THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Voyage Segmentation and Propulsive Power Allocation: A Data-Driven Approach for Short Sea Shipping

Daniel Vergara

Department of Mechanics and Maritime Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden, 2025

Voyage Segmentation and Propulsive Power Allocation:
A Data-Driven Approach for Short Sea Shipping

Daniel Vergara

Chalmers University of Technology
Department of Mechanics and Maritime Sciences
Division of Marine Technology
SE-412 96, Gothenburg
Sweden
Telephone: +46 (0)31-772 1000
www.chalmers.se

Printed by Chalmers Reproservice
Gothenburg, Sweden, 2025

To all who have been important to me at any point and that have stayed in far lands.

Voyage Segmentation and Propulsive Power Allocation:
A Data-Driven Approach for Short Sea Shipping

Chalmers University of Technology
Department of Mechanics and Maritime Sciences
Division of Marine Technology

Abstract

Short-sea shipping, a sustainable alternative to land-based transport, faces strict environmental regulations and operational constraints to reduce fuel consumption, emissions, and costs. This thesis aims to minimise fuel consumption in short-sea shipping while adhering to sailing time constraints by developing a framework for optimising engine power allocation across predefined maritime routes. To address the limitations of existing power allocation methods, specifically their limited adaptability to metocean conditions, performance accuracy challenges, and long optimisation times, three approaches are examined: (1) Data-driven modelling, (2) Power allocation optimisation, and (3) Route segmentation.

The first part of the research project analyses a double-ended ferry. Here, data mining techniques were used to uncover trends in fuel consumption linked to power allocation of the ferry, revealing potential savings of up to 35% compared to actual operational data. Building on these findings, a decision support system (DSS) was developed, combining XGBoost to model fuel consumption and sailing time with Bayesian optimisation to recommend optimal engine speed and engine load. Full-scale experiments validated the DSS, achieving an average 18% reduction in the vessel's fuel consumption through the proposed engine power allocation strategies.

In the second half, the developed data-driven methods were combined with a novel voyage optimisation method performed in two steps. 1) Route segmentation: ship routes were segmented using the metocean score-based pruned exact linear time (MS-PELT) algorithm to identify optimal segments for engine power adjustments; 2) Engine power allocation, a scenario-based analysis grid was generated for each segment, and dynamic programming was used to determine the optimal power allocation for the voyage. The combined approach was tested on three years of data from a chemical tanker. Numerical simulations showed a 14% reduction in fuel consumption compared to measurement data, with sailing time deviations below 1%. This research demonstrates that the proposed framework significantly improves fuel efficiency in short-sea shipping while maintaining time constraints.

Keywords: Bayesian optimisation, double-ended ferry, dynamic programming, machine learning, power allocation optimisation, short sea shipping, voyage segmentation.

Preface

This thesis presents research conducted at the Division of Marine Technology, Department of Mechanics and Marine Sciences, Chalmers University of Technology, from June 2022 to February 2025. The Vinnova project 2021–02768 supported the research and the sustainable shipping program from Lighthouse/Trafikverket.

My PhD journey has been marked by challenges and growth, made possible by the support and mentorship of many. I express my sincere gratitude to Professor Wengang Mao, my main supervisor, for the opportunity of this journey, which has led to a significant learning experience, contributing to both my professional and personal growth. I want to express my sincere gratitude to my assistant supervisor, Xiao Lang, whose support has been indispensable. Thank you for all the hours you have spent teaching me to do research. A special thanks to Professor Jonas Ringsberg, my main examiner, whose dedication to my well-being and academic progress inspired me to improve. Likewise, I would like to thank Professor Per Högström; in hindsight, my career in Naval Architecture started with you accepting me into the program back in 2021; thank you for that. I am also thankful to my colleagues at the Division of Marine Technology. Especially to Rui Lopes, Qais Khraisat, Mohammad Shikholeslami and Stephan van Reen. It has been a nice experience with all our conversations during lunch. For those not mentioned by name:

"I don't know half of you half as well as I should like, and I like less than half of you half as well as you deserve. - Bilbo Baggins to his guests during his 111 birthday party."

To my brethren: Daniel and Anibal, Canonico and Julian, one day we will hunt, drink and feast again together. Edgar, may luck always be in your favour. Jose, that you visit every corner of the earth. Abu and Rahim, you will do great in life. May money always be abundant and your business prosper. To my brother from another mother, Kostas, thank you for the encouragement; I might have given up long ago without you. This list would easily be as long as my whole thesis if it weren't limited to a page, so here is a bunch of people I love... To Angelica, Chacho, David, Dina, Fabiola, Franklin, Gabriela, Gustavo, Jesus, Juan, Laura, Leo, Luis, Manu, Micol, Raul, Ricardo, Rosalia, Ronald, Sira, Yulfred... I care about all of you in no particular order (even when I am not always around).

