NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL & COMPUTER ENIGINEERING
UNIVERSITY OF PIRAEUS
SCHOOL OF MARITIME AND INDUSTRIAL STUDIES
DEPARTMENT OF INDUSTRIAL MANAGEMENT & TECHNOLOGY
INTERDISCIPLINARY POSTGRADUATE STUDIES
“ENGINEERING – ECONOMIC SYSTEMS”

MASTER’S THESIS

DEVELOPMENT OF A MACHINE LEARNING MODEL FOR SHIP SPEED PREDICTION: A DATA-DRIVEN APPROACH

SUBMITTED BY:

ZAGKLIVERI THEANO

SUPERVISOR:

DIMITRIOS ASKOUNIS

PROFESSOR, NTUA

ATHENS

MARCH 2025

~ This page was intentionally left blank. ~

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL & COMPUTER ENIGINEERING
UNIVERSITY OF PIRAEUS
SCHOOL OF MARITIME AND INDUSTRIAL STUDIES
DEPARTMENT OF INDUSTRIAL MANAGEMENT & TECHNOLOGY
INTERDISCIPLINARY POSTGRADUATE PROGRAMME
“ENGINEERING – ECONOMIC SYSTEMS”

MASTER’S THESIS

DEVELOPMENT OF A MACHINE LEARNING MODEL FOR SHIP SPEED PREDICTION: A DATA-DRIVEN APPROACH

ZAGKLIVERI THEANO

SUPERVISOR: DIMITRIOS ASKOUNIS PROFESSOR, NTUA

Approved by the examination committee on March 11th, 2025.

Askounis Dimitrios	Psarras Ioannis	Marinakis Evangelos
Professor NTUA	Professor NTUA	Assistant Professor NTUA

ATHENS

MARCH 2025

Zagkliveri Theano

Graduate of the Interdisciplinary Postgraduate Programme

“Engineering-Economic Systems”

Master of Science

School of Electrical and Computer Engineering

National Technical University of Athens

March 2025

Περίληψη

Η ακριβής πρόβλεψη της ταχύτητας του πλοίου και των απαιτήσεων ισχύος αποτελεί κρίσιμο στοιχείο στις θαλάσσιες μεταφορές, επηρεάζοντας την ενεργειακή απόδοση, τη συμμόρφωση με τους κανονισμούς και τη βελτιστοποίηση του ταξιδιού. Οι παραδοσιακές μέθοδοι εκτίμησης της σχέσης ταχύτητας-ισχύος των πλοίων βασίζονται σε ημι-εμπειρικά μοντέλα αντίστασης και παρεμβολές από δοκιμές στη θάλασσα (Sea Trials), τα οποία συχνά αποτυγχάνουν να λάβουν υπόψη τις πραγματικές συνθήκες στη θαλασσα, όπως ο άνεμος, τα κύματα και η δυναμική απόδοση των μηχανών. Για την αντιμετώπιση αυτών των περιορισμών, η παρούσα μελέτη αναπτύσσει ένα μοντέλο πρόβλεψης της ταχύτητας του πλοίου βασισμένο στη μηχανική μάθηση, συνδυάζοντας τεχνικές φυσικής μοντελοποίησης με σύγχρονες μεθόδους τεχνητής νοημοσύνης.

Η έρευνα ξεκινά με μια εκτενή διερευνητική ανάλυση δεδομένων (EDA) των λειτουργικών δεδομένων του πλοίου, συμπεριλαμβανομένων δεδομένων από αισθητήρες και ημερήσιες αναφορές (noon reports). Εφαρμόζονται τεχνικές όπως ανίχνευση ακραίων τιμών, επιλογή χαρακτηριστικών και ανάλυση συσχέτισης, με στόχο τη βελτιστοποίηση του συνόλου δεδομένων για τις ανάγκες της μοντελοποίησης. Στη συνέχεια, η μελέτη υλοποιεί ένα υβριδικό πλαίσιο πρόβλεψης, ενσωματώνοντας μοντέλα βασισμένα στη φυσική (π.χ. ημιεμπειρικά μοντέλα αντίστασης, παρεμβολές δοκιμών στη θάλασσα και Physics-Informed Neural Networks (PINNs)) καθώς και αλγόριθμους μηχανικής μάθησης (π.χ. Γραμμική Παλινδρόμηση, Random Forest Regressor και XGBoost).

Παρουσιάζεται μια παράλληλη προσέγγιση γκρι-κουτιού (grey-box modeling), η οποία ενσωματώνει φυσικούς περιορισμούς στις προβλέψεις μηχανικής μάθησης με στόχο τη βελτίωση της ερμηνευσιμότητας και της ακρίβειας του μοντέλου. Η διαδικασία επιλογής του βέλτιστου μοντέλου περιλαμβάνει βελτιστοποίηση υπερπαραμέτρων, ανάλυση συναρτήσεων σφάλματος και έλεγχο εγκυρότητας, εξασφαλίζοντας άριστη προγνωστική απόδοση. Πραγματοποιήθηκαν συγκριτικές αναλύσεις μεταξύ καθαρών φυσικών μοντέλων, αμιγώς δεδομενο-κεντρικών μοντέλων μηχανικής μάθησης και υβριδικών προσεγγίσεων grey-box, αξιολογώντας τη γενίκευση και την υπολογιστική αποδοτικότητα.

Η μελέτη συζητά επίσης την πιθανή ενσωμάτωση του αναπτυγμένου μοντέλου πρόβλεψης σε ένα Σύστημα Υποστήριξης Αποφάσεων (DSS), περιγράφοντας πώς θα μπορούσε να χρησιμοποιηθεί για εκτίμηση της ταχύτητας σε πραγματικό χρόνο και βελτιστοποίηση του ταξιδιού.

Η έρευνα αυτή συμβάλλει στην προώθηση της μηχανικής μάθησης στη ναυτιλία, προτείνοντας μια μεθοδολογία που βελτιώνει την απόδοση των πλοίων και υποστηρίζει τη λήψη αποφάσεων με γνώμονα τα δεδομένα.

Λέξεις-κλειδιά: Πρόβλεψη Ταχύτητας Πλοίου, Βελτιστοποίηση Ταξιδιού, Μηχανική Μάθηση, Μοντέλα Δεδομένων, Θαλάσσιες Μεταφορές, Νευρωνικά Δίκτυα

Abstract

The accurate prediction of ship speed and power requirements is a critical aspect of maritime operations, impacting fuel efficiency, regulatory compliance, and voyage optimization. Traditional methods for estimating ship speed-power relationships rely on semi-empirical resistance models and sea trial interpolations, which often fail to account for real-world operational conditions such as wind, waves, and dynamic engine performance. To address these limitations, this study develops a data-driven machine learning model for ship speed prediction, integrating physical modeling techniques with modern artificial intelligence approaches.

The research begins with a comprehensive exploratory data analysis (EDA) of ship operational datasets, including sensor-derived and noon report data. Techniques such as outlier detection, feature selection, and correlation analysis are applied to refine the dataset for modeling purposes. The study then implements a hybrid prediction framework, incorporating both physics-based models (e.g., semi-empirical resistance models, sea trial interpolations, and Physics-Informed Neural Networks (PINNs)) and machine learning algorithms (e.g., Linear regression, Random Forest Regressor, and XGBoost).

A parallel grey-box modeling approach is introduced, integrating physics-based constraints into machine learning predictions to improve model interpretability and accuracy. The model selection process involves hyperparameter optimization, error function analysis, and validation testing, ensuring optimal predictive performance. Comparative analyses between pure physics-based, machine learning-based, and hybrid grey-box models are conducted to evaluate generalization ability and computational efficiency.

The study also discusses the potential integration of the developed prediction model into a Decision Support System (DSS), outlining how it could be used for real-time power estimation and operational optimization.

This research contributes to the advancement of machine learning applications in maritime, offering a methodology that enhances operational efficiency, and supports data-driven decision-making in modern shipping operations.

Keywords: Ship Speed Prediction, Voyage Optimization, Machine Learning, Data-Driven Models, Maritime Transportation, Neural Networks

Table of Contents

List of Figures

List of Tables

Abbreviations

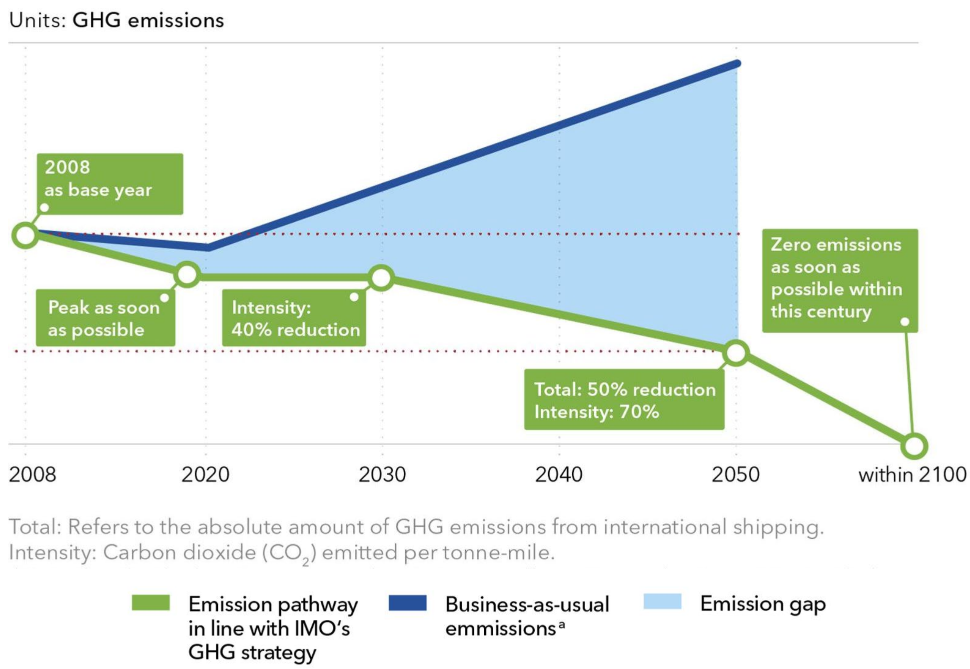
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BBM	Black-Box Model
CFD	Computational Fluid Dynamics
CII	Carbon Intensity Indicators
DSS	Decision Support System
DT	Decision Tree
EDA	Exploratory Data Analysis
EEDI	Efficiency Design Index
EEXI	Energy Efficiency Existing Ship Index
EU ETS	European Union's Emissions Trading System
GBM	Grey-Box Model
GHG	Greenhouse Gas
GPS	Global Positioning System
IMO	International Maritime Organization
ISO	International Standardization Organization
LIDAR	Light Detection and Ranging
LNG	Liquified Natural Gas
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCR	Maximum Continuous Rating
ML	Machine Learning
MSE	Mean Square Error
PDE	Partial Differential Equation
PINN	Physics-Informed Neural Network
QSS	Quasi-Steady-State
RANS	Reynolds-averaged Navier-Stokes
RF	Random Forest
RMSE	Root Mean Square Error
RUL	Remaining Useful Life
SFOC	Specific Fuel Oil Consumption
SOG	Speed Over Ground
STW	Speed Through Water
SVM	Support Vector Machines
UNFCCC	United Nations Framework Convention on Climate Change
UTC	Coordinated Universal Time
UWC	Underwater Cleaning
WBM	White-Box Model

~ This page was intentionally left blank. ~

0. Εκτεταμένη Ελληνική Περίληψη

0.1. Εισαγωγή

Η ναυτιλιακή βιομηχανία αντιπροσωπεύει περίπου το 3% των παγκόσμιων εκπομπών CO_2, γεγονός που καθιστά επιτακτική την ανάγκη για αποδοτικότερες και πιο φιλικές προς το περιβάλλον τεχνολογίες. Ρυθμιστικές πρωτοβουλίες, όπως η MARPOL Annex VI και το EU ETS, επιβάλλουν αυστηρότερα όρια στις εκπομπές, ενώ η Διεθνής Ναυτιλιακός Οργανισμός (IMO) έχει θέσει ως στόχο τη μείωση των εκπομπών κατά 50% έως το 2050.



Εικόνα 1 Όρια στόχων μείωσης εκπομπών

Επιπλέον, το κόστος των καυσίμων αντιπροσωπεύει το 50-60% των λειτουργικών εξόδων ενός πλοίου, ενώ αν δεν ληφθούν μέτρα, οι εκπομπές CO_2 μπορεί να αυξηθούν κατά 90-130% μέχρι το 2050. Σε αυτό το πλαίσιο, η πρόβλεψη της ταχύτητας ενός πλοίου υπό διαφορετικές συνθήκες λειτουργίας καθίσταται κρίσιμη τόσο για τη μείωση της κατανάλωσης καυσίμου όσο και για τη βελτίωση της επιχειρησιακής αποδοτικότητας.

Η παρούσα μελέτη διερευνά τη χρήση μοντέλων μηχανικής μάθησης (Machine Learning) για την ακριβέστερη πρόβλεψη της ταχύτητας πλοίου, αξιοποιώντας δεδομένα από αισθητήρες και noon reports. Ο στόχος είναι η ανάπτυξη ενός αξιόπιστου μοντέλου που θα υποστηρίξει την επιχειρησιακή λήψη αποφάσεων στον ναυτιλιακό τομέα.

0.2. Βιβλιογραφική Επισκόπηση & Θεωρητικό Υπόβαθρο

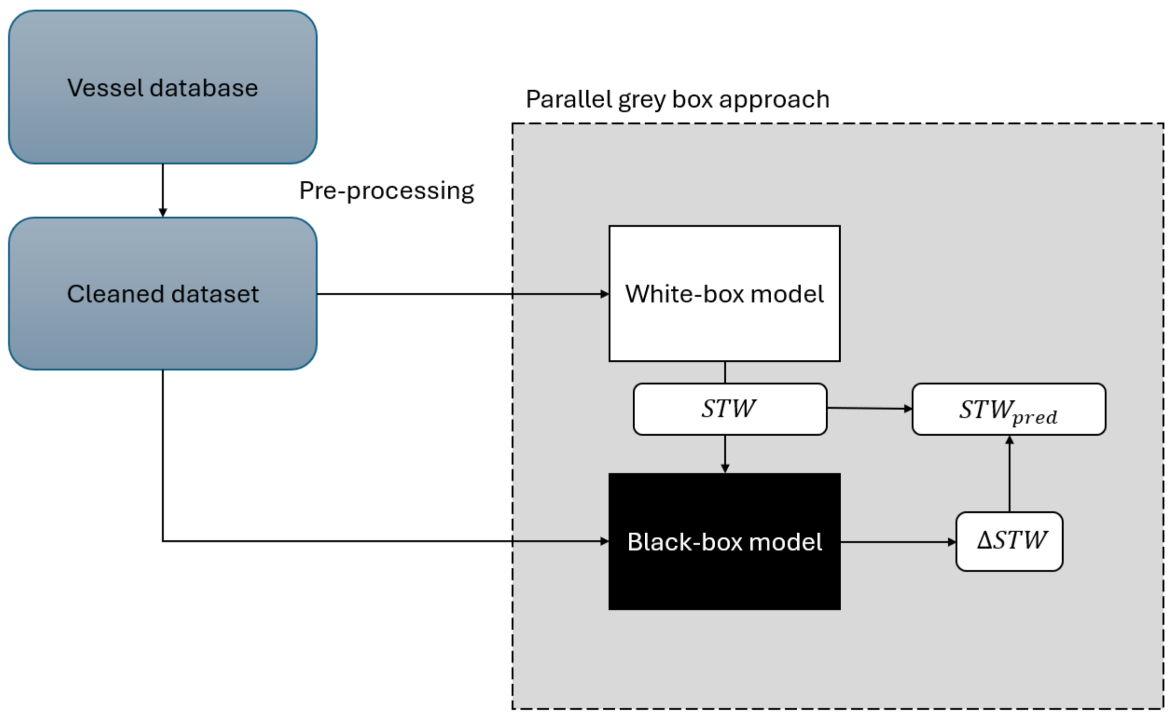
Η πρόβλεψη της ταχύτητας των πλοίων έχει μελετηθεί τόσο μέσω παραδοσιακών προσεγγίσεων, όπως η εμπειρική μοντελοποίηση και οι φυσικές προσομοιώσεις, όσο και με τη χρήση σύγχρονων μεθόδων δεδομένων. Οι φυσικές μέθοδοι βασίζονται σε αναλυτικές εξισώσεις και πειραματικά δεδομένα από θαλάσσιες δοκιμές, ενώ τα δεδομένα πλοίων που συλλέγονται από αισθητήρες επιτρέπουν μια πιο δυναμική και προσαρμοστική προσέγγιση.

Η μηχανική μάθηση έχει αποδείξει την ικανότητά της να αποτυπώνει μη γραμμικές σχέσεις μεταξύ μεταβλητών, προσφέροντας πιο ακριβείς προβλέψεις σε σχέση με τα συμβατικά μοντέλα. Στη ναυτιλία, μέθοδοι όπως τα XGBoost, Random Forest και Physics-Informed Neural Networks (PINNs) έχουν χρησιμοποιηθεί για την πρόβλεψη της σχέσης μεταξύ ταχύτητας, ισχύος και περιβαλλοντικών παραμέτρων.

0.3. Μεθοδολογία

Η έρευνα ακολούθησε μια δομημένη προσέγγιση ανάλυσης δεδομένων, η οποία περιλάμβανε τα εξής στάδια:

- **Συλλογή και Προεπεξεργασία Δεδομένων**: Χρησιμοποιήθηκαν δεδομένα από αισθητήρες και noon reports, λαμβάνοντας υπόψη παραμέτρους όπως το βύθισμα (draft), η ταχύτητα ως προς το νερό (STW), η ταχύτητα ως προς το έδαφος (SOG), οι καιρικές συνθήκες, και η ισχύς της μηχανής. Δόθηκε ιδιαίτερη έμφαση στην εξασφάλιση της ποιότητας των δεδομένων και στην εξάλειψη πιθανών σφαλμάτων καταγραφής ή μη διαθέσιμων τιμών.
- **Εξερεύνηση και Ανάλυση Δεδομένων**: Εφαρμόστηκε ανάλυση συσχέτισης μεταξύ των μεταβλητών για να εντοπιστούν οι κρίσιμοι παράγοντες που επηρεάζουν την ταχύτητα. Επίσης, εξετάστηκε η στατιστική κατανομή των δεδομένων και εντοπίστηκαν ανωμαλίες ή αποκλίσεις που θα μπορούσαν να επηρεάσουν την απόδοση των μοντέλων.
- **Ανάπτυξη Μοντέλου**: Δοκιμάστηκαν διάφορες μέθοδοι μηχανικής μάθησης, όπως τα XGBoost και Random Forest, ενώ εξετάστηκε και η χρήση Physics-Informed Neural Networks (PINNs) για την ενσωμάτωση φυσικών περιορισμών στο μοντέλο. Έγινε λεπτομερής σύγκριση των μεθόδων με στόχο την επιλογή της βέλτιστης προσέγγισης.
- **Εκπαίδευση και Αξιολόγηση**: Χρησιμοποιήθηκαν κατάλληλοι δείκτες σφάλματος (π.χ. RMSE, MAE) για την αξιολόγηση της απόδοσης των μοντέλων, με έμφαση στην ικανότητά τους να προβλέπουν την ταχύτητα με ακρίβεια υπό διαφορετικές συνθήκες λειτουργίας. Η διαδικασία επικύρωσης των αποτελεσμάτων περιλάμβανε τη χρήση ανεξάρτητων συνόλων δοκιμών και τη σύγκριση των μοντέλων με πραγματικά δεδομένα λειτουργίας πλοίων.



Εικόνα 2 Διαγραμματική απεικόνιση της προσέγγισης που ακολουθήθηκε στην εργασία

0.4. Αποτελέσματα και Συζήτηση

Τα αποτελέσματα έδειξαν ότι τα μοντέλα μηχανικής μάθησης προσφέρουν σημαντικά βελτιωμένη ακρίβεια πρόβλεψης σε σχέση με τις παραδοσιακές προσεγγίσεις. Το XGBoost εμφάνισε τη βέλτιστη απόδοση μεταξύ των δοκιμασμένων μεθόδων, ενώ η προσέγγιση PINN έδειξε ενδιαφέρουσες προοπτικές ενσωμάτωσης φυσικών νόμων στη διαδικασία πρόβλεψης.

Επιπλέον, η μελέτη ανέδειξε τη σημασία της σωστής προεπεξεργασίας των δεδομένων, ιδίως όσον αφορά τη διαχείριση των ελλιπών τιμών και την ορθή αντιστοίχιση των δεδομένων από διαφορετικές πηγές (αισθητήρες και noon reports). Ο συνδυασμός αυτών των δεδομένων βελτίωσε την κατανόηση της σχέσης μεταξύ STW, SOG και των εξωτερικών παραγόντων.

Μια ακόμη κρίσιμη διαπίστωση ήταν η διαχείριση των θαλάσσιων ρευμάτων, τα οποία επηρεάζουν τη σχέση μεταξύ STW και SOG. Η έρευνα ανέδειξε τη σημασία της ενσωμάτωσης πρόσθετων μεταβλητών, όπως η δύναμη και κατεύθυνση του ανέμου, για τη βελτίωση της ακρίβειας των προβλέψεων.

