR Method for Filtering of Experimental Data. *Journal of Engineering Thermophysics*, 33(1), 1-8. doi: 10.1134/S1810232824010016. Repéré à <https://doi.org/10.1134/S1810232824010016>

Klacar, N. (2023). L'IQR dans les tests d'automatisation : libérer le potentiel de l'analyse des données. Repéré à <https://www.atlantbh.com/iqr-in-automation-testing-unleashing-the-data-analytics-potential/>

M, A., Totla, T., Renwa, Y., C, R., Subbiah, S., Dharmaraj, T. B., & Prakash, S. O. (2022). Comparative Analysis of Ant Colony Optimization and Particle Swarm Optimization for Test Case Prioritization. Dans 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 680-686). doi: 10.1109/3ICT56508.2022.9990713

Machała, S., Chamier-Gliszczyński, N., & Królikowski, T. (2022). Application of AR/VR Technology in Industry 4.0. *Procedia Computer Science*, 207, 2984-2992. doi: 10.1016/j.procs.2022.09.357

Mhlongo, I. N., Nnachi, G. U., Nnachi, A. F., & Adesola, A. T. (2020). Modelling and Simulation of Conveyor Belt for Energy Efficiency Studies. Dans 2020 IEEE PES/IAS PowerAfrica (pp. 1-5). doi: 10.1109/PowerAfrica49420.2020.9219974

Miao, P., & Fan, L. (2020). Energy efficiency optimization of belt conveyors with bias noise based on beetle antennae search algorithm. Dans 2020 39th Chinese Control Conference (CCC) (pp. 1569-1574). doi: 10.23919/CCC50068.2020.9189153

- Pang, Y., Qiao, T., Yan, G., & Lodewijks, G. (2023). Improving Energy Efficiency of Belt Conveying System by Controlling Transient Operation Sequence. Dans ICBMH 2023 - 14th International Conference on Bulk Materials Storage, Handling and Transportation, Proceedings (pp. 256-262). Repéré à <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85219622022&partnerID=40&md5=d92b59ecfae7188dabe67ec16c47a0e8>
- Schmidl, E., Fischer, E., Steindl, J., Wenk, M., & Franke, J. (2021). Reinforcement learning for energy reduction of conveying and handling systems. *Procedia CIRP*, 97, 290-295. doi: <https://doi.org/10.1016/j.procir.2020.05.240>. Repéré à <https://www.sciencedirect.com/science/article/pii/S221282712031461X>
- Shirong, Z., & Xiaohua, X. (2009). A new energy calculation model of belt conveyor. Dans *AFRICON 2009* (pp. 1-6). doi: 10.1109/AFRCON.2009.5308257
- Smith, M. R., & Priddis, G. (2014). Self-unloading bulk carriers: a novel system for transporting dry bulk cargo. *The Journal of Navigation*, 67(6), 1097-1112.
- Swami, R., Dave, M., & Ranga, V. (2023). IQR-based approach for DDoS detection and mitigation in SDN. *Defence Technology*, 25, 76-87. doi: <https://doi.org/10.1016/j.dt.2022.10.006>. Repéré à <https://www.sciencedirect.com/science/article/pii/S2214914722002239>
- Tran, T. T., Phan, N. Q., & Huynh, H. X. (2025). Enhancing CO₂ Sequestration Modeling in Mangrove Forests Through Random Forest Hyperparameter Optimization and Remote Sensing Band Analysis. Dans 2025 5th Asia Conference on Information Engineering (ACIE) (pp. 37-43). doi: 10.1109/ACIE64499.2025.00013
- Tran, T. T., Phan, N. Q., & Huynh, H. X. (2025). Random Forest Model Parameters Optimization. Dans N. Thai-Nghe, T.-N. Do & S. Benferhat (Éds.), *Intelligent Systems and Data Science* (pp. 237-247). Springer Nature Singapore.
- VALPROD. (2025). Repéré à <https://valprod.fr/convoyeur-a-bande/>
- Yue, M., Yan, H., Han, R., & Wu, Z. (2025). A DDoS attack detection method based on IQR and DFFCNN in SDN. *Journal of Network and Computer Applications*, 240, 104203. doi: <https://doi.org/10.1016/j.jnca.2025.104203>. Repéré à <https://www.sciencedirect.com/science/article/pii/S1084804525001006>
- Zeng, F., & Wu, Q. (2015). A prediction model for the energy consumption of a belt conveyor system based on neural network. Dans *2015 IEEE International Conference on Information and Automation* (pp. 1865-1870). doi: 10.1109/ICInfA.2015.7279592

Zhang, C., Patras, P., & Haddadi, H. (2019). Deep Learning in Mobile and Wireless Networking: A Survey. *IEEE Communications Surveys & Tutorials*, 21(3), 2224-2287. doi: 10.1109/COMST.2019.2904897

Zhang, S., & Xia, X. (2011). Modeling and energy efficiency optimization of belt conveyors. *Applied Energy*, 88(9), 3061-3071. doi: <https://doi.org/10.1016/j.apenergy.2011.03.015>. Repéré à <https://www.sciencedirect.com/science/article/pii/S0306261911001723>