Finally, my heartfelt thanks to my parents, Leonardo Vergara and Gladys Escalona, my brother Alexander, and my sister Sol... We fight a lot, but it is tough love. That money comes in our favour. To my cousin Leo... thanks for teaching me about sailing. To Elvira, thank you for standing by me. I wish you good health.

A special thanks to all the cattle for providing the delicious beef that made life worth living. Thank you for being so tasty... and to God, because even though I am not a believer, I still thank you for not making me a vegan.

Contents

List of appended papers

The author contributed substantially to each of the three attached papers, providing the main ideas and refining them in collaboration with the co-authors. The author managed data processing, conducted formal analysis, and/or developed algorithms to solve each research question, including clear visualisations to present the findings. The author wrote each manuscript, with final reviews and editing completed in collaboration with the co-authors.

Paper I D. Vergara, M. Alexandersson, X. Lang, and W. Mao. "Power allocation influence on energy consumption of a double-ended ferry." *In proceedings of the 33rd International Offshore and Polar Engineering Conference*, Ottawa, Canada, 2022. ISBN 978-1-880653-80-7. ISSN 1098-6189.

Paper II D. Vergara, M. Alexandersson, X. Lang, and W. Mao. "A machine learning-based Bayesian decision support system for efficient navigation of double-ended ferries." *Journal of Ocean Engineering and Science*, 2023. DOI: `https://doi.org/10.1016/j.joes.2023.11.002`

Paper III D. Vergara, X. Lang, M. Zhang, M. Alexandersson, and W. Mao. "Reduced environmental impact of short sea shipping through optimal engine power allocation. " Manuscript submitted for Journal Publication, 2024.

.

CRediT Authorship Contribution Statement

Daniel Vergara (All papers): Conceptualisation, Methodology, Data Curation, Investigation, Visualisation, Writing–Original Draft, Writing–Review and Editing

Martin Alexandersson (All papers): Conceptualisation, Formal analysis, Data Sourcing, Investigation, Writing–Review and Editing.

Xiao Lang (All papers): Conceptualisation, Visualisation, Validation, Supervision, Writing–Review and Editing.

Mingyang Zhang (Paper III): Formal analysis, Visualisation, Writing–Original Draft.

Wengang Mao (All papers): Conceptualisation, Validation, Funding Acquisition, Supervision, Project Administration, Writing-Review and Editing.

Nomenclature

Greek notations

$\alpha_{current}$	Sea current direction [°]
$\overline{\alpha_{current}}$	Average sea current direction [°]
α_{wave}	Wave direction [°]
$\overline{\alpha_{wave}}$	Average wave direction [°]
α_{wind}	Wind direction [°]
$\overline{\alpha_{wind}}$	Average wind direction [°]
β	Direction Score (MS-PELT) [-]
χ	Domain of an objective function/Search space [-]
δt	Time deviation in parallel scenarios [h]
γ	Overfitting factor PELT algorithm [-]
$\Gamma(\nu)$	Gamma function [-]
Γ_j	Power allocated to leg j [kW]
ι	Intensity Score (MS-PELT) [-]
λ	Regularisation parameter [-]
$\mu(\xi_{k+1})$	Mean function of the GP [-]
ν	Mattern kernel smoothness parameter [-]
ϕ	Cost function for the PELT algorithm [-]
ρ	Length scale parameter in the GP kernel
σ^2	Variance parameter of the GP [-]
$\Sigma_{i,j}$	Covariance matrix [-]
τ	Changepoints [-]
Θ	Inverse covariance matrix [-]
θ	Inverse covariance of a cluster [-]
φ_{eng}	Propeller pitch angle [°]
ξ_i	Sampled data points in BO [-]