0.5. Συμπεράσματα

Η έρευνα κατέδειξε ότι οι τεχνικές μηχανικής μάθησης μπορούν να παρέχουν ακριβείς προβλέψεις για την ταχύτητα των πλοίων, προσφέροντας ένα πολύτιμο εργαλείο για την υποστήριξη λήψης αποφάσεων στη ναυτιλία. Η χρήση δεδομένων αισθητήρων σε συνδυασμό με προηγμένες μεθόδους ανάλυσης συμβάλλει στην καλύτερη κατανόηση της απόδοσης του πλοίου και στη βελτιστοποίηση των ναυτιλιακών λειτουργιών.

Μελλοντικές επεκτάσεις της έρευνας θα μπορούσαν να περιλαμβάνουν τη βελτίωση των μοντέλων με περισσότερα δεδομένα, τη χρήση deep learning τεχνικών και την ανάπτυξη ενός ολοκληρωμένου συστήματος υποστήριξης αποφάσεων που θα ενσωματώνει δυναμικά τις προβλέψεις του μοντέλου στο πλαίσιο επιχειρησιακών στρατηγικών. Επιπλέον, η μελέτη της επίδρασης της υδροδυναμικής αντίστασης και η ενσωμάτωση φυσικών μοντέλων θα μπορούσαν να συμβάλουν στην περαιτέρω αύξηση της ακρίβειας των προβλέψεων και στη δημιουργία μιας ολοκληρωμένης προσέγγισης στη ναυτιλιακή απόδοση.

1. Introduction

1.1. Motivation

Climate change has emerged as one of the most pressing global challenges, with international bodies such as the United Nations Framework Convention on Climate Change (UNFCCC) and the International Maritime Organization (IMO) taking significant steps to mitigate its effects. The adoption of the Paris Agreement in 2015 [1] marked a global commitment to reducing greenhouse gas (GHG) emissions, with the goal of limiting global temperature rise to well below 2°C, while striving to keep the increase below 1.5°C relative to pre-industrial levels. Achieving these targets necessitates reaching net-zero emissions by 2050, driving a fundamental transformation across energy production, consumption, and industrial sectors, including the maritime industry.

The Role of the Maritime Industry in Global Trade and Emissions

Maritime transport remains a fundamental component of global trade, accounting for over 80% of total merchandise transportation. In 2023, global seaborne trade volumes reached 12.3 billion tonnes, reflecting a 2.4% increase following a previous contraction. This growth was largely driven by increased demand for commodities, vessel rerouting due to geopolitical tensions, and disruptions in key maritime chokepoints [2]. As of 2024, the global fleet comprises approximately 150,000 vessels, with an estimated total value of $2 trillion [3]. The distribution of fleet value across vessel types is as follows:

- Offshore vessels: 22.4%
- Bulk carriers: 18.4%
- Tankers: 13.2%
- Containerships: 10.6%
- Gas carriers: 9.1%
- Other vessels: 26.3%

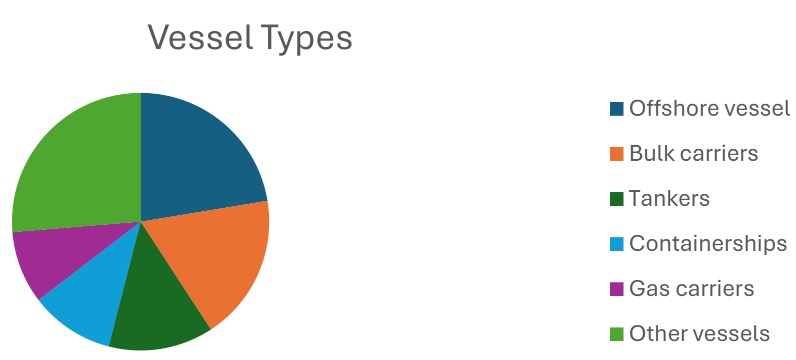


Figure 1 Distribution of fleet across vessel types

In terms of deadweight tonnage (DWT), bulk carriers (41.2%) and oil tankers (28.2%) dominate the fleet composition. Conversely, containerships, gas carriers, and offshore vessels, while constituting a smaller percentage of DWT, hold significant market value.

Fleet segmentation by gross tonnage (GT) is as follows [4] :

- Small vessels (GT < 500): 37%
- Medium vessels (500 ≤ GT < 25,000): 43%
- Large vessels (25,000 ≤ GT < 60,000): 13%
- Very large vessels (GT ≥ 60,000): 7%

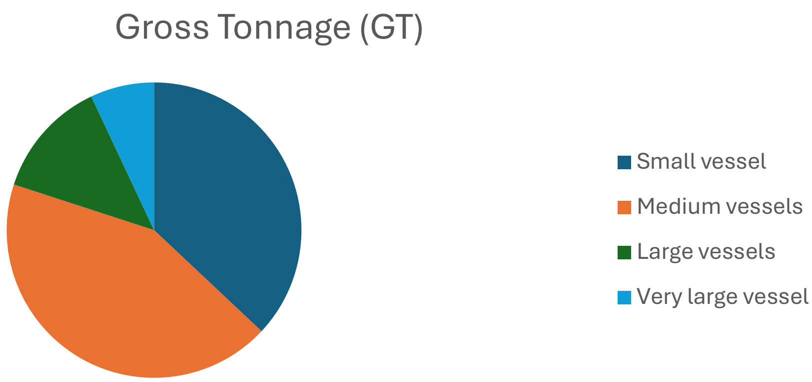


Figure 2 Distribution of fleet by gross tonnage

This classification is significant as vessel size directly influences operational efficiency, fuel consumption, and regulatory compliance, particularly in speed prediction models.

Decarbonization Challenges in the Maritime Sector

The maritime industry experienced a 4.2% increase in ton-miles in 2023, the highest in 15 years, indicating longer shipping distances due to geopolitical rerouting and climate-related disruptions [2]. However, industry challenges persist:

- Geopolitical risks: Ongoing conflicts impact trade routes and increase supply chain volatility.
- Regulatory compliance: The IMO 2020 sulfur cap and upcoming IMO decarbonization targets for 2030 and 2050 require increased adoption of alternative fuels and energy-efficient technologies.
- Economic pressures: Fluctuations in crude oil prices and rising operational costs place financial constraints on shipowners.
- Environmental and climate risks: Increased weather-related disruptions and stricter emissions regulations necessitate improved fleet management strategies.

The maritime sector, responsible for approximately 3% of global total emissions [5], has been identified as a "hard-to-abate" industry due to its reliance on fossil fuels, long asset lifespans, and the operational challenges posed by electrification [6] . Recognizing the need for a more sustainable future, the IMO established the goal of reducing carbon intensity per transport work by at least 40% by 2030, with further reductions of up to 70% by 2050 [7]. However, as these targets fall short of full decarbonization, the international community has called for further action, including the adoption of stricter regulations and incentives for low- and zero-carbon fuels [8].

Industry Response and the Need for Data-Driven Decision-Making

In response, the shipping industry has pursued energy-efficient technologies and alternative fuels, such as liquefied natural gas (LNG), biofuels, and hydrogen-based fuels. Additionally, initiatives like the European Union's Emissions Trading System (EU ETS) have begun to apply market-based measures to reduce carbon emissions from maritime transport [9]. However, the pathway to achieving full decarbonization is complex and requires a multi-faceted approach that combines technological innovation, regulatory frameworks, and operational efficiency improvements.

One of the critical areas of focus is optimizing vessel performance, particularly the relationship between ship speed and power requirements. This relationship is essential for reducing fuel consumption and emissions, as the power needed to propel a ship increases exponentially with speed [10]. Accurately predicting speed-power dynamics under varying operational conditions can significantly improve decision-making for ship operators, enabling more efficient routing and speed adjustments to minimize energy use.

1.2. Problem Statement

The maritime industry faces increasing pressure to reduce greenhouse gas emissions while maintaining operational efficiency. With the introduction of stringent regulations such as the EU Emissions Trading System and the IMO's decarbonization goals, ship operators must find cost-effective ways to comply. One key challenge lies in optimizing the speed-power relationship of vessels, which directly influences fuel consumption and emissions. The power required to propel a ship increases exponentially with speed, yet existing methods of speed-power prediction often fail to account for critical real-world variables, including weather conditions, hull resistance, and engine performance. Moreover, operational complexities such as biofouling, propeller inefficiencies, and dynamic factors like wind and currents can significantly impact a ship's energy consumption [10].

This thesis addresses these challenges by developing machine learning models that predict ship speed based on power input, while considering a wide range of operational and environmental factors. Leveraging data from a real-time fuel optimization system installed on a tanker vessel, the model aims to provide enhanced predictive capabilities, enabling operators to make informed decisions regarding speed adjustments and power management, ultimately improving vessel efficiency and reducing fuel costs. By accurately modeling the speed-power relationship, the research seeks to provide insights that support decision-making processes for speed adjustments and power management, in line with the goals of the IMO and the EU to decarbonize shipping. Furthermore, the study advances data-driven methods in maritime operations, which are essential for navigating regulatory challenges and achieving climate targets.

1.3. Objectives

The primary objective of this thesis is to develop a machine learning model for ship speed prediction based on power inputs, integrating data-driven techniques with maritime operational knowledge. This research aims to enhance vessel performance monitoring, optimize fuel consumption, and support decision-making for speed and power management in line with decarbonization efforts.

First, the study will investigate and compare ship speed prediction methods by reviewing traditional estimation techniques and identifying their limitations. It will explore machine learning approaches and assess their potential improvements over conventional models.

Second, the research will focus on developing and evaluating a machine learning model for ship speed prediction. This includes conducting exploratory data analysis (EDA) to identify key influencing factors, implementing physics-based, data-driven, and hybrid (grey-box) models, and optimizing model performance through feature selection, hyperparameter tuning, and validation with real-world vessel data.

The third objective is to assess the practical applications of the developed model in vessel operations. The study will demonstrate how predictive modeling can assist ship operators in voyage planning, speed optimization, and fuel efficiency strategies, contributing to compliance with IMO and EU decarbonization targets.

Finally, the research will identify challenges and potential improvements in ML-based ship performance modeling. It will analyze the limitations of the developed model in real-world applications and suggest areas for future research and enhancements in maritime data analytics. Additionally, it will highlight how predictive models can be integrated into DSS frameworks to support ship operators in making data-driven speed and power management decisions.

1.4. Work Structure

This thesis is structured into eight chapters, each addressing different aspects of the research on machine learning-based ship speed prediction.

Chapter 1: Introduction provides motivation, problem statement, objectives, and work structure of the study. It outlines the significance of optimizing vessel operations through predictive modeling and highlights the role of machine learning in addressing maritime decarbonization challenges.

Chapter 2: Literature Review examines existing research on machine learning applications in maritime operations, traditional ship speed and power estimation methods, and the evolution of data-driven models in vessel performance analysis. This chapter identifies the gaps in current approaches and establishes the rationale for integrating machine learning techniques into speed prediction models.

Chapter 3: Theoretical Background presents the fundamental principles relevant to the study. It includes a taxonomy of machine learning algorithms, an overview of ship resistance and power dynamics, and a discussion on data preprocessing and feature engineering techniques used to enhance model accuracy.

Chapter 4: Case Study focuses on the dataset used in this research. It describes data sources, preprocessing steps, exploratory data analysis (EDA), feature selection methods, and outlier detection techniques. These steps are crucial for preparing high-quality input data for machine learning modeling.

Chapter 5: Prediction Model details the development and evaluation of the ship speed prediction model. It introduces a semi-empirical physics-based resistance model as a baseline, followed by machine learning approaches to enhance predictive accuracy. The chapter explores a grey-box modeling approach, integrating physics-based data-driven methodologies, and evaluates model performance based on real-world vessel data.

Chapter 6: Decision Support System discusses how the developed speed prediction model could be utilized in decision support frameworks for vessel operations. This thesis highlights potential applications of DSS in voyage planning, speed optimization, and fuel efficiency strategies, aligning with industry sustainability goals.

Chapter 7: Conclusion and Future Work summarizes the key findings, discusses the limitations of the study, and proposes future research directions to further enhance ML-based ship speed and power prediction models.

2. Literature Review

2.1. Machine Learning Applications in Maritime Operations

The maritime industry is increasingly adopting machine learning (ML) and data-driven methodologies to address complex challenges and enhance decision-making processes. ML algorithms, by analyzing large datasets, can predict patterns, identify anomalies, and optimize processes across a wide range of maritime operations. These applications provide innovative solutions that improve efficiency, safety, and environmental sustainability as global shipping continues to expand and confront regulatory pressures related to decarbonization. Machine learning, a subset of artificial intelligence, enables systems to learn from data and make predictions without explicit programming. As global shipping continues to expand, the need for efficient, sustainable practices has become paramount. According to Yan and Wang (2022) in their book “Applications of Machine Learning and Data Analytics Models in Maritime Transportation” [11], data-driven methodologies are transforming traditional maritime operations by providing innovative solutions that improve efficiency, safety, and environmental sustainability.

The importance of ML in the maritime industry cannot be overstated. By leveraging ML techniques, stakeholders can optimize ship performance, enhance fuel efficiency, and facilitate operational decision-making. For instance, predictive maintenance models can reduce downtime by identifying potential failures before they occur, while fuel optimization algorithms enable more efficient routing and speed adjustments, significantly lowering emissions.

One of the most promising applications of ML is vessel health monitoring. Machine learning models trained in real-time sensor data can predict potential failures in engines, thrusters, and other vital mechanical systems. By leveraging predictive maintenance algorithms, ship operators can reduce unplanned downtime and maintenance costs. Techniques such as Long Short-Term Memory (LSTM) networks have shown efficacy in estimating the Remaining Useful Life (RUL) of critical components, ensuring timely maintenance actions before system failures occur. The transition from reactive to predictive maintenance enables more efficient resource management, contributing to both operational safety and cost reduction [12].

ML has also become a key enabler in the development of autonomous ships. These systems rely on machine learning algorithms to analyze real-time data from multiple sensors, such as GPS, radar, and LIDAR, to enable autonomous navigation, obstacle detection, and route adjustments. For instance, machine learning models can process environmental and navigational data in real-time, allowing ships to make operational adjustments autonomously, enhancing both safety and efficiency. The integration of ML into autonomous navigation systems has the potential to reduce human error, lower operational costs, and improve safety in challenging marine environments [12].

In maritime logistics and port operations, machine learning enhances the efficiency of container handling, ship scheduling, and cargo flow prediction. Predictive analytics tools can forecast port congestion, improving berthing schedules and minimizing waiting times. By using data-driven methods to optimize cargo distribution, ML helps ports to streamline operations, reduce delays, and increase throughput [11]. Additionally, ML models can optimize fleet management, ensuring that vessels operate on the most efficient routes, with minimal fuel consumption and optimal cargo loading configurations.

Machine learning is gaining traction in the maritime industry, offering innovative ways to optimize vessel performance, improve operational efficiency, and address regulatory challenges. These advancements mark a significant shift from traditional methods, which, while foundational, are increasingly being supplemented by data-driven approaches. As this chapter introduces the role of ML in shipping, the traditional methods for estimating vessel speed and power, which continue to play a vital role in maritime engineering, will be presented in detail in the following chapter.

2.2. Traditional Methods of Ship Speed and Power Prediction

The prediction of ship speed and power requirements has long relied on traditional methods, such as empirical formulas and physical models, which form the foundation of maritime engineering. These methods provide a reliable basis for ship design and operational planning but struggle to account for the complex and dynamic conditions encountered during real-world voyages.

One of the most widely used empirical methods is the cube law, also known as the propeller law, which approximates the relationship between a ship's required power and its speed through water as cubic:

$$P(v) = kv^3$$

where:

- v is the ship's speed through water (knots),
- $P(v)$ is the ship's required power, including main engines, boilers, and auxiliary engines (kW),
- k is a constant derived from the ship's characteristics [10].

This formula, derived from basic hydrodynamic principles, suggests that small reductions in speed can lead to significant reductions in power and fuel consumption. However, this approximation is primarily under ideal conditions, such as calm water and constant speed, and does not account for real-world variables such as waves, wind, and currents.

More advanced forms of cube law, like the modified admiralty formula, extend this approach by incorporating the ship's displacement and payload:

$$P(v, w) = m(A + w)^{2/3} v^n$$

where:

- w is the ship's payload (tonnes),
- A is the lightship weight (tonnes),
- n is typically ≥ 3,

- m is a constant.

This formula accounts for the wetted surface area of the hull, which affects resistance, and provides a more refined estimate of power requirements, especially at different loading conditions [10].

While the cube law provides a useful approximation for calm water resistance, real-world conditions often introduce additional complexities. For example, a ship's resistance in waves, known as added resistance, is highly nonlinear. Added resistance depends on factors such as hull shape, sea state, and the ship's seakeeping characteristics, which are not captured by simple empirical formulas. Estimating added resistance requires more sophisticated methods, such as strip theory approximation [13] or other hydrodynamic models [10].

The "Basic Principles of Ship Propulsion" further highlight the relationship between resistance and fuel consumption [14]. The total resistance RT of a ship in calm water is a combination of frictional and residual resistance, which can be described as:

$$RT = \frac{1}{2}\rho C_T v^2 S$$

where:

- ρ is the water density,
- C_T is the total drag coefficient (sum of the frictional and residual drag coefficients),
- S is the wetted surface area of the hull.

This equation shows that, even in calm water, the relationship between speed and resistance is quadratic, reinforcing the importance of understanding how real-world conditions affect power and fuel consumption. Traditional methods like the cube law tend to oversimplify these dynamics, leading to less accurate predictions in varying operational environments. Moreover, empirical formulas and hydrodynamic models often fail to account for operational factors like biofouling, propeller slip, and sea state, which can significantly impact performance over time. Biofouling increases hull resistance, while propeller slip—caused by currents or changes in sea conditions—reduces the efficiency of propulsion. As the "Basic Principles of Ship Propulsion" document notes, operators often rely on performance benchmarks from sea trials to measure these factors during normal operations, but these methods provide limited flexibility for real-time adjustments [14].

Computational Fluid Dynamics (CFD) is an advanced method used to simulate fluid flow around a ship's hull and predict its hydrodynamic performance. Unlike empirical formulas, which are based on simplified assumptions, CFD offers a more comprehensive approach by solving the Reynolds-averaged Navier-Stokes (RANS) equations. These equations govern fluid motion and allow for detailed simulations of how water flows around the hull, accounting for factors like turbulence, pressure distribution, and wave-structure interactions [15]. Historically, CFD models have been used at model scale, where simulations are validated using experimental data from testing tanks. However, advancements in computational power have enabled the shift toward full-scale simulations, which provide more accurate predictions of ship performance under real-world conditions. As highlighted in the Siemens white paper, full-scale CFD simulations eliminate the scaling errors associated with model tests, particularly in terms of boundary layer behavior and drag coefficients. Full-scale simulations also better represent the effects of environmental conditions such as wave heights, currents, and wind [15].

CFD models offer significant advantages over traditional empirical methods by providing detailed insights into specific operational conditions. For example, CFD can predict how hull resistance changes under different wave amplitudes, directions, and speeds, which is particularly valuable for voyage optimization. Ship operators can use CFD simulations to determine optimal speeds, heading angles, and power settings under varying sea states, leading to more efficient fuel usage and reduced emissions. Despite its accuracy, CFD simulations are computationally intensive, especially for full-scale simulations that require millions of grid cells to resolve the flow field around the hull. The complexity of these simulations limits their application for real-time operational decisions. In most cases, CFD is used during the design phase to optimize hull shapes, propeller designs, and engine configurations rather than during voyages. This makes CFD a powerful tool for improving the energy efficiency of new ship designs but less suitable for dynamic, on-the-fly operational adjustments [15]. Furthermore, CFD's reliance on high-quality numerical grids and turbulence models introduces additional challenges. Simulations must be carefully set up to ensure grid convergence and accurate resolution of boundary layers, particularly in the turbulent flow regime. Even slight errors in grid resolution can lead to significant discrepancies in predicted drag forces, making it essential to balance computational cost with simulation accuracy.

Semi-empirical models represent a middle ground between empirical formulas and CFD, combining experimental data with theoretical principles to improve accuracy. These models often incorporate specific experimental results, such as wave resistance or hull form tests, and apply them within a theoretical framework to predict ship performance under a variety of conditions. For example, Lang (2021) developed a semi-empirical model to estimate a ship's added resistance in waves based on experimental data [16]. Although this model improved accuracy compared to purely empirical methods, it still struggled to adapt to real-time variations in environmental factors. Lang's work highlighted that more accurate predictions, especially in complex sea states, required machine learning models that could process large datasets and respond to a wider range of conditions.

Despite the utility of these traditional methods, they rely on simplifying assumptions that limit their accuracy in real-world conditions. Assumptions such as calm water, steady-state conditions, or a constant power-speed relationship make these models less effective in handling dynamic factors like changing sea states, wind, currents, or biofouling. As a ship's hull fouls over time or as environmental conditions fluctuate, the predictions of traditional models become less reliable, often resulting in suboptimal performance and higher fuel consumption. Moreover, while CFD and semi-empirical models offer a more accurate representation of ship performance, they remain computationally expensive and require significant resources to set up and run simulations. This makes them impractical for real-time operational adjustments during voyages, where quick and adaptable decisions are needed to optimize speed and fuel consumption. This gap has driven the development of data-driven models, particularly those based on machine learning, which offer more accurate predictions by incorporating a wide range of operational and environmental factors in real-time. These models can process vast amounts of operational data, accounting for a broader range of variables, including weather patterns, hull conditions, and traffic patterns, in real time. By continuously learning from both historical and real-time data, machine learning models offer the potential for more dynamic, responsive ship performance optimization.