Latin notations

3D-ACA	Three Dimensional Ant Colony Algorithm.
AIS	Automatic Identification System.
AI	Artificial Intelligence
BO	Bayesian Optimisation
$\mathcal{C}$	Cost function [-]
CFD	Computational Fluid Dynamics.
CO	Carbon Oxides.
C_f	Fuel cost
C_o	Operation cost
C_p	Delay penalties
DSS	Decision Support System
DP	Dynamic Programming.
$d_{1\to j}$	Accumulated sailing distance between departure and leg j.
$\Delta d_{j,k}$	Haversine/Great circle distance between two waypoints [km]
$\Delta t_{j,k}$	Sailing time between two waypoints [h]
ΔT_{j+1}	Allowable time of arrival window for leg $j+1$ [h]
E	Emissions [gCO_2]
EA	Evolutionary Algorithms.
ECA	Emissions Control Area.
EMO	Evolutionary Multi-Objective Optimisation.
ETA	Estimated Time of Arrival [h]
ETR	Extra Trees Regressor.
E_{max}	Maximum allowed emissions [gCO_2]
FFNNs	Feedforward Neural Networks.
FOC	Fuel Oil Consumption.
f	Generic objective function [-]
F_j	Bellman recursion of the fuel consumption up to leg j [ton]
f_j	Fuel consumption function for leg j [ton]
f_i	Measured quantity.
$\hat{f}_i$	Prediction estimation of f_i from χ_i

f_{fuel}	Fuel consumption prediction model [ton]
f_V	Speed overground prediction model [kt]
GA	Genetic Algorithm.
GBM	Grey Box Model.
GP	Gaussian Process
GPR	Gaussian Process Regressor.
H_s	Significant Wave Height [m]
$\overline{H}_s$	Average wave height [m]
i	Index i, generic index.
J	Generic objective function.
j	Index j, often refers to a leg.
KNN	K-nearest neighbours.
k	Index k, often refers to a waypoint inside a leg.
LOESS	Locally Estimated Scatterplot Smoothing.
LSTM	Long Short-Term Memory.
$\mathcal{L}$	Loss function [-]
MC	Monte Carlo simulations
MCR	Maximum continuous rating (MCR) [kW]
MILP	(Mixed Integer) Linear Programming.
MLR	Multi-Linear Regression.
MPC	Model Predictive Control.
M_{fuel}	Total fuel consumption of the voyage [ton]
M_{east}	Fuel consumption eastbound.
M_{west}	Fuel consumption westbound.
MSE	Mean squared error
m_{fuel}	Fuel consumption rate [$\frac{ton}{h}$ or $\frac{l}{h}$]
m_j	The total number of waypoints in a leg
n	Index n
n_j	Total number of legs
$NLCB$	Negative Lower Confidence Bound
NM	Nautical Miles.
NNs	Neural Networks.

NSGA-III	Non-dominated Sorting Genetic Algorithm III.
Q	Total number of reference Monte Carlo voyages
PELT	Pruned Exact Linear Time.
PSO	Particle Swarm Optimisation.
P	Engine power [kW]
$P_{\mathbf{bow}}$	Power bow engine [kW]
P_j	Power at leg j.
$P_{\mathbf{max}}$	Maximum allowed power [kW]
$P_{\mathbf{min}}$	Minimum allowed power [kW]
$P_{\mathbf{stern}}$	Power stern engine [kW]
P_T	Thrust power [kW]
Q	Total number of reference Monte Carlo voyages
$\mathcal{R}$	Route set [-]
RPM	Engine speed [rpm]
R_P	Double-ended ferry power ratio [-]
S	Ship state vector
SVM	Support Vector Machine.
$SFOC$	Specific fuel oil consumption $[\frac{ton}{kWh}]$
T	Thrust [kN]
t	Time [h]
$t^0_{j+1}, 1$	Nominal departure time from leg $j + 1$ [h]
$t_{earliest}$	Earliest time of arrival
t_{latest}	Latest time of arrival
$TICC$	Toeplitz Inverse Covariance-Based Clustering
$T_{\mathbf{arrival}}$	Predicted time of arrival [h]
$T_{\mathbf{opt}}$	Optimal sailing time for a specific voyage [h]
T_{mean}	Mean draught [m]
T_z	Wave period [s]
V	Ship speed overground
$V_{\mathbf{average}}$	Average speed [kt]
$V_{current}$	Surface current speed $[\frac{m}{s}]$
$\overline{V_{current}}$	Average surface current speed $[\frac{m}{s}]$

$V_{j,k}$	Speed overground between two waypoints [kt]
$V_{\mathbf{max}}$	Maximum allowed speed [kt]
$V_{\mathbf{min}}$	Minimum allowed speed [kt]
V_{wind}	Wind speed [$\frac{m}{s}$]
w.r.t.	With respect to.
$\overline{V_{wind}}$	Average wind speed [$\frac{m}{s}$]
W	Weather vector [-]
x	Longitude [°]
$\hat{x}$	Approximate longitude [°]
y	Latitude [°]
$\hat{y}$	Approximate latitude [°]

CHAPTER 1

Introduction

1.1 Background and motivation

Short-sea shipping refers to transporting cargo and passengers over relatively short distances by sea, usually within the same continent or region. It offers a sustainable alternative to road transport, reducing environmental impact (Brooks et al., 2014; Douet & Cappuccilli, 2011; Papadimitriou et al., 2018). Additionally, it alleviates motorway congestion and provides logistical advantages in coastal regions (Comi & Polimeni, 2020; Fadda et al., 2020). This form of maritime transport primarily involves moving goods between nearby ports without crossing oceans, thereby benefiting coastal areas by reducing reliance on trucks and railways (Fadda et al., 2020; Mulligan & Lombardo, 2006; Papadimitriou et al., 2018; Raza, 2020). It offers an effective solution for moving cargo over short distances and relieving the strain on road infrastructure, particularly in areas where transport networks are often congested. Ships provide higher fuel efficiency and larger cargo capacity than other forms of transportation, making them a sustainable choice with significantly lower emissions per tonne-mile (Comi & Polimeni, 2020; International Maritime Organization, 2009). Studies have further highlighted their effectiveness in reducing environmental impacts through improved energy efficiency (Raza, 2020; H. Wang, 2020). Short-sea voyages benefit from higher reliability in metocean forecast data, as noted in (H. Wang, 2020). These metocean forecasts can then be integrated into operational optimisation and decision support systems to improve energy efficiency. Integrating real-time metocean data further optimises routes, reduces delays, and enhances safety, especially in areas with significant metocean conditions (H. Wang et al., 2020).

Optimisation algorithms to improve energy efficiency in voyages have been a common subject of academic research, with voyage optimisation research mainly focusing on optimising speed profiles and optimal routes to reduce fuel consumption. Extensive methods for solving the minimum speed-fuel cost problem are available in the literature. However, speed optimisation is a limited approach since ship speed is indirectly controlled by the propulsion system, which generates thrust to overcome resistance. Achieving target speeds requires regulating engine speed, power, torque, propeller pitch, or combinations of these parameters. Though less explored, engine power optimisation offers potential benefits in actual operations. However, unlike speed optimisation, it demands precise speed estimation to maintain reliable ETA predictions (H. Wang et al., 2019). Many modelling approaches for speed prediction are also available in the literature, with data-driven models becoming a dominant force in the field, at least when data is available (Lang et al., 2024). These models are necessary to describe voyage optimisation problems. In general, a voyage optimisation problem consists of three key components:

- A mathematical description of the voyage.
- Defining the optimisation *Objective Function*, *Constraint Function*, and *Control Variable(s).*
- Selecting a suitable *Optimisation Algorithm* to solve the optimisation problem.

A mathematical model represents each part of the optimisation problem: find the efficient operational strategy when sailing between ports A and B so that the ship arrives at time T. The objective function is expressed as the voyage cost or fuel consumption (K. Wang et al., 2020; Zaccone et al., 2018), calculated by estimating the fuel consumption rate and integrating it across a sequence of decisions from departure to destination. The constraint function is clearly defined, including limits on ship speed, engine power, and sailing time or ETA. The control variable is obtained through an optimisation algorithm applied to the voyage. Typically, the voyage is modelled as a series of stages (legs or waypoints), with a constant value assigned to the optimisation control variable for each stage.

For optimal navigation of short-sea shipping, one of the big challenges is understanding those ships' performance in terms of their complex/flexible propulsion systems. However, there is currently limited research on performance modelling for double-ended ferries. With the increase in shipping digitalisation, data collection systems have been installed onboard to monitor ship operation performance and systematically collect useful data. Therefore, data analysis techniques were used to understand ships' regular operation and identify hidden trends for fuel savings. Then, machine learning models were used to estimate ship speed and fuel consumption under varying power settings and sea conditions.

Furthermore, many voyage optimisation studies rely on fixed waypoints, defined by specific distances/time intervals or simple port-to-port segmentation. Such approaches fail to account for changing metocean conditions that can substantially impact fuel efficiency, and it is often unclear how route segmentation is conducted in these studies. This work employs adaptive segmentation methods based on real-time metocean data, ensuring that each segment experiences relatively similar metocean conditions while simultaneously reducing the number of segments to enhance optimisation effectiveness.