2.3. Data-Driven Models in Vessel Performance

Data-driven models, particularly those based on machine learning (ML) techniques, offer a promising alternative by providing more accurate, real-time predictions of ship speed, power, and fuel consumption under varying operational and environmental conditions.

Machine learning applications in vessel performance optimization can be categorized into three key areas: fuel consumption prediction, speed prediction, and overall operational efficiency.

In fuel consumption prediction, ML models help predict fuel usage based on various operational parameters, such as speed and sea conditions. For example, Madureira (2021) developed ML models using historical operational data to optimize fuel consumption on ships [17]. Additionally, the study “A Deep Learning Method for the Prediction of Ship Fuel Consumption in Real Operational Conditions” demonstrates a deep learning approach to fuel consumption prediction, showcasing how advanced techniques can significantly enhance predictive accuracy [18].

For speed prediction, algorithms such as artificial neural networks (ANNs) and support vector machines (SVM) are utilized to predict ship speed under different operational conditions, aiding in voyage planning and optimization. The paper “Development of Neural Networks for Ship Speed Prediction” focuses on creating models that can enhance the accuracy of speed predictions using various neural network architectures [19]. Lang (2021) focused on developing speed-power performance models, leveraging the XGBoost algorithm, resulting in a significant reduction in prediction errors compared to traditional models [16]. Furthermore, the work on “Speed-Power Models – A Bayesian Approach” highlights the use of Bayesian methods for modeling speed-power relationships, which can provide valuable insights into optimizing vessel performance [20]. Additionally, Lang et al. (2022) proposed physics-informed ML models that integrate physical principles with machine learning, providing a more accurate approach to speed predictions [21].

Regarding operational efficiency, ML is applied to improve overall operational efficiency by integrating data from various sources, including weather forecasts, vessel performance, and market dynamics, to inform strategic decision-making. The study titled “Machine Learning Techniques for Modeling Ships Performance in Waves” explores the integration of real-time data to enhance operational performance [22]. The review of various applications in “Machine Learning for Naval Architecture, Ocean, and Marine Engineering” also provides insight into how ML can optimize different operational facets [23].

In the detailed review of key studies, Zhang et al. (2024) developed a deep learning method to predict ship fuel consumption in real operational conditions [18]. The model incorporates an attention mechanism into a Bi-directional Long Short-Term Memory (Bi-LSTM) network to capture the complex relationships between operational data inputs, such as sailing speed, heading, displacement, trim, weather, and sea conditions, and fuel consumption. The method uses data from sensors, voyage reporting, and hydrometeorological information comprising 266 variables. This approach demonstrated a significant improvement in prediction accuracy when compared to existing methods, highlighting the potential of Bi-LSTM with attention mechanisms in optimizing fuel consumption and supporting decision-making for environmentally sustainable ship operations.

Lang (2021) developed speed-power performance models for arbitrary wave headings, focusing on ship voyage optimization [16]. The thesis introduces two models: a semi-empirical model and a machine learning-based model. The semi-empirical model estimates a ship's added resistance in head waves and extends this to different wave headings, verified by experimental model tests. A significant wave height-based correction factor is introduced to account for the nonlinear effects of irregular waves on a ship's resistance and power increase. The machine learning model, developed using the XGBoost algorithm, leverages full-scale measurement data from a PCTC vessel, with input features including operational profiles, metocean conditions, and motion responses. The machine learning model outperformed the semi-empirical model, reducing the discrepancy between power predictions and actual values from over 40% to less than 1%. These models significantly improve voyage optimization, leading to reduced fuel consumption and increased energy efficiency.

Madureira (2021) aimed to develop machine learning models to represent the operation of a fuel optimization system and to create a prototype decision support system that predicts optimal fuel consumption [17]. The study used a one-year dataset collected from a ship's automated fuel optimization system, which included data on propulsion system parameters, environmental conditions, and fuel consumption. After pre-processing and analysis, the data were used to train machine learning models using Artificial Neural Networks (ANN) and Support Vector Machines (SVM) algorithms. The performance of these algorithms was evaluated, and a two-stage model was developed to predict both ship speed and fuel consumption under operational conditions. These models were integrated into a decision support system, which was demonstrated in different operational scenarios, showing potential for optimizing fuel efficiency and supporting operational decision-making.

Lang et al. (2023) proposed a novel hybrid model that integrates physics-informed approaches with machine learning to improve ship speed predictions [21]. The study utilized a grey-box model (GBM) approach, where the expected ship speed in calm water was modeled using Physics-Informed Neural Networks (PINNs) based on speed-power model tests. This was combined with the XGBoost machine learning algorithm to estimate ship speed reduction under actual weather conditions. The results demonstrated that the GBM significantly improved prediction accuracy compared to traditional black-box models, especially when sufficient data was available. Even with limited data, the GBM showed considerable improvements in speed prediction accuracy, making it a robust method for practical applications. The model was further validated through its implementation for Estimated Time of Arrival (ETA) predictions for cross-Pacific and North Atlantic voyages, showing a maximum cumulative error of only 5 hours

Grubišić et al. (2018) developed a system for monitoring and recording the influence of sea conditions on a vessel in motion using machine learning techniques to model the ship's performance in waves [22]. The system correlates measured wave parameters such as encounter angle, wave height, and wave amplitude with the vessel's motion characteristics. High-quality GRIB data from regions like the North Sea and Adriatic were used to generate training sets, and these correlations were stored in a neural network. The model is then applied alongside performance indicators like the root mean square (RMS) of linear acceleration, roll or pitch angle, and fuel consumption. This data helps create historical performance charts that assist in rational route planning and optimization. The study's experiments demonstrated its effectiveness in enhancing operational decision-making and optimizing ship performance under various sea conditions.

When comparing these studies, it is evident that different ML methodologies can yield significant improvements in predictive accuracy and operational efficiency. For example, deep learning models like RNNs have shown high accuracy in fuel consumption predictions, while tree-based methods like XGBoost excel in modeling speed-power relationships. However, each approach comes with its advantages and limitations, such as the need for extensive data for training deep learning models compared to the relatively simpler requirements of traditional regression models.

Despite the promising applications of ML in maritime operations, several challenges persist. Data quality and availability remain critical issues, as incomplete or inaccurate data can significantly affect model performance. Additionally, integrating ML solutions into existing operational frameworks poses its own set of difficulties, particularly regarding staff training and system compatibility.

In summary, the applications of machine learning in maritime operations are diverse and impactful, offering solutions that enhance efficiency and sustainability. The studies reviewed demonstrate the significant potential of ML techniques to optimize fuel consumption, improve speed predictions, and enhance operational decision-making. As industry continues to evolve, the integration of ML will play a crucial role in shaping the future of maritime transportation.

A critical focus in maritime operations is the relationship between ship speed, power, and fuel consumption, which is central to both operational efficiency and policy decisions. As discussed in "Ship Speed vs Power or Fuel Consumption: Are laws of physics still valid? Regression analysis pitfalls and misguided policy implications" [10], the non-linear relationship between speed and fuel consumption is key to reducing greenhouse gas emissions from ships. Reducing speed can lead to a disproportionate reduction in fuel consumption and emissions, making it a powerful tool in the short to medium term while the industry transitions to low-carbon fuels. The paper emphasizes that this speed reduction is critical in achieving compliance with regulations such as the Energy Efficiency Design Index (EEDI), Energy Efficiency Existing Ship Index (EEXI), and Carbon Intensity Indicators (CII), which are mandated by the International Maritime Organization (IMO) and the European Union (EU).

Furthermore, the paper highlights the risks of misinterpreting the speed-fuel relationship and the potential for misguided policies based on flawed analyses. By ensuring that policy decisions are grounded in scientifically sound models, the maritime sector can better navigate the complex trade-offs between operational efficiency and environmental sustainability. Machine learning models, which incorporate large-scale operational data and environmental factors, offer a robust solution for accurately predicting fuel consumption and optimizing speed to meet regulatory requirements while minimizing emissions. These models, by enhancing the accuracy of speed-power-fuel consumption predictions, can help ship operators comply with regulatory frameworks and optimize operations in a more sustainable manner.

Ultimately, machine learning techniques, when properly integrated into decision-making processes, hold great potential for bridging the gap between current operational practices and future sustainability goals. As the maritime industry continues to adopt more sophisticated data-driven methods, ML models will play a pivotal role in enabling more precise and effective measures to reduce GHG emissions and improve the energy efficiency of ships.

3. Theoretical Background

3.1. Ship Resistance and Power Dynamics

A ship's energy system is a complex interplay of various factors, including propulsion, heating, and auxiliary equipment, all of which contribute to overall energy consumption. In real-world conditions, a vessel's fuel consumption is influenced by multiple parameters such as marine engine performance, propeller efficiency, and total hydrodynamic resistance [24]. The relationship between propulsion power and ship speed is directly affected by environmental factors, including wind, waves, and currents. A standard approach to estimating a vessel's speed-power performance follows a structured workflow, as illustrated in Figure 3.

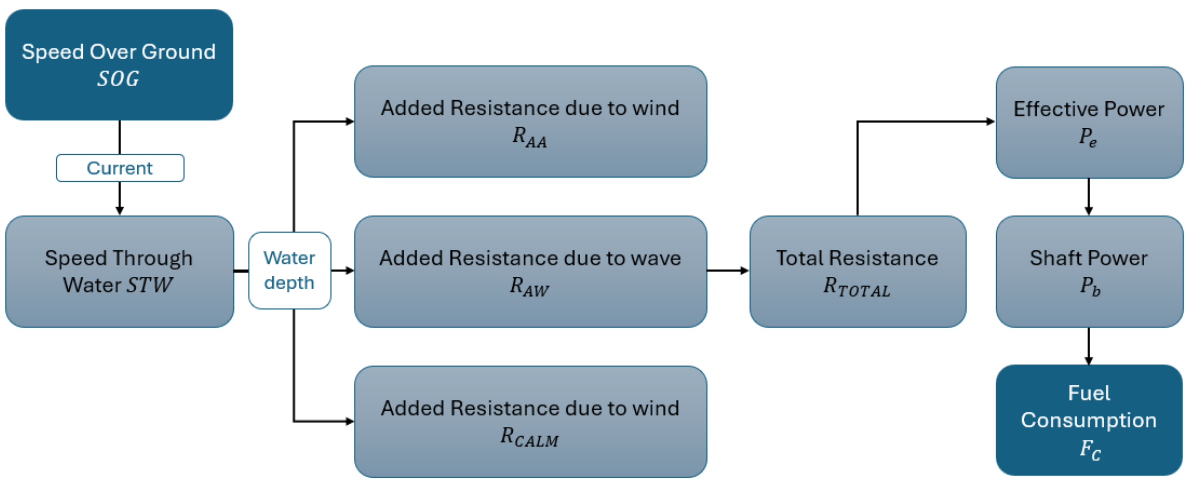


Figure 3 Typical workflow for the conventional estimation of a ship's speed to power/fuel consumption

The estimation of a ship's speed-power relationship begins with determining its resistance across various sailing speeds. This can be achieved through model testing, numerical simulations, or semi-empirical formulas. When operating in real sea conditions, external environmental factors such as wind and wave-induced resistance significantly impact the vessel's performance over its voyage. Therefore, added resistance due to wind (R_{AA}) and waves (R_{AW}) must be considered to accurately compute the total resistance (R_{TOTAL}) acting on the ship.

To overcome this resistance and maintain forward motion, the ship relies on thrust force generated by its propeller, which is powered by marine engines operating at specific RPMs under varying load conditions. The propulsion power required for the ship to sustain a given speed through water (V) is defined as the effective power (P_e):

$$P_e = R_{TOTAL} * V$$

This effective power is provided by the brake power (P_b) of the main engine, which is responsible for delivering the shaft power necessary to drive the propeller. The relationship between these parameters is given by:

$$P_b = \frac{P_e}{\eta_s \cdot \eta_h \cdot \eta_r \cdot \eta_o}$$

where η_s, η_h, η_r, η_o represent the shaft efficiency, hull efficiency, relative rotative efficiency, and open-water efficiency, respectively.

Finally, fuel consumption is estimated by multiplying the engine's brake power (P_b) with the specific fuel oil consumption (SFOC) and the operational time. The overall propulsion efficiency, which depends on engine type and propeller characteristics, is typically provided by manufacturers or shipowners and serves as a key factor in determining the vessel's total fuel usage under different operational and environmental conditions.

3.1.1. Calm water resistance

Holtrop and Mennen [25] proposed an approximate method for calculating a ship's calm water resistance based on full-scale trials and model experiments. This method considers key ship characteristics, including main dimensions, hull type, appendage configuration, and immersed transom sterns. The total resistance in still water is decomposed into six primary components:

$$R_{CALM} = R_F\,(1 + k_1) + R_{APP} + R_W + R_B + R_{TR} + R_A$$

where:

- R_F represents frictional resistance, estimated using the ITTC-1957 correlation line [26].
- $1 + k_1$ is the form factor, accounting for viscous pressure effects.
- R_{APP} denotes the resistance of appendages.
- R_W is the wave-making resistance of the bare hull.
- R_W corresponds to the wave resistance due to the bulbous bow.
- R_{TR} represents the additional resistance from immersed transoms.
- R_A accounts for the model-ship correlation resistance to correct for scale effects.

When towing tank resistance test data is available, it is generally preferred to interpolate the measured values rather than relying solely on empirical formulas. This approach helps minimize potential deviations caused by variations in hull form and ship type, ensuring a more accurate estimation of resistance.

3.1.2. Added resistance due to wind

The additional resistance induced by wind is primarily influenced by the ship's superstructure area and the relative wind conditions. The relative wind is calculated as the vector sum of the ship's speed and direction with the wind speed and direction. This added resistance can have a significant impact on the ship's propulsion power requirements, especially in adverse weather conditions. According to the ISO [27] guideline is used to estimate wind-induced resistance:

$$R_{AA} = \frac{1}{2} \cdot \rho_A \cdot [C_{AA}(\theta_{WR}) A_{XV} V_{WR}^2 - C_{AA}(0) A_{XV} V_{OG}^2]$$

where:

- ρ_A, air mass density
- A_{XV}, transverse projected area of the ship above the waterline, including the superstructure
- V_{WR}, relative wind speed
- θ_{WR}, relative wind direction
- V_{OG}, ship speed over ground
- $C_{AA}(\vartheta_{WR})$, wind resistance coefficient at a given relative wind angle
- $C_{AA}(0)$, wind resistance coefficient for headwind conditions

The wind resistance coefficients are determined based on extensive wind tunnel experiments and model tests, providing empirical values for different ship types and wind angles. These coefficients play a crucial role in accurately predicting the impact of wind on ship performance, particularly for voyage planning and fuel consumption estimation.

3.1.3. Added resistance due to wave

The resistance of a ship experiences due to waves is influenced by the wave spectrum and spreading function, which define the distribution of wave energy in different directions. To model irregular wave conditions encountered in real sea states, this study applies the JONSWAP wave spectrum [28] along with a Cosine-Squared spreading function:

$$D(\theta) = \begin{cases} \frac{2}{\pi} cos^2(\theta), & -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \\ 0, & otherwise \end{cases}$$

The wave spectrum is defined as:

$$S(\omega|H_s,T_p,\gamma)D(\theta) = \frac{320\,H_s^2}{T_p^4\,\omega^5}\exp\left(\frac{-1950}{T_p^4\,\omega^4}\right)\gamma^{exp\left[\frac{-(\omega-\omega_p)^2}{2\sigma^2\omega_p^2}\right]}D(\theta)$$

where:

- H_s, significant wave height
- T_p, wave peak period
- γ, peakedness factor (typically set to 3.3 for JONSWAP spectrum)
- σ, spectral width parameter, with values 0.07 for $\omega \le \omega_p$ and 0.09 for $\omega > \omega_p$

The added wave resistance (R_{AW}) in irregular waves is typically estimated by integrating the resistance due to regular waves ($R_{AW}(\omega)$), weighted by the wave spectrum $S(\omega)$, across the full range of wave frequencies:

$$R_{AW}(\omega|H_s,T_p,\gamma,V,\beta) = 2\int_0^{\infty}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} S(\omega|H_s,T_p,\gamma)\ \frac{R_{\alpha w}(\omega|V,\beta)}{\zeta_\alpha(\omega)^2}\ D(\theta-\beta)d\theta d\omega$$

where:

- $\zeta_\alpha(\omega)$, amplitude of the regular wave
- $R_{\alpha w}/\zeta_\alpha^2$ transfer function (RAOs) describing the ship's added resistance response to waves

This formulation provides a realistic estimate of wave-induced resistance, crucial for accurately modeling ship performance in various sea states.

3.1.4. Correction Factor for Ship Resistance and Power

The conventional integration methods developed by ITTC [29] and ISO [27] have played a crucial role in the implementation of the Energy Efficiency Design Index (EEDI) for ship design. These methods incorporate the effects of waves based on extensive benchmark studies using experimental test data and sea trials. While semi-empirical approaches provide reasonable estimates of the average wave resistance for large ship datasets, they are primarily designed for ship design applications rather than real-time operational performance assessment.

For operational applications such as voyage optimization, the focus shifts to predicting the specific ship's actual resistance and power demand rather than relying on generalized mean wave resistance estimates. One key limitation of traditional integration methods is their assumption of linear wave superposition, which does not fully capture the nonlinear nature of ship responses such as wave reflections, ship motions, and dynamic propulsion efficiency variations. These nonlinear effects become particularly significant in harsher sea conditions, where larger waves (H_S) lead to:

- Increased added resistance due to waves (R_{AW})
- Reduced propulsion efficiency due to excessive ship motions
- Surf-riding and instability, further increasing power requirements

To account for these nonlinear effects, a correction factor (C_{H_S}) is introduced, adjusting the semi-empirical wave resistance component (R_{AW}) to better reflect real-world ship performance in varying sea states. The total resistance is thus modified as follows:

$$R_{TOTAL} = R_{CALM} + R_{AA} + R_{AW} * C_{H_S}$$

where the correction factor is wave-height dependent and defined as:

$$C_{H_S} = \sqrt[3.5]{H_s}$$

For broader applicability, further experimental validation and full-scale measurements are necessary to refine and generalize this correction factor, ensuring its effectiveness across various ship types and operational conditions.

3.2. Machine Learning Techniques for Ship Speed Prediction

In recent years, artificial intelligence (AI) has emerged as a rapidly advancing field, revolutionizing various industries and solving complex problems. AI has been transformative in applications ranging from medical diagnostics and autonomous vehicles to speech recognition and personalized recommendations on e-commerce platforms and streaming services.

At its core, AI refers to the automation of tasks traditionally performed by humans, enabling machines to learn from data, recognize patterns, and make decisions. While AI's popularity has surged in recent years, its fundamental concepts trace back to the 1950s. Early AI techniques, such as rule-based systems and Support Vector Machines (SVMs), were often constrained by limited computational power. However, significant advancements in hardware (e.g., high-performance GPUs, faster storage solutions), data availability (large-scale datasets), and optimization algorithms have enabled researchers to develop more sophisticated machine learning models. These breakthroughs have led to remarkable improvements in predictive accuracy, making AI an indispensable tool across a wide range of scientific and industrial domains [30].

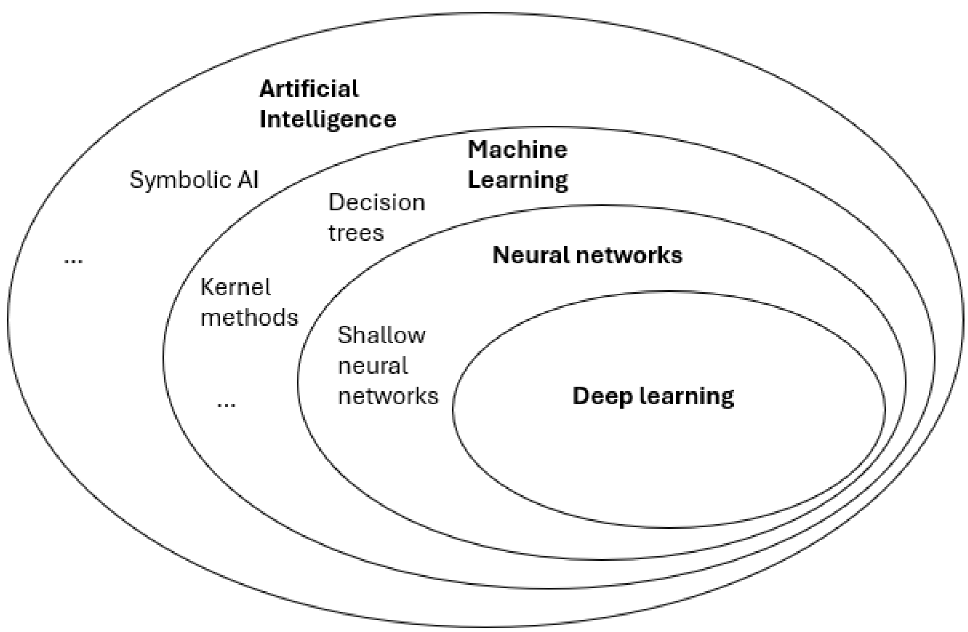


Figure 4 Artificial intelligence, Machine Learning, and Deep Learning [30]

While AI encompasses various approaches, machine learning (ML) is a specific subset focused on enabling systems to learn patterns from data without explicit programming. Unlike traditional rule-based programming, where humans define complex rules for a task (e.g., pattern recognition), machine learning models are trained on input-output pairs and automatically extract the underlying patterns and relationships in the data.