To consider such types of ship navigations in voyage optimisation systems requires a shift from traditional speed-based voyage optimisation to more sophisticated algorithms that account for reliable metocean forecasts and precise engine control. The optimisation methods should help to determine how to optimally allocate the engine power in terms of dynamic metocean environments encountered by the ship.

1.2 Literature review

1.2.1 Data–driven ship performance modelling

The evolution of ship performance modelling has gone through a transition from theoretical assumptions (white box models) and simplified formulas based on statistics of low-resolution data (semi-empirical models) to the applications of Big Data and Artificial Intelligence (AI) techniques to increase the quality of predictions.

White-box models rely on fundamental physical principles to estimate propulsion power, speed, and fuel consumption. Standard methods include empirical hull resistance calculations and energy balance equations (Lang & Mao, 2021). Notable models by Holtrop and Mennen (1982) and Hollenbach (1998) have been widely applied to conventional merchant vessels, such as tankers and bulk carriers, for initial predictions under calm water conditions. However, these models rely on many simplifications, making them less effective in dynamic or complex situations (Lang & Mao, 2020).

Machine-learning models do not rely on rigid physical assumptions, unlike white-box and semi-empirical models. Instead, they learn patterns directly from historical operational data, allowing them to capture complex, non-linear relationships between ship speed, propulsion power, and environmental conditions.

Recent advancements in describing a ship's performance at sea have been dedicated to developing statistical and machine learning models for both ship performance Malte Mittendorf et al. (2023) and Mao et al. (2016) and ship response at sea Mao et al. (2015) and Schirmann et al. (2023), resulting in more robust prediction tools compared to white box models. Comparative studies have shown the importance of selecting appropriate ML models for ship performance tasks. For instance, Lang et al. (2022) compared multiple ML models for predicting ship speed and power and concluded that XGBoost outperforms other methods due to its robustness across varying conditions. Extending this idea, Lang et al. (2024) integrated XGBoost with Physics-Informed Neural Networks, forming a Grey Box Model (GBM) that combines data-driven flexibility with the interpretability of physics-based models.

Laurie et al. (2021) evaluated five ML models—Adaboost, Multiple Linear Regression (MLR), K-Nearest Neighbors (KNN), classical Neural Networks (NNs), and Random Forest—and found that Random Forest achieved the best performance for predicting ship biofouling effects. Similarly, Abebe et al. (2020) trained and compared tree-based ML models such as Decision Trees, Random Forest, Extra Trees, Gradient Boost, and Extreme Gradient Boost using Automatic Identification System (AIS) data, indicating that the Extra Trees Regressor (ETR) provided the most accurate predictions of ship speed.

Artificial Neural Networks (ANNs) have been the focus of attention for various ship performance predictions. Beşikçi et al. (2016) developed an ANN-based Decision Support System to estimate fuel consumption using noon reports. In contrast, using three months of ship measurement data, Parkes et al. (2018) applied ANNs to predict shaft power. Their results suggested that increasing the depth of the networks could improve accuracy. Similarly, Karagiannidis and Themelis (2021) employed Feed Forward Neural Networks (FFNNs) to predict speed and power, introducing an imputation algorithm to address missing data during preprocessing. However, the nonlinearity of ANNs can sometimes hinder generalisation, as observed by Kim et al. (2021), who compared Multiple Linear Regression and ANNs for predicting fuel consumption and emphasised the need for careful model design to avoid overfitting.

Broader algorithmic comparisons have also been conducted to assess the effectiveness of different ML techniques. Bassam et al. (2022) compared linear regression models, regression trees, ensembles, Gaussian Process Regression (GPR), and Support Vector Machines (SVMs) in terms of their prediction accuracy, highlighting that no single method universally outperforms others without considering the specific dataset and context. These findings underscore the importance of model selection and tuning when applying ML techniques to ship performance modelling tasks.

However, ML models require many processing steps to guarantee data quality, requiring careful preprocessing to remove noise and outliers. Techniques such as Locally Estimated Scatterplot Smoothing (LOESS) (Cleveland, 1979), Savitzky-Golay filtering (Savitzky & Golay, 1964), and rolling averages are commonly used to prepare data for analysis (Karagiannidis & Themelis, 2021). Finally, proper feature selection further enhances model performance (Abebe et al., 2020; Kim et al., 2021).

1.2.2 The voyage optimisation problem

Optimising propulsion power during a voyage is crucial for enhancing energy efficiency and reducing operational costs. Traditional voyage optimisation methods have focused on the ship's speed as the primary control variable, and extensive literature covers methods for solving speed optimisation problems to identify optimal speed profiles for minimal fuel consumption and computational efficiency. Methods such as evolutionary algorithms (EA), dynamic programming (DP), particle swarm optimisation (PSO), and linear programming (MILP) are among the methods covered in the literature.