The machine learning process consists of two primary phases:

1. **Training Phase:** The model is initially provided with a dataset containing known input-output relationships. During training, the model's parameters (weights) are iteratively adjusted to minimize the difference between its predictions and actual values using a feedback mechanism.

2. **Inference Phase:** Once trained, the model is deployed to make predictions on new, unseen data by applying the learned relationships.

3.2.1. Taxonomy of data-driven algorithms

Data-driven models can be broadly categorized into three main learning paradigms, depending on how they learn from data:

1. Supervised Learning

2. Unsupervised Learning

3. Reinforcement Learning

This classification is based on the availability of labeled data and the learning approach used by each method.

- Supervised Learning requires that each observation in the training dataset is associated with a known target value (ground truth). The model learns by mapping input features to the correct outputs, making it well-suited for predictive modeling tasks, such as ship speed estimation based on historical operational data.

- Unsupervised Learning does not rely on labeled data. Instead, it identifies hidden structures, patterns, or clusters within the dataset. This approach is useful for anomaly detection, pattern recognition, and feature extraction, where the model explores relationships between variables without predefined categories.

- Reinforcement Learning (RL) differs from both supervised and unsupervised learning in that it does not use labeled data. Instead, the model learns through interaction with an environment, where it receives rewards or penalties based on its decisions. The goal of RL is to determine the optimal strategy (policy) that maximizes long-term rewards, making it applicable in scenarios requiring dynamic decision-making, such as autonomous voyage optimization.

Additionally, a hybrid approach known as semi-supervised learning combines elements of both supervised and unsupervised learning by using a mix of labeled and unlabeled data for model training. This technique is particularly useful in situations where acquiring labeled data is expensive or time-consuming.

While reinforcement learning has shown promise in various applications, its practical use in ship speed prediction and performance monitoring remains limited. Therefore, for the scope of this discussion, the primary focus will be on supervised and unsupervised learning methods, which have more direct applications in maritime data analysis.

3.2.2. Artificial Neural Networks (ANNs) and Deep Learning

Neural networks are a class of machine learning algorithms designed to mimic the structure and functionality of the human brain [31]. They consist of interconnected units called neurons, which are arranged in multiple layers:

1. Input Layer – Receives raw data and passes it to the next layer.
2. Hidden Layers – Process the information through weighted connections, applying transformations to extract relevant features.
3. Output Layer – Produces the final prediction or classification based on the learned representations, as shown in Figure 5.

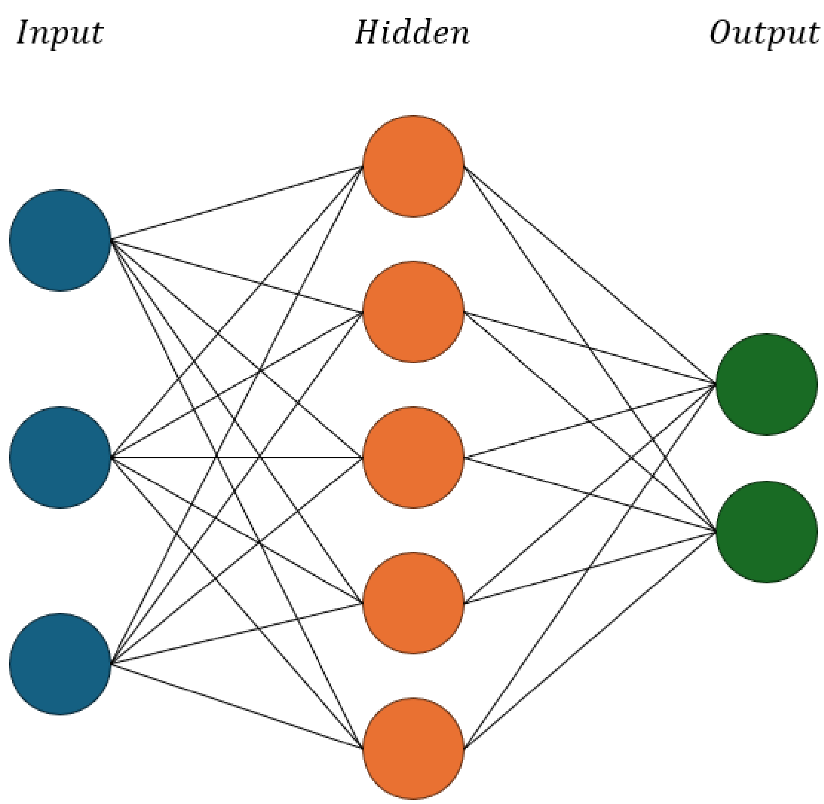


Figure 5 Representation of neural network

Each connection between neurons is associated with weights and biases, which determine the strength of influence an input has on the output. During training, these weights are adjusted through backpropagation and gradient descent, minimizing the difference between predicted and actual values [32]. To capture complex non-linear relationships, neural networks utilize activation functions, which introduce non-linearity into the learning process. This enables the network to model intricate dependencies between input features and improve predictive performance.

During training, a neural network is exposed to labeled data and iteratively adjusts its weights and biases using backpropagation and gradient descent to minimize prediction errors [32]. This learning process enables the network to model complex, non-linear relationships between input features and outputs, making it highly effective for a wide range of tasks, including image recognition, natural language processing, and strategic decision-making in games [31].

Despite their advantages, neural networks also have notable limitations. They can be computationally demanding, requiring high processing power and large datasets for effective training [33]. Additionally, they are susceptible to overfitting, where the model becomes too specialized to the training data and fails to generalize to new, unseen inputs. To mitigate overfitting, techniques such as regularization, dropout, and early stopping are commonly used to enhance model robustness and generalization [32].

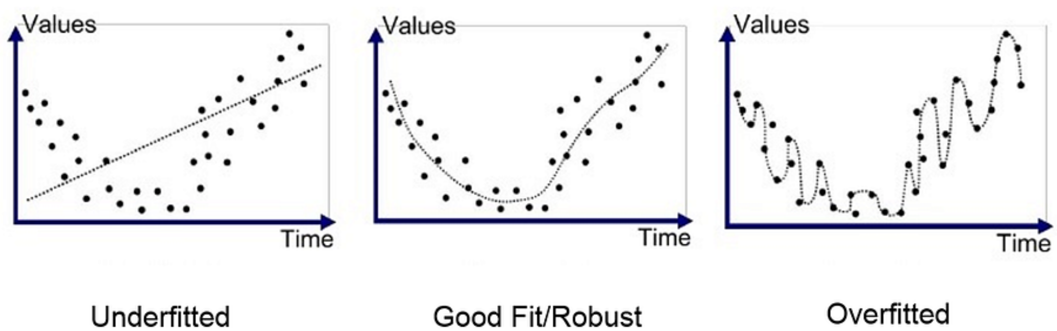


Figure 6 Model Fitting: Overfitting, Underfitting, and Balanced

In summary, neural networks are a powerful tool for solving complex problems, capable of learning intricate patterns and modeling non-linear relationships between inputs and outputs. Their effectiveness across various applications, from predictive modeling to autonomous decision-making, has solidified their role as a cornerstone of modern machine learning.

Despite certain challenges, such as computational demands and overfitting, ongoing research continues to drive significant advancements in neural network architectures, optimization techniques, and training methodologies. These improvements are enhancing model performance, efficiency, and generalization, making neural networks an increasingly robust and versatile solution for a wide range of real-world applications.

Hyperparameter Optimization

Hyperparameter optimization is a critical step in training machine learning models, as it directly impacts their accuracy, generalization, and overall performance. Selecting the appropriate hyperparameter values helps achieve a balance between underfitting and overfitting, ensuring that the model performs well not only on training data but also on unseen test data. Fine-tuning these parameters is essential for maximizing model efficiency and predictive accuracy.

Number of Hidden Layers

The number of hidden layers is a key hyperparameter in neural networks, particularly in Multi-Layer Perceptrons (MLPs) and Deep Neural Networks (DNNs). For many tasks, a single hidden layer is sufficient, provided it has enough neurons to capture the underlying relationships in the data. However, for more complex problems, deep networks (i.e., those with multiple hidden layers) tend to perform better while using fewer parameters, as they can learn hierarchical representations more efficiently.

Deep networks exploit the hierarchical structure of real-world data:

- Lower layers capture low-level features.
- Intermediate layers learn higher-level representations.
- Upper layers extract the most abstract and meaningful features before reaching the final output.

This hierarchical design improves generalization and enables faster convergence to an optimal solution. The common approach for determining the optimal number of hidden layers involves incrementally adding layers until the error stabilizes, ensuring that the model remains efficient while maintaining high performance on unseen data [34].

Number of Neurons per Hidden Layer

The number of neurons per hidden layer plays a crucial role in determining a neural network's ability to learn and generalize effectively. Earlier approaches often followed a pyramidal structure, where successive layers had fewer neurons. However, this practice has largely been abandoned, as using the same number of neurons across all hidden layers has been found to work just as well—if not better—while simplifying hyperparameter tuning.

In some cases, having a larger first hidden layer can enhance performance, but this depends on the dataset. If a hidden layer has too few neurons, the model may lose important information, limiting its ability to capture complex relationships in the data. Conversely, too many neurons can lead to overfitting and increased computational cost.

A common approach for selecting the optimal number of neurons is incrementally adding units until the model's error stabilizes, ensuring the smallest number of neurons that achieves optimal performance [34].

Activation Function

An activation function determines how the weighted sum of inputs from a neuron is transformed into an output. It is a key component that affects both the learning capability and efficiency of a neural network. Activation functions can be linear or nonlinear, with hidden layers typically using nonlinear functions to enable the network to capture complex patterns in the data.

Common activation functions include:

- ReLU (Rectified Linear Unit) – A widely used function due to its simplicity and computational efficiency. It effectively avoids vanishing gradients but can suffer from the "dying ReLU" problem, where inactive neurons output zero for all inputs.
- Logistic (Sigmoid) – Used in some applications but prone to vanishing gradients, making deep networks difficult to train.
- Tanh (Hyperbolic Tangent) – Similar to the sigmoid function but centered around zero, allowing for stronger gradients in deep networks.

For output layers, the choice of activation function depends on the prediction task:

- Linear activation for regression problems.
- Softmax or Sigmoid for classification tasks.

The selection of an activation function is critical for ensuring model stability and convergence, with ReLU being the preferred choice for hidden layers due to its balance of performance and efficiency [34].

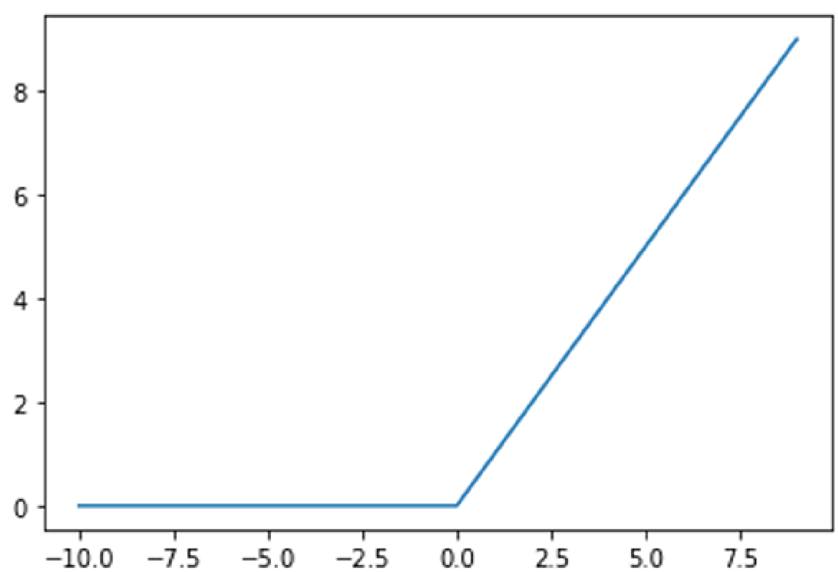


Figure 7 ReLU function

Number of Epochs & Batch Size

The number of epochs refers to the total number of times a machine learning model processes the entire training dataset. Each epoch allows the model to update its parameters based on the training data, gradually improving its predictive accuracy.

Batch size defines the number of samples processed before the model updates its parameters. Instead of updating the model after processing the entire dataset, training is done in mini batches, where updates occur after processing a subset of the data. The batch size must be a positive integer that is less than or equal to the total number of training samples.

- Larger batch sizes improve computational efficiency and stabilize gradient updates, especially in noisy datasets.
- Smaller batch sizes introduce more variability in gradient updates but can help the model generalize better.

While the number of epochs can be any positive integer, an excessively high number may lead to overfitting, where the model learns the training data too well and fails to generalize to new data. The optimal values for both epochs and batch size vary depending on the dataset and model architecture, and they are typically tuned experimentally using validation performance as a guide.

3.2.3. XGBoost: An Advanced Gradient Boosting Algorithm

XGBoost, developed by Chen and Guestrin [35], is an enhanced version of gradient boosting that offers higher computational efficiency and improved regularization techniques to mitigate overfitting. While machine learning methods such as Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) are widely applied in the maritime industry, XGBoost remains underutilized in ship performance modeling despite its robust handling of heterogeneous data and different feature scales.

Gradient Boosting and the XGBoost Framework

XGBoost is an advanced implementation of the Gradient Tree Boosting (GTB) algorithm. Gradient boosting is an ensemble learning method that combines multiple weak learners, typically decision trees, to develop a strong predictive model. Unlike bagging techniques, such as random forests, where trees are trained independently in parallel, boosting operates sequentially, with each new tree correcting the errors of its predecessors. The process begins by training an initial weak model, commonly a simple decision tree. Additional trees are then iteratively added, each designed to compensate for the limitations of the previous ones. The final model aggregates the predictions from all trees, producing a highly accurate outcome. This iterative refinement process makes gradient boosting one of the most effective techniques for classification and regression tasks.

XGBoost introduces several enhancements over traditional Gradient Boosting Decision Trees (GBDTs), primarily aimed at improving model accuracy, efficiency, and regularization. One of the key improvements is its ability to control overfitting through the integration of L1 regularization (LASSO) and L2 regularization (Ridge Regression), ensuring better generalization across datasets. Furthermore, XGBoost optimizes tree construction by utilizing Classification and Regression Tree (CART) algorithms, where binary splits are applied, and leaf nodes store prediction values (leaf weights) to enhance computational efficiency.

To further improve performance, XGBoost incorporates several computational optimizations, such as sparse-aware computations, weighted quantile sketching, and parallelization, making it significantly faster and more scalable compared to standard gradient boosting methods. These enhancements make XGBoost particularly effective for handling large datasets with high-dimensional features, making it a robust choice for predictive modeling tasks such as ship speed prediction and vessel performance analysis.

Gradient Boosting Decision Tree (GBDT) is a boosting algorithm designed specifically for regression tasks. The model-building process starts with a single decision tree, and additional trees are sequentially added in an effort to reduce residual errors from previous iterations. Each tree refines the prediction by learning from the mistakes of its predecessor, and the final model aggregates the weighted sum of all tree predictions.

XGBoost, an advanced implementation of GBDT, introduces several key improvements, particularly in handling regularization, computational efficiency, and predictive accuracy. Unlike conventional GBDTs, where each leaf node represents the average value of all samples assigned to that node, XGBoost assigns leaf weights, which serve as the regression values for each sample. These leaf weights, denoted as $f_k(x_i)$, represent the prediction score from the k-th decision tree for a given input x_i. When a single decision tree is used, the prediction is often inaccurate, but as more trees are added, the ensemble model accumulates the outputs from all trees, leading to more refined predictions. The final prediction for a sample is computed as:

$$\hat{y}_i^{(k)} = \sum_{k=1}^{K} f_k(x_i)$$

where K is the total number of trees in the ensemble.

3.2.4. Random Forest: A Robust Ensemble Learning Method

Random Forest (RF) is a widely used tree-based ensemble learning algorithm that builds upon the bagging (Bootstrap Aggregating) technique, introducing an additional layer of randomness in the training process. It is particularly effective for handling nonlinear relationships, reducing overfitting, and improving predictive performance in both classification and regression tasks [11].

Ensemble learning methods fall into two main categories: bagging and boosting. Bagging creates multiple variations of the training dataset through sampling with replacement and trains a weak learner on each subset independently. The final prediction is obtained by averaging results (for regression) or majority voting (for classification). Boosting, on the other hand, builds models sequentially, where each new model corrects errors from the previous one. Random Forest is an extension of bagging applied to Decision Trees (DTs). Unlike traditional Decision Trees (DTs), where each node is split based on all available features, Random Forests introduce feature randomness by selecting only a subset of features at each node split. This additional randomness enhances model diversity and robustness, making RF less sensitive to variations in the training data. If a dataset contains m features, the recommended number of features to consider at each node is typically $d = log_2 m$.

Random Forest models derive their randomness from two main sources:

1. Bootstrap Sampling – Each tree in the forest is trained on a random subset of the training data, drawn with replacement.
2. Feature Subsampling – Instead of considering all features when splitting a node, only a random subset is used, improving generalization.

This combination significantly reduces variance, leading to improved prediction accuracy and model stability, especially in datasets with high dimensionality or collinear features. By training multiple weak models on different data subsets, RF significantly reduces variance and prevents overfitting. Since training datasets may contain noise, outliers, or underrepresented samples, the RF approach ensures that the final model is less sensitive to anomalies, leading to improved robustness and generalization [36].

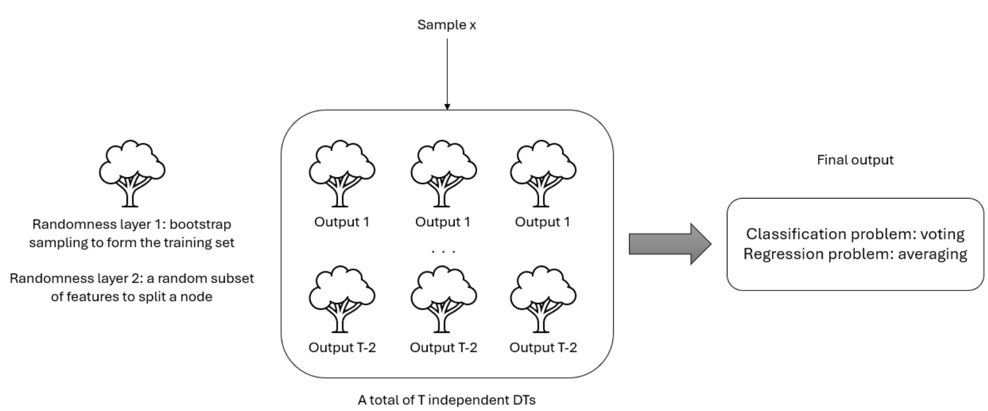


Figure 8 An illustration of an RF model

3.2.5. Error function

The error function is a crucial metric used to evaluate the performance of a machine learning model by comparing its predicted values to the actual values. The choice of the error function significantly impacts the reliability and interpretability of the model's predictions.

For regression problems, the most commonly used error functions include:

- **Mean Squared Error (MSE)**

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

MSE squares the differences between predicted and actual values, emphasizing larger errors more heavily.

- **Root Mean Squared Error (RMSE)**

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

RMSE provides a more interpretable error measure by converting squared differences back to the original unit of measurement but still penalizes large errors.

- **Mean Absolute Error (MAE)**

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$$

MAE measures absolute differences between actual and predicted values, making it more robust to outliers than MSE or RMSE.

- **Mean Absolute Percentage Error (MAPE)**

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

MAPE expresses errors as a percentage, making it scale-independent, but can be problematic when actual values are close to zero.

- **Coefficient of Determination (R^2 Score)**

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2}$$

R^2 measures how well the independent variables explain the variance in the dependent variable. It ranges from 0 to 1, where 1 indicates a perfect fit and values close to 0 suggest poor predictive power. Negative values indicate the model performs worse than a simple mean prediction.

Each error function has specific advantages depending on the use case. MSE and RMSE penalize larger errors more heavily, making them useful when minimizing large deviations is critical. MAE, being less sensitive to extreme values, is preferred when robustness to outliers is required. MAPE is commonly used in forecasting applications, where percentage-based errors provide a meaningful performance metric. R^2 Score is useful for assessing the model's overall explanatory power but should be used alongside absolute error metrics for a complete evaluation.

4. Case Study

4.1. Dataset Description

This chapter discusses the data sources for this project. The dataset consists of high-frequency sensor data collected every five (5) minutes from two sister Aframax crude oil tankers (Vessel_1 and Vessel_2). The data spans from June 2023 to May 2024, with 99,196 data points for Vessel_1 and 103,620 data points for Vessel_2. The vessels' main particulars are presented in Table 1.

Table 1 Main particulars for both vessels

Main Particulars		
Ship Type	Tanker	
Length Between Perpenticulars (L_{BP})	242	[m]
Breadth (moulded)	44	[m]
Depth (moulded)	21.2	[m]
Scantling Draught	15.2	[m]
Main Engine Type	MAN B&W 6G60ME - C10.5	
MCR Power	12690	[kW]
MCR RPM	85.8	[RPM]
Propeller Type	Fixed Pitch Propeller	
YOB	2022	

The dataset includes a variety of operational parameters as presented in Table 2.