EAs, especially Genetic Algorithms (GA), have been extensively employed in ship voyage optimisation due to their ability to handle complex, multi-variate problems. Lee et al. (2018) formulated a simultaneous path and speed optimisation model, addressing the interdependencies between speed adjustments and route planning using RPM and headings as decision variables in a GA. Szlapczynska and Szlapczynski (2019) applied evolutionary multi-objective optimisation (EMO) for ship weather routing, utilising a trade-off-based approach to balance objectives like passage time, fuel consumption, and safety. Similarly H. Wang et al. (2021) integrated a GA with DP to minimise fuel consumption and emissions. Z. Li et al. (2024) proposed a collaborative optimisation framework combining GA-improved Long Short-Term Memory (GA-LSTM) with the Non-dominated Sorting Genetic Algorithm III (NSGA-III) to optimise speed, route, and trim simultaneously.

DP-based methods have also been a common approach to optimise both route and speed. Zaccone et al. (2018) used dynamic programming to develop a three-dimensional energy-efficient route-speed optimisation. Y. Du et al. (2019) used DP to solve a trip-speed optimisation problem of tankers considering weather data. Ma et al. (2020) proposed a method that optimises route and speed by considering emission control areas (ECAs) and changing weather conditions. Tzortzis and Sakalis (2021) introduced a dynamic optimisation approach with time horizon segmentation to account for the declining accuracy of long-term weather forecasts and improve the optimisation in dynamic sea environments. Fan et al. (2022) used DP in a two-dimensional approach to simultaneously solve the optimal speed and trim of the vessel along a fixed route using ANNs to model the ship's performance.

PSO has also become popular for solving the voyage optimisation problem. K. Wang et al. (2020) introduced a method to optimise speed and route using PSO while considering a fixed grid and environmental conditions. Dai et al. (2022) combined an ML power load forecasting model using a support vector machine (SVM) and used a PSO algorithm to determine the optimal power load and improve prediction accuracy for ship power management. W. Du et al. (2023) extended the use of PSO in route optimisation by proposing a second-order oscillating PSO algorithm that considers real-time sea conditions, resulting in fuel oil consumption (FOC) and CO emission reductions.

Other methods in the literature include H. Wang et al. (2019), which used a three-dimensional Djikstra algorithm for simultaneous speed and route optimisation and fuel efficiency. C. Zhang et al. (2022), which developed a multi-objective optimisation model for Arctic ice routing using a three-dimensional ant colony algorithm (3D-ACA) to minimise fuel consumption and navigation risks under time-varying ice conditions. K. Wang et al. (2021) introduced a dynamic, collaborative optimisation method that combines spatial and temporal analysis of environmental factors with Model Predictive Control (MPC) and swarm intelligence algorithms for fuel saving. Bahrami and Siadatmousavi (2024) applied an iterative Dijkstra algorithm to adjust the network weights based on met-ocean parameters dynamically to reduce fuel consumption. Vergara et al. (2023) combined Bayesian optimisation (BO) with XGBoost models to solve the power allocation problem of a double-ended ferry, showing an application outside of parameter tuning in ML. Yu et al. (2024) introduced a proxy method to optimise trim by fitting splines on samples from CDF simulations. Most recently, Shang et al. (2024) used reinforcement learning to optimise an electric ship's power generation and sailing speed, improving their optimisation under uncertainty.

However, the previous methods are limited in practice because the propulsion system indirectly controls ship speed, and maintaining a constant or piecewise set-point speed requires continuous propulsion adjustments to respond to changes in the metocean conditions. Frequent engine speed, power, and propeller pitch adjustments can reduce energy efficiency and accelerate component wear (Sørensen, 2013; Sørensen et al., 1997; Yu et al., 2024).

1.2.3 Route segmentation

Optimal segmentation of maritime routes can enhance a ship's performance and energy efficiency. Route segmentation enables optimal voyage speed or power allocations that account for varying metocean conditions along a voyage. In speed optimisation studies, existing methods typically allocate speed to route legs using equal distance or time intervals, structured around grids or waypoints (H. Wang et al., 2019; Zaccone et al., 2018). In liner shipping or cargo allocation scenarios, each leg usually spans between two ports (Guericke & Tierney, 2015; Qi & Song, 2012; S. Wang & Meng, 2012; Wu, 2020). These strategies often disregard the influence of varying metocean conditions on ship performance during a single voyage.

Several methods have been introduced in academic research for segmenting fixed routes. Clustering algorithms applied to time-series metocean data are commonly used to distinguish different sea states. K-means clustering, one of the most popular algorithms in this context, groups data points into clusters based on their proximity, allowing for the segmentation of routes by grouping consecutive waypoints with similar metocean conditions (K. Wang et al., 2017; Yan et al., 2018). Similarly, turning point detection methods have been employed to identify significant changes in metocean conditions along a route (X. Li et al., 2022; M. Zhang et al., 2024). These methods analyse time-series data to detect points where the statistical properties change, thereby segmenting the route at these points. While effective at dividing routes into segments, they often overlook how metocean conditions directly impact ship performance, particularly the non-linear and complex relationship between metocean variables and the vessel's fuel consumption and speed. Moreover, when detailed metocean conditions are considered, processing extensive datasets can reduce computational efficiency. This approach often leads to excessively segmented routes, impractical for real-world implementation due to increased complexity in voyage planning and execution (X. Li et al., 2023; K. Wang et al., 2020). These challenges show the need for more efficient segmentation methods that balance accuracy with computational feasibility.

Alternative methods for time-series segmentation from other domains hold potential for application in maritime route segmentation. The Pruned Exact Linear Time (PELT) algorithm, introduced by Killick et al. (Killick et al., 2012), is an efficient technique for change point detection in time-series data. PELT optimises segmentation by minimising a cost function and balancing model fit and complexity while maintaining computational efficiency even with large datasets. It has been applied to classify wave data by detecting shifts in statistical properties, which is crucial for classifying distinct sea states. Similarly, the Toeplitz Inverse Covariance-Based Clustering (TICC) method proposed by Hallac, Leskovec, and Boyd (2017) provides a framework for segmenting multi-variate time-series data by modelling temporal dependencies and clustering structures. TICC can detect non-linear relationships between metocean conditions along the voyage directly related to ship performance. These methods are underexplored in the context of maritime route segmentation but have the potential to yield improved results with an increased computational performance.

1.3 Objectives, goals and contributions

The main objective of this thesis is to develop and demonstrate an integrated power allocation optimisation framework to support decision-making in short-sea shipping. The framework aims to minimise fuel consumption, reduce operational costs, and enhance efficiency while adhering to ETA constraints. The proposed framework applies data-driven machine learning models, which enable enhanced ship performance modelling. It incorporates a power allocation-based voyage optimisation approach, providing a more practical alternative to speed-based optimisation by reducing the need for frequent engine adjustments. The framework should enable voyage segmentation based on relatively consistent metocean conditions while minimising the number of segments to maintain computational efficiency.

To achieve the overall objectives, this thesis investigates the following several specific research goals:

1. Investigate historical operational data from a double-ended ferry to identify power allocation patterns and fuel consumption trends.
2. Implement machine learning models based on operational and environmental factors to predict ship fuel consumption and sailing time.
3. Introduce a metocean-based voyage segmentation method to ensure power allocation adjustments align with environmental conditions.
4. Formulate power allocation optimisation framework combining scenario-based analysis and dynamic programming.
5. Evaluate the proposed methodology through numerical simulations and full-scale experiments to quantify potential fuel savings and operational improvements.

1.4 Assumptions and limitations

The following assumptions and limitations are applied to streamline the power allocation optimisation process:

1. Measurement data, including propulsion and motion-related metrics, are assumed to represent the ground truth. Sensor uncertainties and measurement noise are disregarded in this analysis.
2. Environmental data from hindcast databases, without onboard measurements, are used to approximate actual conditions along the ship's route.
3. Effects such as biofouling on the hull and rudder-induced resistance are neither available nor identifiable from the sensor data and are therefore assumed minimal. They are excluded from the optimisation model.

4. Metocean data included in the optimisation is limited to wave, wind, and current information. Other environmental factors are considered negligible in their impact on fuel consumption and are omitted.
5. Only open waters data are utilised for data-driven modelling; transient conditions and manoeuvres are not considered in the models. The models are only valid for the quasi-steady state operation.
6. Operational parameters and hindcast data between waypoints are assumed constant.
7. The effects of Earth's curvature are considered negligible for distances of up to 15 NM.

The methodology proposed in this thesis was developed using data from open waters under quasi-steady state conditions and common metocean conditions the ship may encounter during regular sailing. As such, the method has not been validated for extreme conditions such as severe storms, unusually high current speeds, or ice-covered waters. Users should exercise caution when applying the optimisation framework outside these typical operational scenarios.