Table 2 Recorded sensor data parameters

Label	Units - Format
Timestamp	YYYY-MM-DD HH:MM
Draught Mid	[m]
Speed Through Water (STW)	[kn]
Speed Over Ground (SOG)	[kn]
M/E RPM	[RPM]
M/E Power	[kW]
M/E Cylinder exhaust Gas Temperature	[°C]
M/E Consumption	[t/h]
Wind Speed	[kn]
Wind Direction	[deg]
Vessel's Course	[deg]

Additionally, the dataset is complemented by Noon Reports from both vessels for the same period, providing information on:

Table 3 Recorded noon report data parameters

Label	Units - Format
Timestamp	DD/MM/YYYY HH:MM:SS
Time zone	
Report Type	
Draught Fore	[m]
Draught Mid	[m]
Draugght Aft	[m]
Speed Through Water (STW)	[kn]
Speed Over Ground (SOG)	[kn]
M/E RPM	[RPM]
M/E Power	[kW]
Wind Speed	[kn]
Wind Direction	[deg]
Wave Height	[m]
Wave Direction	[deg]
Swell Height	[m]
Swell Direction	[deg]
Vessel’s Course	[deg]
Ambient Temperature	[°C]
E/R Temperature	[°C]

The Noon Reports serve as a validation source, enabling cross-checking of idle periods, speed variations, and power consumption trends.

To enhance the accuracy of predictions and to shape the final grey model, external data sources are included:

- **Sea Trials Data**: Provides performance benchmarks under controlled conditions.
- **M/E Shop Tests**: Contains manufacturer test results for power performance.
- **Model Tests**: Offers hydrodynamic testing data for ship behavior modeling.
- **Underwater Cleaning (UWC) and Propeller Polishing Records**: Includes maintenance logs affecting ship resistance and performance.

4.2. Exploratory Data Analysis & Feature Selection

This chapter outlines the data preprocessing methodology employed to ensure high-quality input for the machine learning model. First, an Exploratory Data Analysis (EDA) is conducted to examine the structure of the dataset, detect missing values, and identify potential inconsistencies. Then, statistical filtering techniques for outlier detection and removal are applied to eliminate erroneous data points. Finally, a feature selection process is performed, leveraging correlation analysis and domain knowledge to retain only the most relevant and independent variables for model training.

A visual representation of the multi-stage preprocessing pipeline and feature selection methodology is shown in Figure 9. The diagram illustrates the integration of various data sources, including sensor data, noon reports, and UWC logs, followed by the data processing steps applied before feature

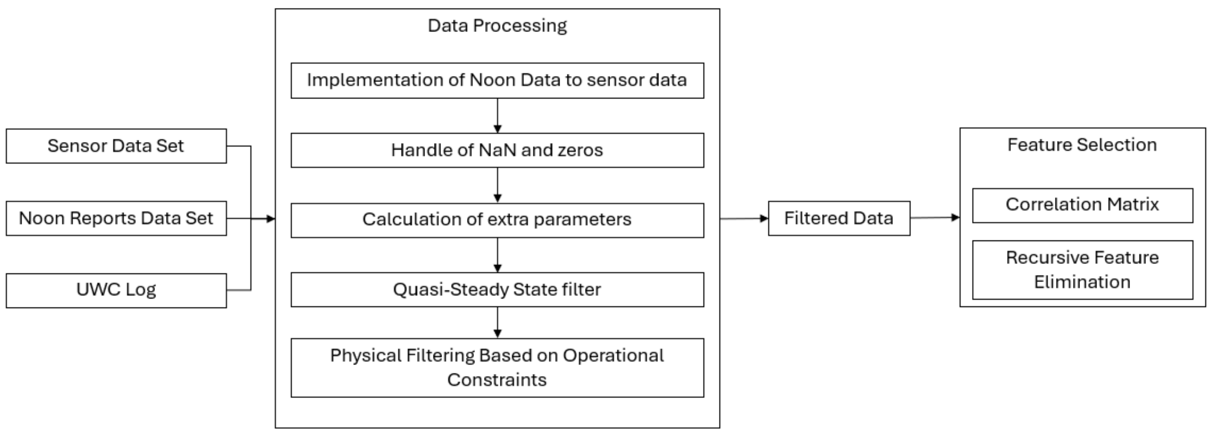


Figure 9 Visual representation of data preprocessing and feature selection

4.2.1. Outlier Detection and Removal

An outlier is an observation that lies an abnormal distance from other values in a random sample from a population. In a sense, this definition leaves it up to the analyst (or a consensus process) to decide what will be considered abnormal. Before abnormal observations can be singled out, it is necessary to characterize normal observations [37] A first step in this process is computing descriptive statistics, as shown in Table 4, which provides insight into the characteristics of each parameter from sensor data analyzed.

Table 4 Descriptive statics of the raw dataset for Vessel_1

	WIND_FORCE_KN	DRAFT_M	ME_RPM	M1_POW	WIND_DIR	STW
count	99050	98015	99129	99129	99055	99057
mean	11.26	9.73	35.62	3472.52	173.87	7.12
min	0.00	-2.21	-371.00	-3710.00	0.00	-4.12
25%	6.10	7.30	0.00	0.00	46.67	0.60
50%	9.45	11.02	0.50	1.00	157.50	5.24
75%	14.80	12.44	75.20	7499.00	309.09	13.65
max	46.06	14.67	81.10	10900.00	360.00	16.54
std	7.10	3.33	36.66	3769.75	126.52	6.37

Additionally, handling missing values is a critical step in data preprocessing, as they are among the most common issues encountered. Similarly, outliers should be addressed appropriately to prevent skewed analyses, and negative values that lack physical meaning must be identified and corrected where necessary.

Table 5 Total number of NaN instances for Vessel_1

	NaN Values
WIND_FORCE_KN	146
DRAFT_M	1181
ME_RPM	67
ME_POW	67
WIND_DIR	141
STW	139

Additionally, Figure 10 presents a time-series visualization of the dataset over its full duration, highlighting inconsistencies and missing values. Notably, a significant data gap is observed in February 2024, affecting most of the examined variables.

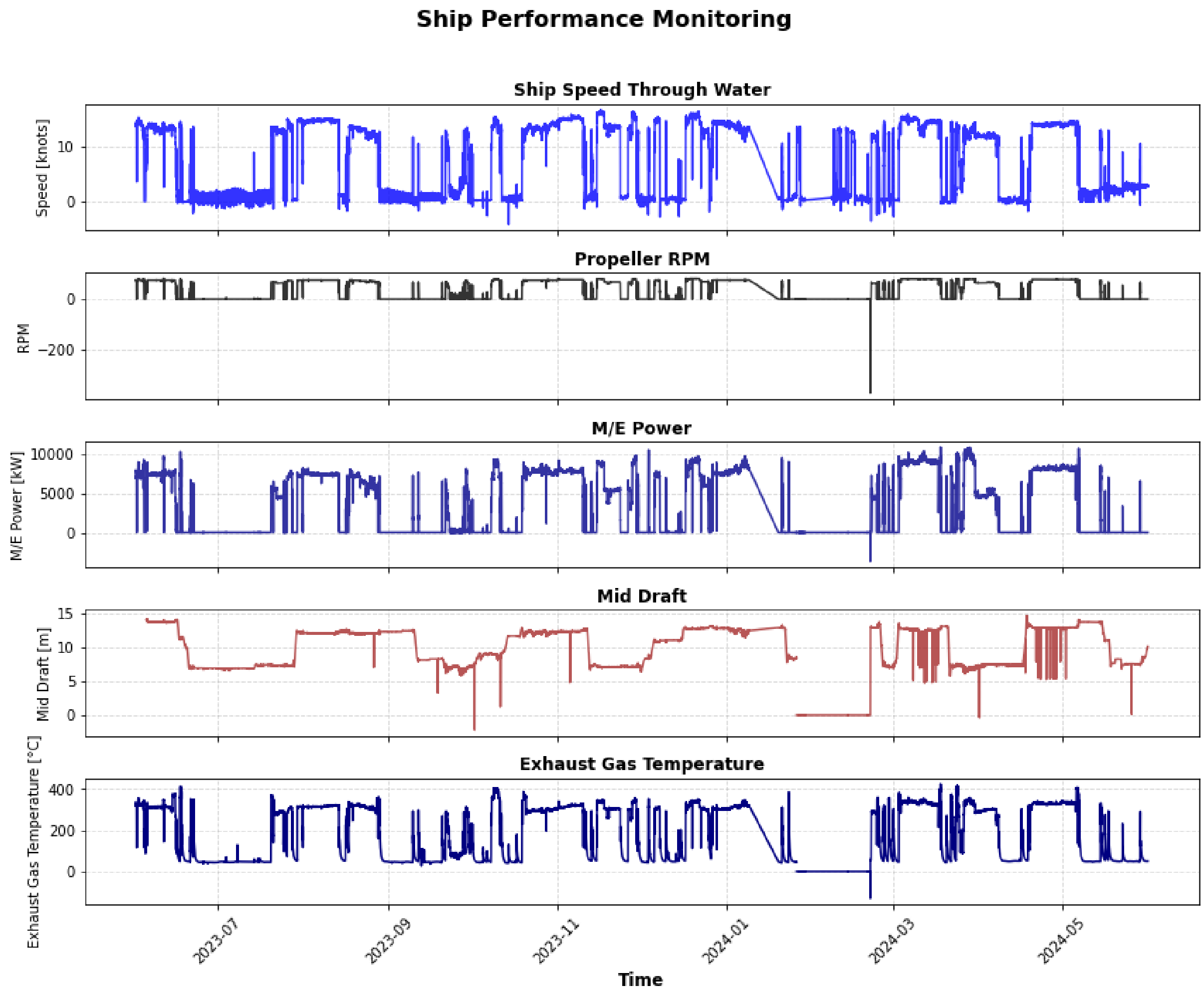


Figure 10 Raw data of key vessel parameters over time

To further explore feature relationships, scatter plots are used to visualize dependencies identified in the correlation matrix. In Figure 11, the first plot shows a strong linear relationship between Speed Over Ground (SOG) and Speed Through Water (STW), confirming expected physical behavior, due to the current effects. The second plot, depicting the relationship between STW and RPM, follows a trend consistent with propeller hydrodynamics. The spread and variation in the data indicates the influence of propeller slip, which varies depending on factors such as hull fouling, sea state, and propulsion efficiency. Slip increases when there is greater resistance acting against the vessel, causing the actual speed through water to be lower than the theoretical value derived from propeller RPM. Finally, the third plot demonstrates the cubic relationship between RPM and Main Engine Power, which is consistent with the Propeller Law. Although the dataset largely follows the expected relationships between the analyzed parameters, observed outliers should be removed to improve model reliability and accuracy.

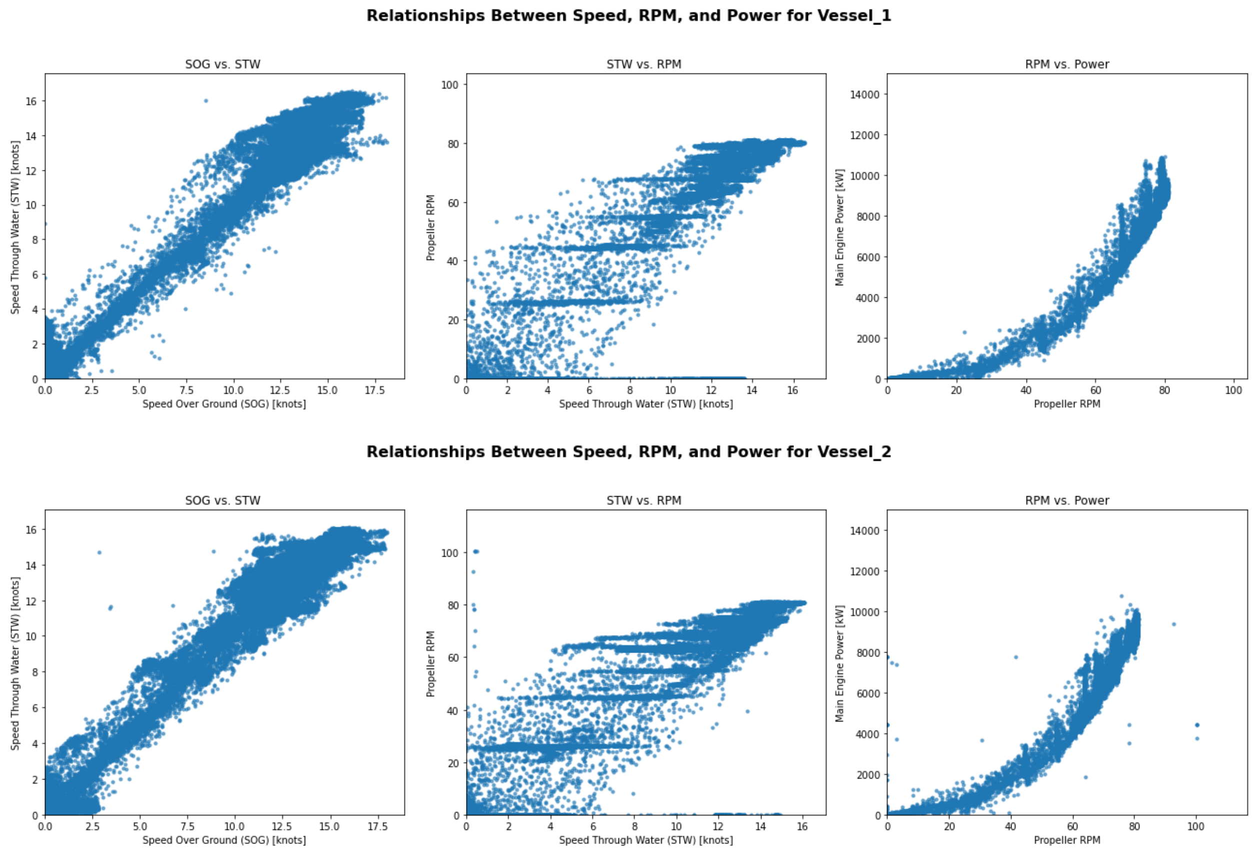


Figure 11 Scatter plots showing key vessel relationships—SOG vs. STW, STW vs. RPM, and RPM vs. Power

The previous analysis focuses on the sensor dataset; however, this dataset is complemented by the Noon Reports, which provide additional parameters not captured by the sensors, such as swell, wave height, and direction. Additionally, the Noon Reports are utilized to fill in data gaps when sensor measurements are unavailable.

For instance, if the mid-draft value is missing in the sensor dataset for specific records, it is replaced with the corresponding value from the Noon Reports for the same timestamp. To ensure accurate data alignment, it is crucial to understand the structure and reporting mechanism of Noon Reports. These reports are typically generated at noon each day, except in cases where a key event occurs (e.g., "Arrival"), in which case an additional report is created. Each Noon Report corresponds to the period since the previous report, meaning that a recorded value, such as M/E fuel consumption of 40 metric tons, represents the total consumption from the previous report to the current one.

Furthermore, Noon Reports are recorded in local time, reflecting the vessel's current time zone. To align them with sensor timestamps, which are typically in Coordinated Universal Time (UTC), the Noon Report timestamps are converted to UTC before merging with the sensor dataset. Subsequently, for each missing value in the sensor dataset, the corresponding Noon Report value within the same reporting period is used as a replacement. This methodology ensures that the sensor dataset is enriched by incorporating reliable Noon Report data.

After integrating the Noon Report data, additional data correction and filtering procedures were applied to ensure data quality before analysis. As previously mentioned, Mid Draft values from sensors were supplemented with Noon Report data in cases where the sensor readings were missing (NaN) or recorded as zero. Similarly, negative values in SOG were identified as physically implausible and were replaced with zero. The same correction was applied to negative RPM and M/E Power values, as such readings do not hold physical meaning in this context.

As part of the data preprocessing, additional parameters were computed to improve the accuracy of the vessel power prediction model. These derived parameters ensure a more comprehensive dataset for analysis.

Trim Calculation

Trim is a crucial factor in vessel hydrodynamics, affecting fuel efficiency and propulsion power. It is defined as the difference between the fore draft and the aft draft:

$$Trim = Draught\ Fore - Draught\ Aft$$

Wind Component in Vessel's Direction

The impact of wind on a vessel's performance depends on its relative direction to the vessel's movement. To quantify this effect, the wind component acting along the vessel's course was computed using the relative wind angle:

$$Wind\ Component = Wind\ Speed * \cos\left(Relative\ Angle\right)$$

Wave and Swell Components

To account for the impact of waves and swells on vessel resistance, two new features were

$$Wave\ Component = Wave\ Height * \cos\left(Wave\ Direction\right)$$

$$Swell\ Component = Swell\ Height * \cos\left(Swell\ Direction\right)$$

Fouling Factor Calculation

Biofouling on a vessel's hull increases hydrodynamic resistance, leading to greater fuel consumption over time. To quantify this effect, a Fouling Factor was introduced using an exponential decay function:

$$Fouling\ Factor = 1 - e^{-k*Days\ Since\ Last\ UWC}$$

where:

- k is a decay rate constant, empirically set to 0.005.
- Days Since Last UWC represents the number of days since the last hull and propeller cleaning.

This function ensures that:

- When recently cleaned (Days Since Cleaning = 0), the fouling factor is near zero.
- Over time, fouling gradually accumulates toward a maximum limit.

Current Effect Calculation

The impact of ocean currents was estimated by analyzing the difference between Speed Over Ground (SOG) and Speed Through Water (STW):

$$Current\ effect = Speed\ Over\ Ground - Speed\ Through\ Water$$

where:

- SOG represents the vessel's actual movement relative to the seabed.
- STW represents the movement relative to water.

If SOG > STW, the vessel benefits from a favorable current. If SOG < STW, the vessel faces an adverse current, increasing resistance.

Systematic Filtering Approach

Beyond basic corrections, a systematic approach to outlier detection was implemented to ensure data consistency and reliability. Outlier removal is particularly important in machine learning applications, where incorrect data points can distort model predictions. In ship performance monitoring, outliers may arise due to:

- Sensor malfunctions or calibration errors, leading to unrealistic readings.
- External environmental influences, such as sudden gusts of wind, strong currents, or extreme weather conditions.
- Transient operational states, where the vessel undergoes acceleration, deceleration, or abrupt course changes.

To address these issues, a combination of filtering methods was employed, including Quasi-Steady-State (QSS) filtering, physical constraints filtering, and statistical outlier removal based on industry standards (e.g., ISO 19030).

A widely used statistical approach for outlier detection, Chauvenet's Criterion, is explicitly referenced in ISO 19030-2:2016 [38], which recommends its use for filtering consecutive, non-overlapping 10-minute data blocks. According to this approach, an observation is considered an outlier if its probability of occurrence is below a defined threshold based on the complementary error function:

$$P(d_i) = erfc(\frac{\Delta_i}{\sigma * \sqrt{2}})$$

where,

- $P(d_i)$ is the probability of data point d_i occurring, $\delta_i = |d_i - \mu|$ represents the deviation from the mean, and
- σ is the standard deviation of the dataset.

A value is considered an outlier if the inequality:

$$P(d_i) * N < 0.5$$

is satisfied, where N is the total number of data points.

However, despite its theoretical validity, several practical limitations arise when applying Chauvenet's Criterion to vessel performance datasets [39]. The method assumes that data is normally distributed, yet real-world vessel data often deviates significantly from Gaussian assumptions. Ship speed, fuel consumption, and engine load are influenced by external factors such as sea state, wind, and hull fouling, leading to skewed or multi-modal distributions that Chauvenet's Criterion does not account for. Furthermore, Chauvenet's Criterion treats each data point independently, failing to consider time dependencies in vessel operation. Ship performance data is inherently time-series-based, where variations in speed, power, and resistance occur due to evolving operational conditions rather than sensor errors. Consequently, applying a probability-based outlier detection method without considering temporal dynamics may lead to the removal of valid performance data.

For these reasons, Chauvenet's Criterion was not implemented in this study. Instead, an alternative multi-stage filtering approach was adopted to identify and exclude unreliable data, ensuring that only physically meaningful outliers were removed. This filtering process includes:

1. Quasi-Steady-State (QSS) filtering [40], which eliminates transient operational states such as acceleration, deceleration, and course changes,

2. Physical filtering, applying thresholds based on vessel speed, wind force, wave height, and swell height, ensuring that extreme environmental conditions are excluded.

3. Statistical filtering, with error thresholds to align with real-world vessel behavior.

By implementing this methodology, outlier detection focuses on operationally significant inconsistencies, rather than removing data based solely on probabilistic assumptions. The approach preserves genuine variations in ship performance while ensuring that spurious anomalies do not distort modeling and analysis.

Quasi-Steady-State (QSS) Filtering

The QSS filter was applied to remove data corresponding to transient operational phases, where the vessel undergoes significant changes in speed, shaft RPM, or heading. These transient states are characterized by rapid variations in the dataset, which introduce non-stationary behavior and increase model uncertainty. Since modeling transient states explicitly would require a significantly more complex approach, they were eliminated before further analysis.

The QSS filter, adapted from Gupta et al. [40], operates in two stages. In the first stage, a sliding window regression is applied to detect periods where the rate of change (slope) in shaft RPM exceeds a defined threshold. A t-test is performed to determine whether a given data window exhibits a statistically significant change in state. To avoid misclassification due to near-zero variance in the data, the t-value t_1 is computed as:

$$t_1 = \frac{b_1}{1 + \sigma_b}$$

where b_1 is the estimated slope within the window, and σ_b is the standard deviation of the slope. This approach prevents cases where flat-line data with near-zero standard deviation would incorrectly be classified as transient.

In the second stage, a backward gradient check is applied to samples that failed the first test. This step helps retain some observations near sudden state changes that might otherwise be incorrectly excluded. The gradient is calculated as:

$$\frac{\partial x_i}{\partial t} = \frac{x_i - x_{i-1}}{t_i - t_{i-1}}$$

where x_i and x_{i-1} represent consecutive data points, and t_i , t_{i-1} are their respective timestamps. If the absolute gradient falls below a certain threshold, the sample is retained.

Figure 12 illustrates the impact of the QSS filtering process. The original dataset contains both transient and steady-state conditions, while the filtered dataset removes acceleration and deceleration effects, leaving only steady-state periods.

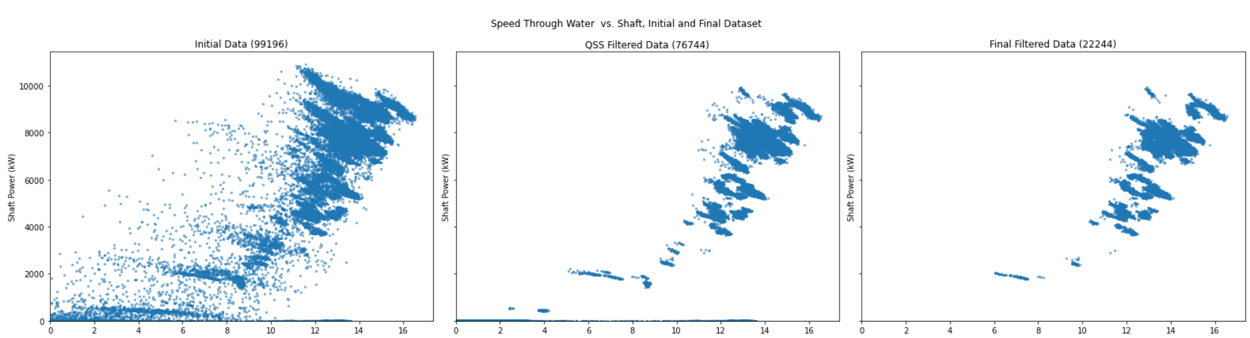


Figure 12 Data filtering process

Physical Filtering Based on Operational Constraints

After eliminating transient conditions, an additional filtering step was applied to exclude data recorded under extreme environmental conditions, which could significantly affect vessel performance. These conditions include high winds, rough seas, and low vessel speeds, where external forces dominate the vessel's response, making performance modeling unreliable.

The following constraints were imposed:

- Speed Through Water (STW) ≥ 6 knots to exclude periods of slow maneuvering or drifting.
- Wind Force ≤ 15.5 knots to remove cases where excessive wind resistance impacts performance.
- Wave Height ≤ 3 meters and Swell Height ≤ 3 meters to ensure that vessel behavior is not significantly altered by heavy seas.

4.2.2. Feature Selection

To improve model efficiency and accuracy, a feature selection process was applied to remove redundant or irrelevant variables. The selection was conducted in two stages: correlation analysis and Recursive Feature Elimination (RFE).

Initially, a correlation matrix (Figure 13) was computed to identify dependencies between the recorded parameters.

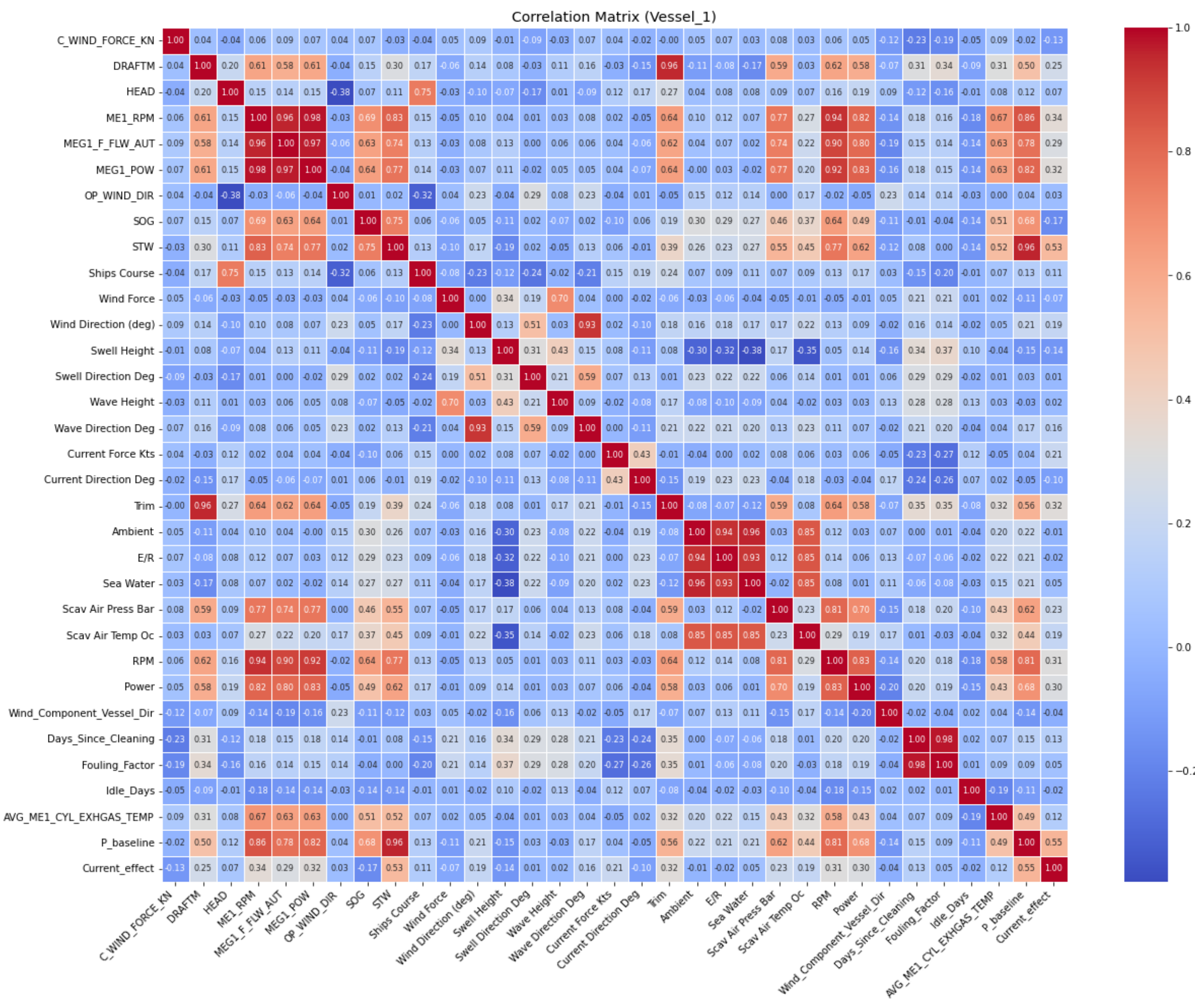


Figure 13 Correlation matrix of key vessel parameters

The analysis revealed strong relationships between certain variables, indicating potential redundancy. Specifically, a near-perfect correlation was observed between Main Engine RPM and Main Engine Power ($\rho \approx 0.98$), suggesting that both variables convey similar information, and one may be omitted without a loss of predictive capability. A similarly high correlation was found between Trim and Mid Draft, implying that both parameters represent hull immersion effects and may not be necessary together in the final feature set. Additionally, environmental parameters such as Wave Height and Swell Height exhibited moderate correlation, indicating that their combined effect should be assessed further. Following this initial analysis, the dataset was refined to retain only the most informative variables, as shown in Figure 14.

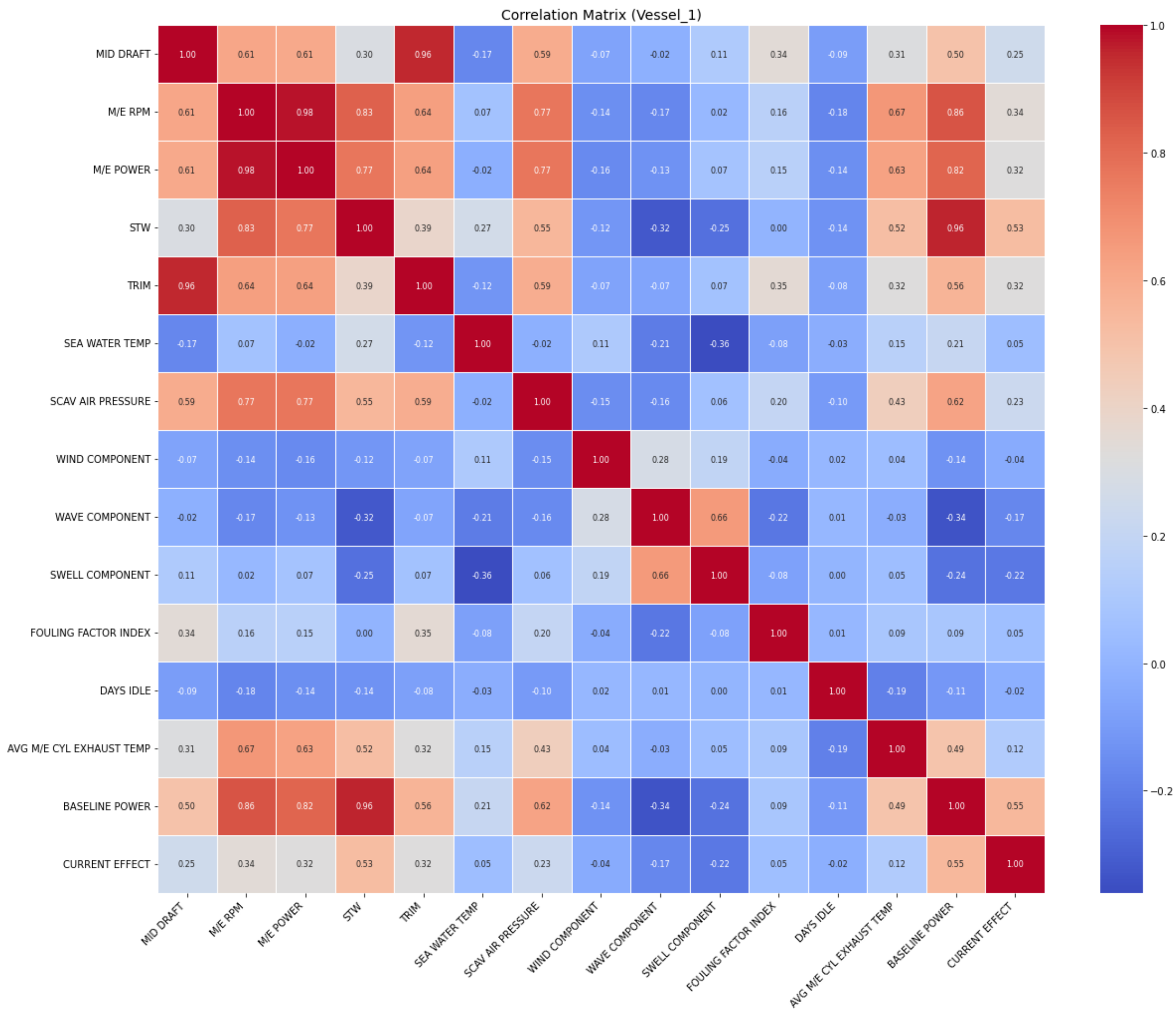


Figure 14 Remaining features after the examination of the initial correlation matrix

To refine feature selection, Recursive Feature Elimination (RFE) was applied using a Random Forest Regressor to determine the most predictive variables for modeling Main Engine Power (M/E Power). This method iteratively removes less significant features while retaining those with the highest contribution to prediction accuracy. The process ranked all available features based on their importance to the model and retained the ten most significant variables. The implementation was carried out using the sklearn.feature_selection package, and the final selected features are presented in Figure 15.

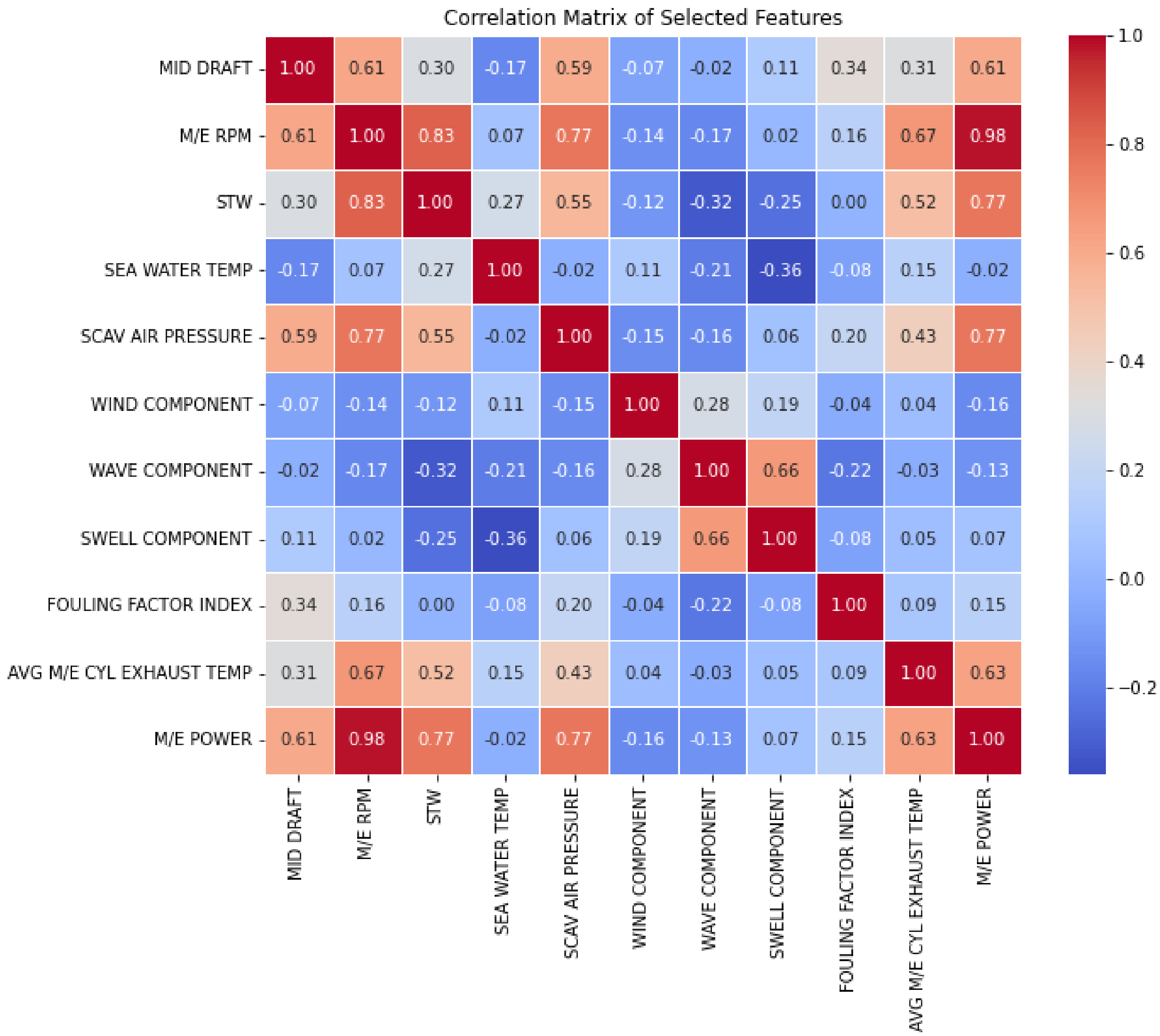


Figure 15 Final feature selection

After applying RFE, the final dataset included Main Engine RPM, Speed Through Water (STW), Trim, Mid Draft, Sea Water Temperature, Wind Component, Wave Component, Swell Component, Current Effect, and Fouling Factor Index. The correlation matrix for the reduced feature set (Figure 15) illustrates the improvement in feature independence, demonstrating that highly correlated parameters were successfully removed. This selection ensures that the dataset remains representative of vessel performance while minimizing redundancy and improving model interpretability.

5. Prediction Model

Accurate prediction of ship speed is a crucial component of maritime performance modeling, as it directly impacts fuel efficiency, voyage planning, and regulatory compliance. A ship's speed over ground (SOG) is influenced by multiple factors, including propulsion power (P_D), hull resistance, and environmental conditions such as wind, waves, and currents. In ideal calm water conditions, the ship achieves speed through water (STW) purely based on propulsion power (P_D) and draft (D). However, in real-world conditions, added resistance from wind and waves reduces the ship's effective speed through water. Ocean currents further affect the final speed over ground.

Predicting ship speed accurately is a complex task that requires balancing theoretical modeling with real-world data observations. Traditionally, ship performance models fall into three categories: white-box models (WBMs), black-box models (BBMs), and grey-box models (GBMs). White-box models are based on first principles or semi-empirical formulations derived from physics. These models rely on well-established hydrodynamic relationships and empirical correlations but are limited by the assumptions and simplifications introduced during the modeling process. Their accuracy depends on how well the resistance, propulsion, and environmental forces are represented, but they often fail to capture real-time operational deviations caused by unpredictable external factors.

On the other end of the spectrum, black-box models are purely data-driven, relying on machine learning and statistical regression techniques trained on experimental or full-scale sailing data. These models do not require prior theoretical knowledge, as they infer patterns directly from large datasets. However, black-box models suffer from poor interpretability and limited extrapolation ability, making them unreliable in unseen scenarios. If the training dataset does not include extreme or rare operating conditions, the model may produce highly inaccurate predictions when extrapolated beyond its data range.

To bridge the gap between WBMs and BBMs, grey-box models (GBMs) have been proposed as a hybrid approach that integrates both physical principles and data-driven corrections. GBMs combine the interpretability and theoretical foundation of white-box models with the adaptability and accuracy of black-box models. Unlike purely data-driven approaches, grey-box models require significantly less training data while still providing superior accuracy compared to physics-only models. They also exhibit better extrapolation capability, ensuring that predictions remain physically meaningful even in conditions where data is scarce.

In this study, two complementary modeling approaches are developed to predict ship speed requirements: a physics-based white-box model, which derives power-speed relationships from first principles and sea trials, and a data-enhanced grey-box model, which refines the physics-based estimates using machine learning algorithms trained on historical ship performance data. This chapter details the development, performance, and comparative analysis of both models, demonstrating how a hybrid approach can optimize predictive accuracy while maintaining physical consistency.

5.1. Physics Model

The white-box model (WBM) is based on first principles and empirical formulations derived from ship hydrodynamics. It provides a baseline estimation of the ship speed required for a given operating condition, assuming calm water conditions and no external disturbances. The accuracy of the WBM depends on the assumptions and simplifications made during the modeling process. While WBMs provide high interpretability and generalization ability, they often fail to account for real-world deviations caused by environmental influences, hull fouling, and variations in propulsion efficiency.

In this study, the baseline speed is estimated using a physics-based approach that incorporates:

1. A semi-empirical resistance model, which calculates calm water resistance and derives the corresponding power requirement.
2. Sea trial-based interpolation, which uses experimental speed-power curves obtained under controlled conditions.
3. Physics-informed neural networks (PINNs), which integrate hydrodynamic principles into a data-driven model.

These approaches allow the white-box model to provide an initial speed estimation, which is later refined using a black-box machine learning model in the grey-box framework.

5.1.1. Semi-Empirical Resistance Model for Speed Prediction

The semi-empirical approach estimates calm water resistance (R_{calm}) and converts it into required propulsion power using total propulsion efficiency (η_D). The resistance components are computed using the Holtrop and Mennen method, which decomposes total resistance into multiple components:

$$R_{calm}(STW, D) = R_F\,(1 + k_1) + R_{APP} + R_W + R_B + R_{TR} + R_A$$

where:

- R_F is the frictional resistance, estimated from the ITTC-1957 correlation line.
- k_1 is the form factor, which accounts for viscous effects.
- R_{APP} represents appendage resistance, calculated empirically.
- R_W is the wave-making resistance due to the hull shape.
- R_B represents the resistance contribution from a bulbous bow.
- R_{TR} is the additional transom resistance caused by immersion.
- R_A is the model-ship correlation resistance, used for full-scale corrections.

Once the total resistance is computed, the required propulsion power in calm water is estimated using:

$$P_D = \frac{STW * R_{calm}(V, D)}{\eta_D}$$

where:

- STW is the ship's speed through water,
- D is the vessel's draught,
- η_D is the propulsion efficiency, provided by ship operators.

This equation assumes that no external disturbances such as waves, wind, or currents affect the vessel's movement. As a result, it provides an idealized estimate of power consumption, which requires further refinement for real-world conditions.

5.1.2. Sea Trial Interpolation for Speed Estimation

Sea trials are conducted under controlled conditions to measure a vessel's performance across different operating points. These trials typically include speed-power tests at various drafts, allowing ship operators to establish empirical relationships between propulsion power, speed, and loading conditions. The measured data is corrected for environmental factors such as wind and currents to ensure accuracy and reliability.

Traditionally, sea trial data has been used for interpolation, where polynomial regression or lookup tables provide estimates of propulsion power for intermediate speeds and drafts. However, these methods have limitations when applied to real-world conditions, as they do not capture nonlinear interactions influenced by hull fouling, weather effects, or operational variations.

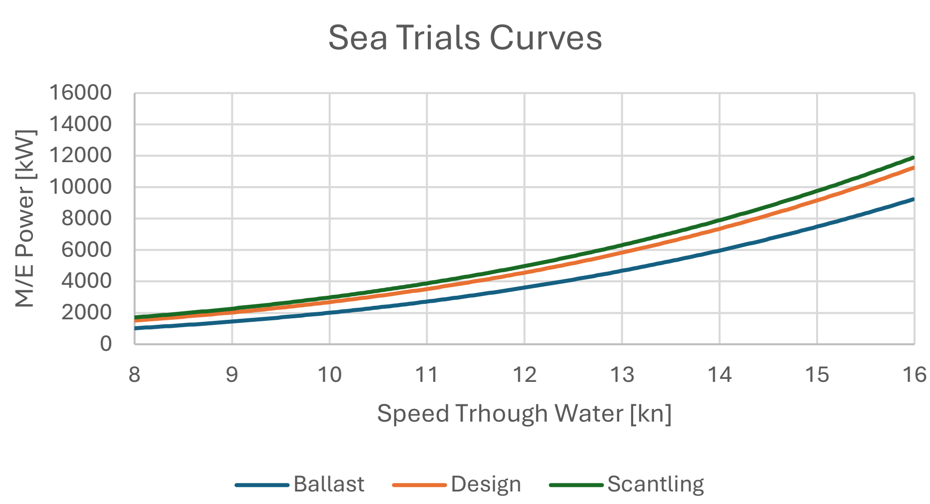


Figure 16 Sea Trials Curves

A Physics-Informed Neural Network (PINN) is employed in this study as an alternative to explicit polynomial regression for learning the relationship between propulsion power, speed, and draft [21]. Unlike polynomial regression, which relies on predefined mathematical equations to fit observed data, the PINN incorporates fundamental ship physics to guide the learning process, ensuring hydrodynamically meaningful predictions.

By integrating sea trial data into the PINN framework, the model is trained to learn from real-world measurements while being constrained by established hydrodynamic relationships. This approach enhances the robustness and generalization capability of the model, allowing it to provide accurate power predictions beyond the range of tested conditions, making it more adaptable to operational variability.

5.1.3. Physics-Informed Neural Networks (PINNs) for Speed Prediction

A more advanced approach integrates physical principles into a neural network framework. Physics-informed neural networks (PINNs) enforce known hydrodynamic relationships while learning from data, providing a balance between theory-based modeling and data-driven adjustments.

Instead of purely relying on empirical regressions, PINNs attempt to solve a partial differential equation (PDE) that governs the relationship between propulsion power (P), speed (V), and draft (D):

$$\alpha_1 \frac{\partial P}{\partial V} + \alpha_2 \frac{\partial P}{\partial T} + \mu[P, V, D; \lambda] = 0$$

where $\mu[P, V, D; \lambda]$ is a nonlinear function representing propeller-hull interactions, efficiency losses, and resistance effects.

The PINN model is implemented using TensorFlow, which enables automatic differentiation and optimization. The neural network consists of three hidden layers, each containing 64 neurons, with the hyperbolic tangent (tanh) activation function applied to introduce nonlinearity while preserving smooth gradients. The output layer consists of a single neuron representing propulsion power, utilizing a linear activation function to ensure continuous regression output.

The loss function is designed to balance data accuracy and physics consistency. Mean squared error (MSE) is used to minimize the difference between predicted and actual power values, while a physics-informed regularization term enforces the PDE constraint by penalizing deviations from the known hydrodynamic relationships. The total MSE loss combines two components:

$$MSE = MSE_{bc} + MSE_g$$

where:

- MSE_{bc} represents the boundary conditions, the loss on the boundary conditions can be expressed as:

$$MSE_{bc} = \frac{1}{N_{bc}} \sum_{i=1}^{N} |P_D\left(V_{bc(i)}, D_{bc(i)}\right) - P_{Dbc(i)}|^2$$

- MSE_g enforces physics-based constraints, penalizing deviations from known hydrodynamic relationships:

$$MSE_g = \frac{1}{N} \sum_{i=1}^{N} |g(V_{(i)}, D_i|^2$$

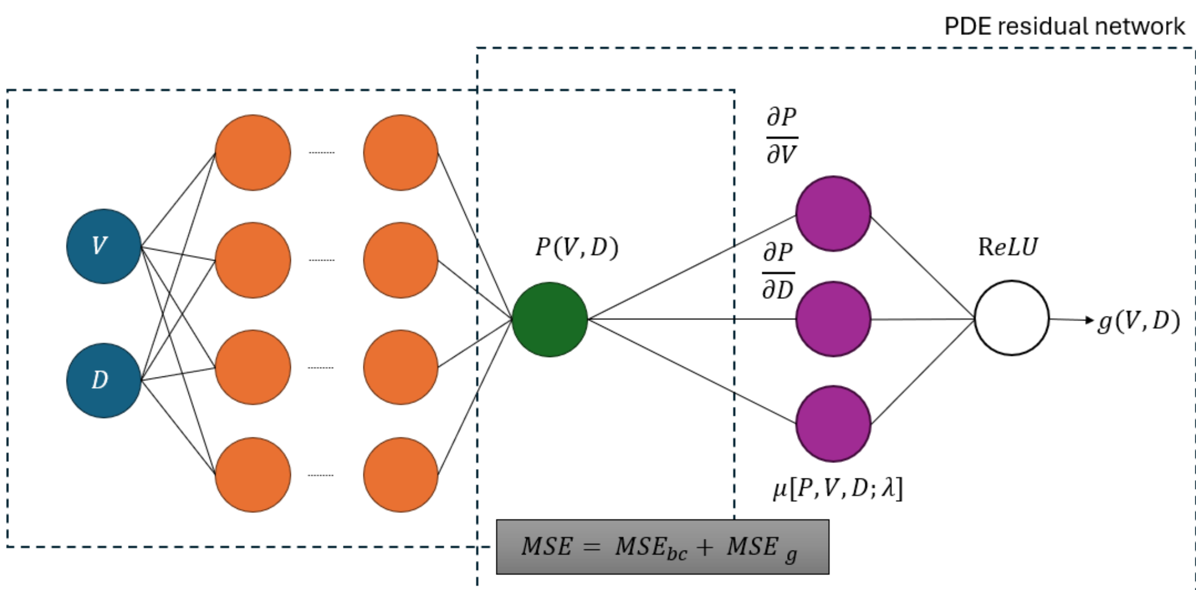


Figure 17 Schematic representation of the applied PINN for speed-power baseline modeling

The model training process is optimized using the Adam optimizer with a learning rate of 0.0001 to ensure stable convergence. Training is conducted over 500 epochs with a batch size of 16, enabling effective learning while mitigating the risk of overfitting. Power values are normalized before training to maintain numerical stability.

By integrating sea trial data and enforcing physics constraints, the PINN model enhances generalization and robustness. This approach ensures that speed predictions remain consistent with hydrodynamic principles while adapting to real-world operational conditions.

5.1.4. Limitations of the White-Box Model

While the white-box model provides a strong theoretical foundation, it has inherent limitations that make it insufficient as a standalone predictive tool. The model assumes calm water conditions, neglecting environmental disturbances such as wind resistance, wave effects, and hull fouling. Additionally, resistance-based formulations rely on empirical coefficients, which can vary significantly between ship designs, leading to potential inaccuracies. Furthermore, the model lacks adaptability, as it does not dynamically adjust to changes in operational conditions, including engine performance degradation or real-time weather variations.

To overcome these limitations, a black-box machine learning model is introduced to capture residual speed deviations caused by real-world factors. The final speed prediction is obtained by integrating the physics-based baseline with data-driven corrections in a grey-box modeling framework. The use of Physics-Informed Neural Networks (PINNs) further strengthens this approach by embedding physical constraints within the learning process, ensuring predictions remain both accurate and hydrodynamically consistent.

5.2. Machine Learning Model

While the PINN model effectively integrates hydrodynamic principles, it does not fully account for operational deviations resulting from factors such as hull fouling, weather conditions, and variations in propulsion efficiency. These deviations introduce systematic errors in the estimated ship speed (STW).

To address this, a data-driven black-box model is introduced, designed to learn and correct these residual deviations based on historical performance data. The black-box model is trained to predict the correction term, $STW_{correction}$, which represents the difference between the actual measured speed and the baseline speed estimated by the PINN.

This correction term accounts for environmental and operational uncertainties that the PINN does not explicitly model, such as variations in wind speed, wave height, and current effects. By learning these patterns from data, the black-box model improves speed predictions under real-world dynamic conditions.

5.2.1. Grey-Box Model: Integrating Physics-Based and Machine Learning Approaches

The grey-box model (GBM) represents a hybrid approach that combines the strengths of both white-box (physics-based) and black-box (machine learning-based) models. While the white-box model (WBM) provides physically interpretable estimates of ship speed (STW) based on hydrodynamic principles, it does not fully capture real-world deviations caused by environmental factors, operational inefficiencies, and hull fouling. Conversely, the black-box model (BBM) effectively learns these deviations but lacks the physical interpretability and generalization beyond its training data.

By integrating these models, the grey-box framework ensures that speed predictions remain physically consistent while being dynamically adaptable to real-world variations. This approach not only improves prediction accuracy but also maintains the relevance of hydrodynamic models for ship speed estimation.

5.2.2. Parallel Grey-Box Model Architecture

The parallel grey-box model is designed to run the white-box and black-box models simultaneously, using their combined outputs to refine power predictions. This approach differs from a serial architecture, where a machine learning model would first process raw inputs before passing them into a physics-based model. Instead, in the parallel grey-box model (Figure 18):

1. The WBM calculates baseline power ($STW_{baseline}$), assuming calm water conditions with no external disturbances.
2. The BBM predicts power deviations (ΔSTW), adjusting for real-world conditions such as wind, waves, and hull fouling.
3. The final predicted power (STW_{pred}) is obtained by summing the two outputs:

$$STW_{pred} = STW_{PINN} + \Delta STW$$

This parallel architecture ensures that the physics-based model anchors the predictions within known hydrodynamic principles, while the machine learning model dynamically refines the estimate based on real-world conditions.

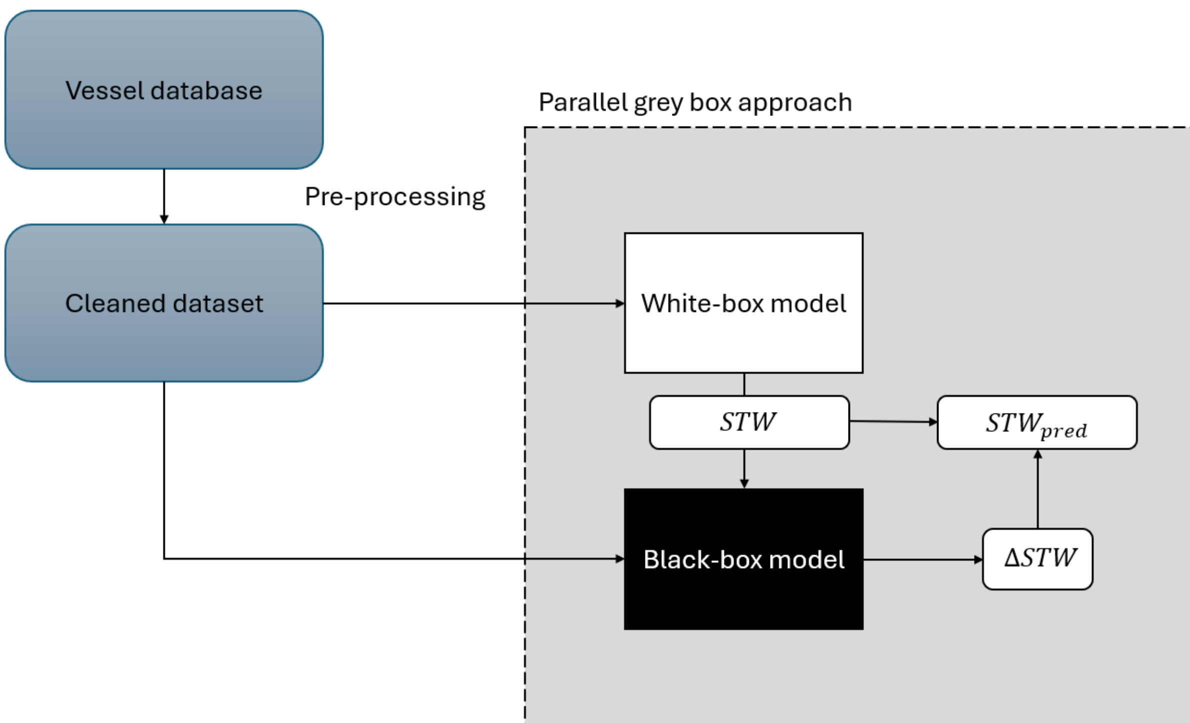


Figure 18 The parallel grey-box modeling procedure for ship speed prediction

5.2.3. Machine Learning Model Selection and Training

To determine the most effective model for predicting speed deviations (ΔV), multiple machine learning algorithms were trained and evaluated. The goal was to select a model that could accurately capture the complex, nonlinear relationships between operational conditions and power deviations.

The dataset was split into 80% training data and 20% testing data to ensure reliable model evaluation. To avoid overfitting, cross-validation techniques and hyperparameter tuning were applied during model training.

Several machine learning models were evaluated for their ability to predict, ΔV, including Linear Regression, Random Forest, and XGBoost. The goal was to determine the most effective model for capturing the complex, nonlinear relationships between operational conditions and power deviations. Performance metrics such as the coefficient of determination (R^2) and RMSE were employed to assess model accuracy and select the best-performing algorithm.

1. **Linear Regression**: Used as a baseline model, Linear Regression exhibited poor predictive performance (R^2 = 0.124, RMSE = 0.336). Its inability to capture the strong nonlinear dependencies between features and ΔV made it unsuitable for accurate predictions.

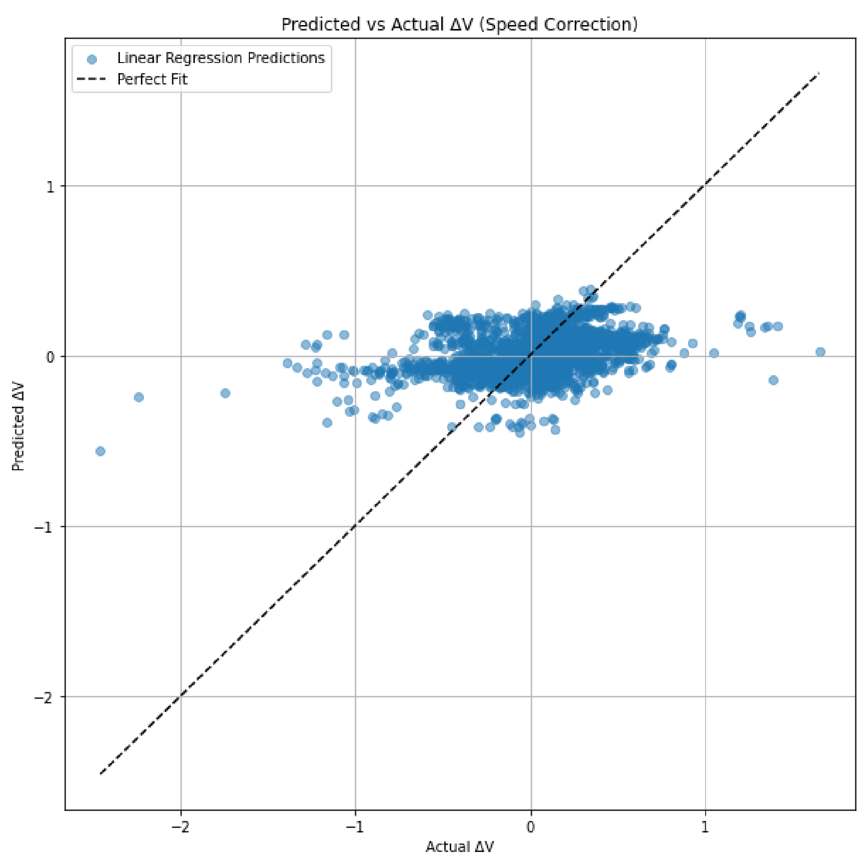


Figure 19 Linear Regression performance, showing the predicted power versus the baseline and measured power

2. **Random Forest Regressor**: This model performed exceptionally well (R^2 = 0.929, RMSE = 0.096), successfully capturing the nonlinear effects in the data. However, further analysis is required to assess potential overfitting.

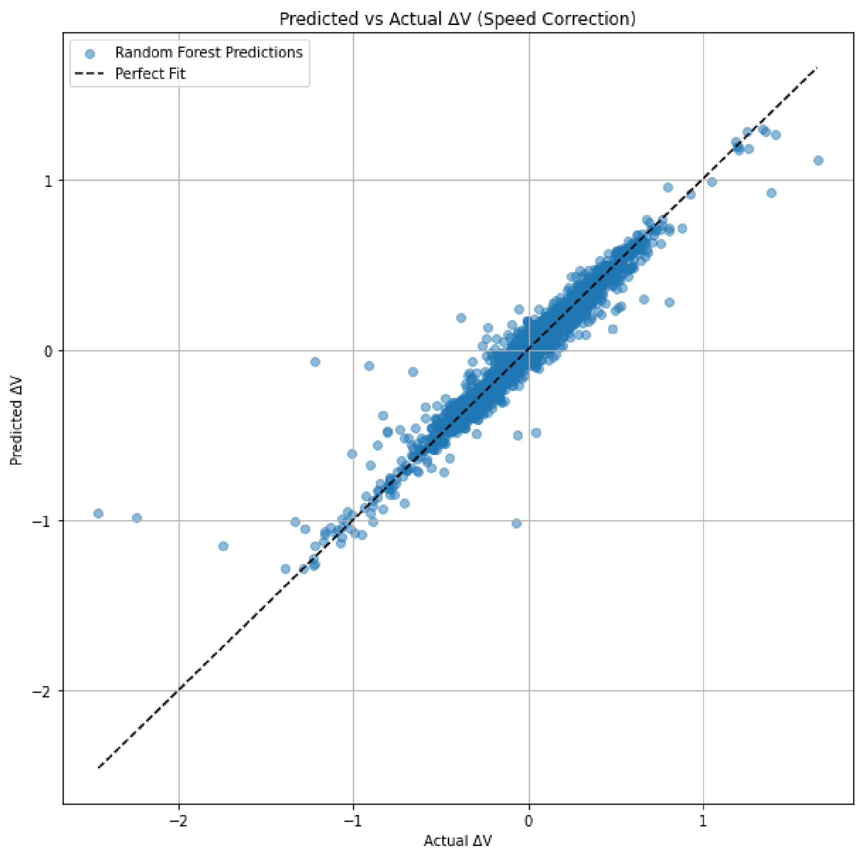


Figure 20 Random Forest performance, comparing predicted power with baseline and measured values.

3. **XGBoost (Extreme Gradient Boosting)**: Also demonstrated strong predictive accuracy (R^2 = 0.915, RMSE = 0.105), performing slightly worse than Random Forest but offering a more robust and generalizable approach.

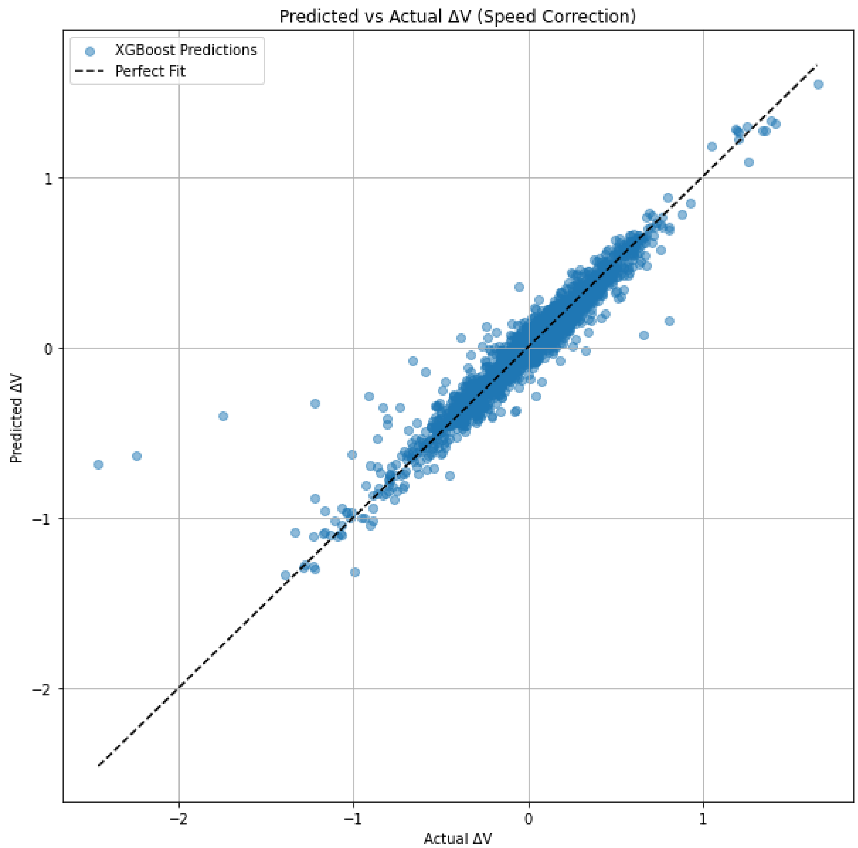


Figure 21 XGBoost model performance, illustrating the predicted power relative to the baseline and measured power.

The Random Forest and XGBoost models both performed well, but Random Forest was ultimately selected due to its slightly superior performance and easier interpretability in comparison to XGBoost.

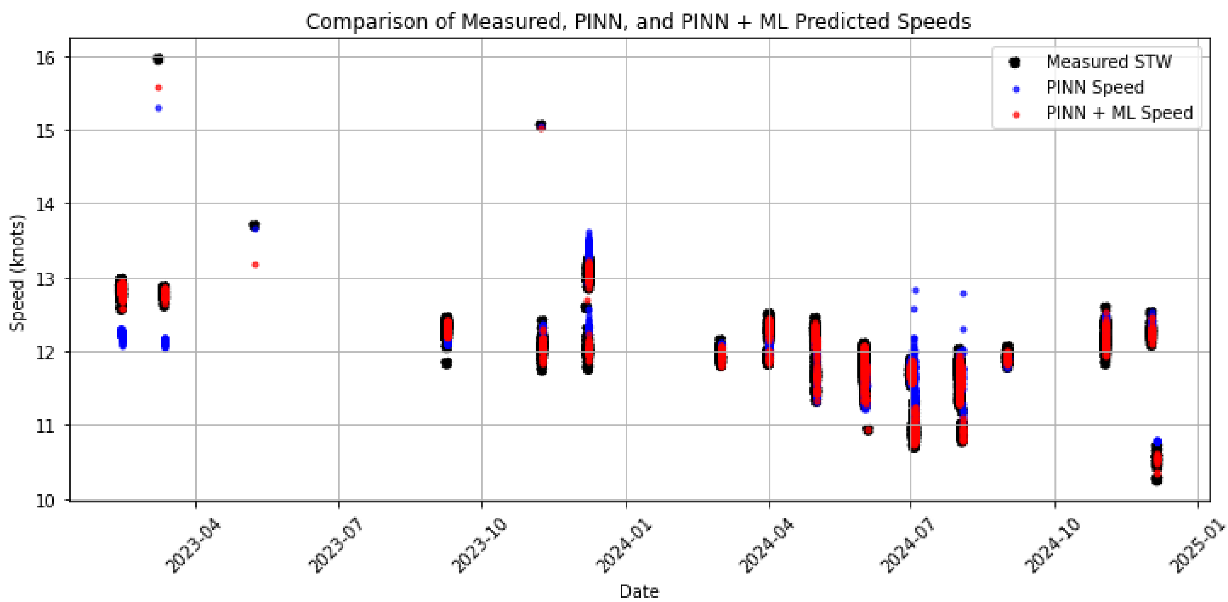


Figure 22 Comparison of sensor data, with results from PINN model, and the combination of PINN and ML models

Figure 22 presents a time-series comparison of the measured Speed Through Water, the PINN-derived baseline speed, and the final predicted speed, which incorporates machine learning corrections to the PINN output.

- The black dashed line represents the measured STW, providing a reference for actual vessel performance.
- The blue dots denote the PINN-predicted baseline speed, which is derived from a physics-based model without considering real-time operational variations.
- The red dots correspond to the final predicted speed, which results from the combination of the PINN baseline speed and the ML-predicted speed correction (ΔV).

The graph illustrates how the PINN + ML model aligns more closely with the measured STW, capturing variations in vessel performance that are not accounted for in the purely physics-based model. The integration of machine learning enables a more accurate representation of operational conditions, improving the predictive capability of the model.

6. Decision Support System Design

The maritime industry is undergoing a digital transformation, with vessels increasingly equipped with sensors, real-time monitoring capabilities, and advanced decision support systems (DSS). The continuous development of digital technology and connectivity enables real-time ship performance tracking and historical data storage in cloud-based systems. These advancements are driving both industry and academia to develop smart ship solutions, including digital twins, automation systems, and data-driven decision-making frameworks. As part of this trend, the proposed DSS in this thesis integrates a physics-informed machine learning model to enhance operational decision-making, fuel efficiency monitoring, and voyage planning.

A DSS is a computerized system that assists decision-makers by providing domain-specific insights, analytical tools, and real-time recommendations. In the context of smart shipping, the DSS can be installed onboard the vessel or at a remote-control center, where it provides operators with actionable intelligence on vessel speed, energy efficiency, and environmental impact [12]. Figure X illustrates how a DSS framework connects vessel systems to a remote monitoring center, facilitating predictive maintenance, fleet management, and optimal route planning.

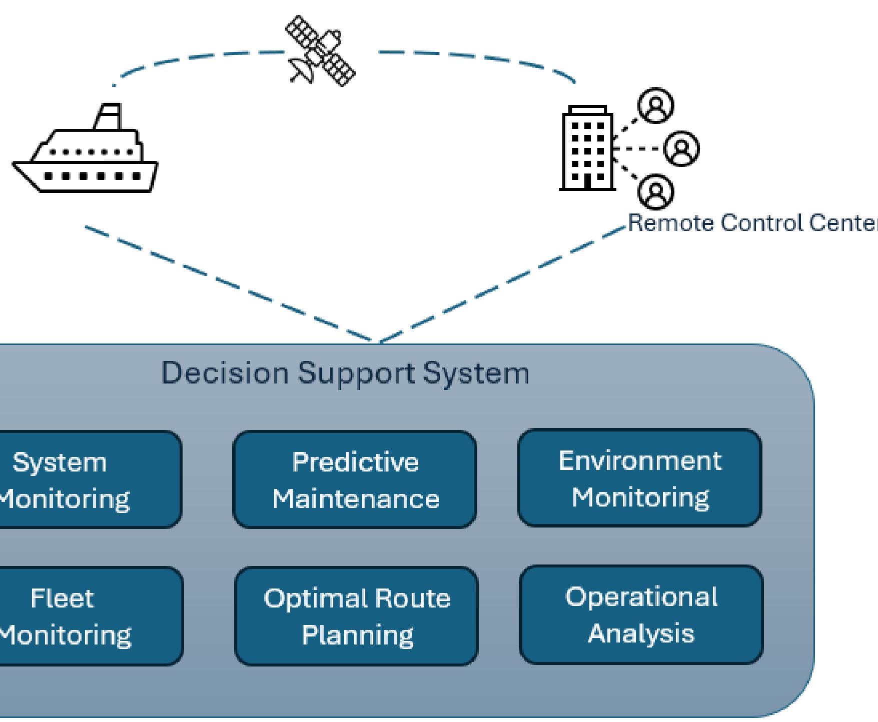


Figure 23 An illustration of smart ship and decision support system

The DSS proposed in this thesis consists of three core components:

- **Data Acquisition and Preprocessing –** This module continuously collects real-time operational data from onboard sensors, such as engine parameters, environmental conditions, and hull performance metrics. It ensures data integrity by handling missing values, aligning timestamps between sensor and noon report data, and filtering out anomalies.
- **Predictive Analytics –** This component integrates a PINN-based baseline speed model and a machine learning correction model (Random Forest/XGBoost) to estimate ΔV, refining speed predictions based on real-world operational variations. The system continuously updates predictions as new data becomes available.
- **Decision Support & Visualization –** The final component delivers insights through an interactive dashboard, enabling operators to compare measured STW, PINN-derived baseline speed, and ML-enhanced speed predictions over time. The DSS also includes anomaly detection features, alerting users to unexpected speed deviations that may indicate issues such as hull fouling, adverse weather effects, or inefficient fuel usage.

By leveraging this framework, the DSS enhances voyage planning and fuel optimization. Traditional voyage planning often relies on historical assumptions and static models, whereas the proposed DSS dynamically updates speed predictions based on real-time vessel conditions, weather forecasts, and fuel consumption trends. This capability ensures that operators can adjust engine load and routing strategies to minimize fuel costs while maintaining efficiency and regulatory compliance. The system supports IMO's Carbon Intensity Indicator (CII) and Energy Efficiency Operational Indicator (EEOI) frameworks, enabling shipowners to proactively monitor emissions and fuel efficiency trends.

Beyond speed prediction, DSS applications in smart shipping extend to condition monitoring, predictive maintenance, and risk assessment. The continuous digitization of vessels has transformed them into floating sensor hubs, enabling advanced machine learning applications. Several state-of-the-art DSS implementations include:

- Trajectory Prediction: Predicting vessel trajectory is critical for collision avoidance and navigation planning. Modern DSS solutions employ time-series forecasting models such as Long Short-Term Memory (LSTM) networks, which have demonstrated strong performance in AIS-based ship trajectory prediction [41].
- Fuel and Power Consumption Prediction: Fuel consumption models are typically formulated as regression problems, utilizing inputs such as speed, draft, weather, and engine conditions. Research has shown that machine learning models, ranging from linear regression to deep neural networks, can significantly improve fuel efficiency predictions, ultimately supporting route planning and energy management.
- Condition Monitoring of Machinery Systems: Advanced DSS implementations incorporate fault diagnostics and predictive maintenance algorithms to track the operational status of vessel machinery. Multi-label classification algorithms have been used for diagnosing

propulsion system faults [41], while ensemble machine learning methods have been applied for wear detection in marine diesel engines [42].

- Ocean Wave Forecasting and Estimation: Maritime operations are heavily influenced by sea state conditions, making wave estimation and prediction essential for voyage optimization. Machine learning approaches, including neural networks and random forests, have been employed to classify sea states and predict wave patterns, improving ship routing decisions in rough seas.

While this thesis focuses primarily on the modeling and predictive analytics component of a DSS, future extensions could enhance the system's adaptability and automation capabilities. One promising direction is adaptive machine learning, where models continuously retrain themselves using new operational data to improve long-term predictive accuracy. Additionally, reinforcement learning techniques could be explored for dynamic decision-making, allowing the DSS to recommend optimal speed and power settings based on real-time conditions.

Another key area for future research is cloud-based deployment and real-time API integration. Deploying the DSS on a cloud computing infrastructure would enable real-time inference and monitoring, ensuring minimal computational delays. An API-based system would allow integration with existing fleet management platforms, providing operators with a unified dashboard for monitoring vessel performance and making informed operational decisions.

In conclusion, the proposed DSS represents a significant step forward in integrating physics-informed machine learning into smart ship operations. By providing real-time speed predictions, fuel efficiency insights, and anomaly detection, the system enhances voyage planning, regulatory compliance, and predictive maintenance capabilities. As the maritime industry continues to embrace digitalization and AI-driven decision-making, such DSS implementations will play an increasingly vital role in optimizing fleet performance and supporting decarbonization efforts. Further advancements in adaptive learning, real-time deployment, and multi-vessel scalability will continue to refine the DSS framework, ultimately contributing to the evolution of autonomous and intelligent maritime operations.

7. Conclusion and Future Work

The maritime industry faces increasing pressure to optimize operational efficiency, reduce costs, and comply with stringent environmental regulations. One of the critical challenges in ship performance monitoring is accurately predicting vessel speed under varying operational and environmental conditions. Traditional speed-power models, while effective in controlled conditions, often fail to capture the complexity of real-world factors such as fluctuating weather, hull fouling, and dynamic sea states.

This research proposed a data-driven vessel speed prediction model, combining Physics-Informed Neural Networks (PINNs) and Machine Learning (ML) techniques to improve predictive accuracy. By leveraging sensor-derived ship performance data alongside physics-based estimations, the model effectively captured the nonlinear relationships governing vessel speed and power dynamics. The approach successfully addressed limitations associated with traditional empirical models, providing a more adaptable and robust framework for real-world operational conditions.

The study developed a PINN-based baseline speed model to establish an initial estimation of vessel speed, incorporating fundamental hydrodynamic principles. To refine this estimate, a machine learning model was trained to predict the necessary speed correction (ΔV), which accounted for deviations caused by external and operational influences. Several ML algorithms were tested, with Random Forest achieving the highest accuracy, demonstrating its ability to generalize well across diverse operational conditions. The final model was validated using measured Speed Through Water (STW) data, confirming its reliability in predicting vessel speed more accurately than standalone physics-based methods.

Beyond the development of the predictive model, this research also explored its integration into a Decision Support System (DSS). The DSS was designed to provide real-time insights for ship operators and fleet managers, assisting in voyage optimization, fuel efficiency monitoring, and proactive maintenance planning. The implementation of an interactive visualization tool enabled direct comparison of measured STW, PINN-estimated speed, and ML-corrected speed predictions over time. This system empowers maritime stakeholders with enhanced decision-making capabilities, ultimately leading to more efficient vessel operations and improved regulatory compliance.

Despite its promising results, the study faced several challenges. Data quality and availability were key constraints, as sensor inconsistencies and missing values introduced noise into the training dataset. Aligning noon report data with sensor readings required careful preprocessing to ensure data integrity. Additionally, while the PINN component ensured some level of generalization, the model was primarily trained on a specific vessel or fleet segment. Future research should explore adaptations for different vessel types, as variations in hydrodynamic characteristics may necessitate additional model tuning.

Computational complexity was another consideration. The integration of physics-informed models with machine learning introduced processing overhead, which may impact real-time implementation in a fleet management system. Deploying the model on a cloud-based architecture could enhance scalability and ensure minimal processing delays. Additionally, while the model incorporated key environmental and operational parameters, external factors such as hull aging, maintenance schedules, and fuel type variations may still influence its predictions. A self-learning mechanism that continuously updates the model with new operational data could further enhance its adaptability.

To advance the application of this research, several future directions are proposed. Expanding the model to multiple vessel types would improve its generalization, enabling application across various fleet segments. Enhancing model interpretability is another priority, as tree-based machine learning models like Random Forest and XGBoost, while highly accurate, function as black-box predictors. Techniques such as SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-Agnostic Explanations) could provide deeper insights into how individual features influence speed deviations, aiding in model validation and operational decision-making.

Further integration with voyage optimization systems could allow real-time speed and fuel efficiency recommendations based on weather forecasts and fuel prices. Developing a prescriptive analytics module that suggests optimal speed profiles under different conditions would enhance the system's practical utility. Additionally, transitioning the model to an API-based deployment would facilitate seamless integration with existing fleet management platforms, enabling real-time inference and monitoring.

An interesting avenue for exploration is the incorporation of reinforcement learning for adaptive optimization. Unlike traditional machine learning models, which generate static predictions, reinforcement learning could enable dynamic decision-making, allowing the system to adjust speed recommendations in response to changing environmental and economic conditions. Hybrid modeling approaches that integrate Computational Fluid Dynamics (CFD) simulations with machine learning corrections could also enhance predictive accuracy while maintaining physical consistency.

In conclusion, this research has demonstrated the effectiveness of combining physics-informed modeling with machine learning to improve vessel speed prediction. The proposed approach provides a scalable, data-driven solution that enhances operational efficiency, reduces fuel consumption, and supports regulatory compliance in the maritime industry. As the sector continues to embrace digitalization and AI-driven decision-making, the integration of such predictive models into decision support systems will play a vital role in optimizing fleet performance. Addressing the identified challenges and pursuing the proposed enhancements will further refine the model, paving the way for a real-time, intelligent system capable of predictive analytics and operational optimization. Ultimately, this research contributes to the ongoing transformation of maritime operations, supporting the industry's transition toward greater efficiency, sustainability, and compliance with global decarbonization efforts.

8. References

[1] United Nations Framework Convention on Climate Change (UNFCCC), “Paris Agreement,” 2015.

[2] UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT, *Review of maritime transport 2024: Navigating maritime chokepoints*. Geneva, 2024. [Online]. Available: https://shop.un.org

[3] Clarkson PLC, “2023 Annual Report Overview,” London, 2023.

[4] Equasis, “The 2022 World Merchant Fleet,” 2022.

[5] International Maritime Organization (IMO), “Fourth IMO GHG Study 2020 Full Report,” 2020.

[6] International Energy Agency (IEA), “World Energy Outlook 2020,” 2020. [Online]. Available: www.iea.org/weo

[7] International Maritime Organization (IMO), “RESOLUTION MEPC.304(72) (adopted on 13 April 2018) INITIAL IMO STRATEGY ON REDUCTION OF GHG EMISSIONS FROM SHIPS,” 2018.

[8] I. Energy Agency, “Net Zero Roadmap: A Global Pathway to Keep the 1.5 °C Goal in Reach - 2023 Update,” 2023. [Online]. Available: www.iea.org/t&c/

[9] European Union, “DIRECTIVE (EU) 2023/959 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL,” 2023.

[10] H. N. Psaraftis and S. Lagouvardou, “Ship speed vs power or fuel consumption: Are laws of physics still valid? Regression analysis pitfalls and misguided policy implications,” *Cleaner Logistics and Supply Chain*, vol. 7, p. 100111, Jun. 2023, doi: 10.1016/j.clscn.2023.100111.

[11] R. Yan and S. Wang, *Applications of Machine Learning and Data Analytics Models in Maritime Transportation*. The Institution of Engineering and Technology, 2022.

[12] P. Han, “Data-driven Methods for Decision Support in Smart Ship Operations,”

[13] N. Salvesen, “Added Resistance of Ships in Waves.,” *J Hydronaut*, vol. 12, no. 1, pp. 24–34, 1978, doi: 10.2514/3.63110.

[14] MAN Energy Solutions, “Basic principles of ship propulsion,” 2023.

[15] Siemens, “Full-scale simulation for marine design,” 2022.

[16] X. Lang, “Development of Speed-power Performance Models for Ship Voyage Optimization,” Gothenburg, Sweden, 2021. [Online]. Available: www.chalmers.se

[17] R. M. L. Madureira, “A data-driven approach for prediction and optimization of ship fuel consumption,” 2021.

[18] M. Zhang, N. Tsoulakos, P. Kujala, and S. Hirdaris, "A deep learning method for the prediction of ship fuel consumption in real operational conditions," *Eng Appl Artif Intell*, vol. 130, p. 107425, Apr. 2024, doi: 10.1016/j.engappai.2023.107425.

[19] C. Selimai, "Development of Neural Networks for Ship Speed Prediction," 2021.

[20] M. Levantis, S. Paereli, and W. Betz, "Speed-Power Models-A Bayesian Approach."

[21] X. Lang, D. Wu, and W. Mao, "Physics-informed machine learning models for ship speed prediction," *Expert Syst Appl*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.121877.

[22] G. Luka, M. Dino, M. Luka, and G. Izvor, "Machine learning techniques for modeling ships performance in waves," 2018.

[23] J. P. Panda, "Machine learning for naval architecture, ocean and marine engineering," Mar. 01, 2023, *Springer*. doi: 10.1007/s00773-022-00914-5.

[24] J. Carlton, "Marine Propellers and Propulsion," 2007.

[25] J. Holtrop and G. G. J. Mennen, "AN APPROXIMATE POWER PREDICTION METHOD," 1982.

[26] ITTC, "Procedures and Guidelines Testing and Extrapolation Methods Resistance Uncertainty Analysis, Example for Resistance Test," 2002.

[27] ISO, "ISO 15016 - Ships and marine technology — Guidelines for the assessment of speed and power performance by analysis of speed trial data," 2015. [Online]. Available: www.iso.org

[28] D. E. Hasselmann, M. Dunckel, and J. A. Ewing, "Directional Wave Spectra Observed during JONSWAP 1973," *J Phys Oceanogr*, Mar. 1980.

[29] ITTC, "Recommended Procedures and Guidelines Analysis of Speed/Power Trial Data," 2014.

[30] F. Chollet, *Deep Learning with Python*. MANNING, 2021.

[31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[32] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning."

[33] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 8, pp. 1798–1828, 2013, doi: 10.1109/TPAMI.2013.50.

[34] A. Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow."

[35] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[36] A. Burkov, "The Hundred-Page Machine Learning Book".

[37] "NIST/SEMATECH e-Handbook of Statistical Methods," National Institute of Standards and Technology. Accessed: Feb. 14, 2025. [Online]. Available: http://www.itl.nist.gov/div898/handbook/

[38] International Organization of Standarization, "ISO 19030-2: Ships and marine technology — Measurement of changes in hull and propeller performance — Part 2: Default method," 2016. Accessed: Feb. 15, 2025. [Online]. Available: https://www.iso.org/standard/63774.html

[39] P. Karagiannidis, "Data-driven Ship Propulsion modeling with applications in the Performance Analysis and Fuel Consumption prediction," National Technical University of Athens, Athens, 2019.

[40] P. Gupta, Y. R. Kim, S. Steen, and A. Rasheed, "Streamlined semi-automatic data processing framework for ship performance analysis," *International Journal of Naval Architecture and Ocean Engineering*, vol. 15, Jan. 2023, doi: 10.1016/j.ijnaoe.2023.100550.

[41] H. Tang, Y. Yin, and H. Shen, "A model for vessel trajectory prediction based on long short-term memory neural network," *Journal of Marine Engineering & Technology*, vol. 21, no. 3, pp. 136–145, May 2022, doi: 10.1080/20464177.2019.1665258.

[42] X. Xu *et al.*, "Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models," *Knowl Based Syst*, vol. 190, p. 105324, Dec. 2019, doi: 10.1016/j.knosys.2019.105324.