1.5 Thesis outline

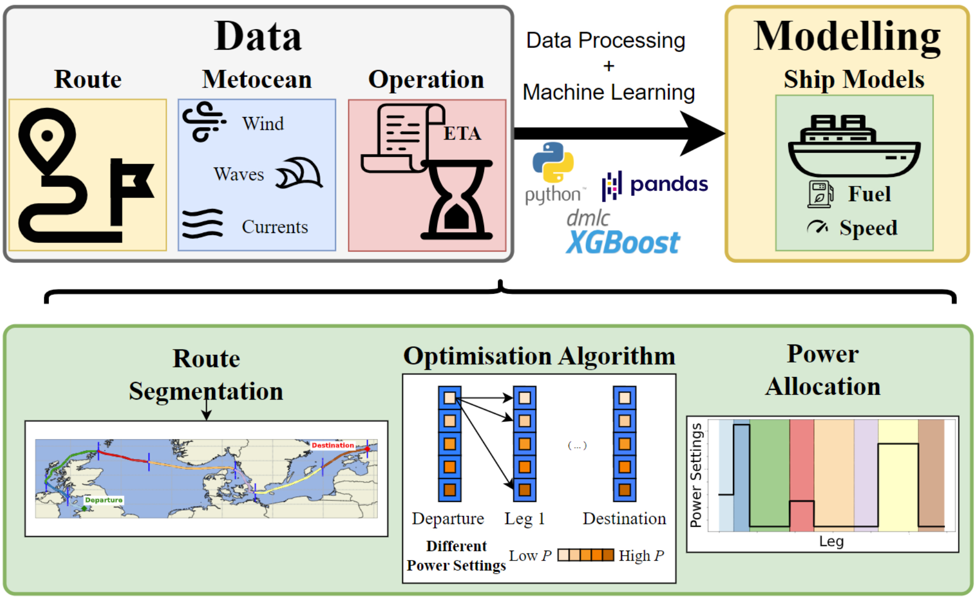


Figure 1.1: Complete thesis framework.

Figure 1.1 presents the general framework of the thesis to optimise ship performance using data-driven techniques and machine learning-based power allocation. It is presented as follows:

Chapter 2 outlines the methods for power allocation optimisation used throughout the research. Chapter 3 describes the case study ships, the available data, and the data-driven vessel modelling techniques. Chapter 4 presents the key findings from each of the papers. Finally, Chapters 5 and 6 conclude the work and outline the next steps for this research project.

CHAPTER 2

Method for optimal power allocation in ship operations

2.1 Power allocation optimisation

Power allocation optimisation refers to planning a ship's power usage before a journey in a way that minimises the operational costs (fuel consumption, crew salaries, etc), maximises the sailing efficiency and meets operational and contractual requirements. The optimisation framework used across all the document and papers is shown in Fig. 2.1. Mathematically, it can be described as solving the optimisation problem from Equation 2.1.

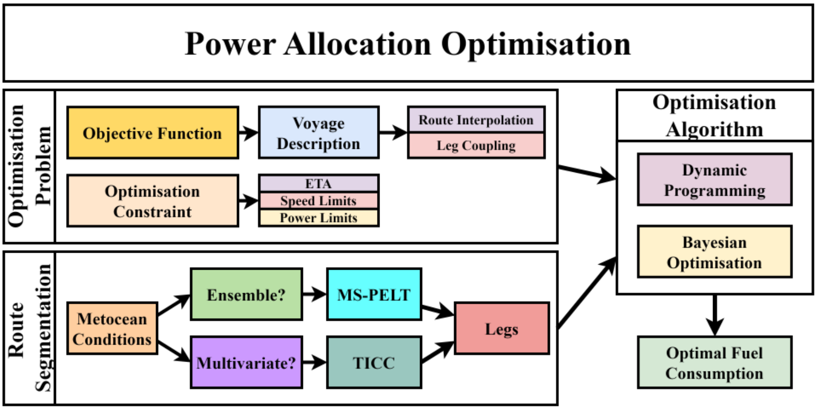


Figure 2.1: Framework for power allocation optimisation.

$$
\begin{aligned}
&\min_{P(t)} J = \int_{t_0}^{t_f} \Big(C_f(V(t), P(t)) + C_o(t) + C_p(D(t)) \Big)\, dt \\
&\text{s.t.} \\
&V_{\min} \leq V(t) \leq V_{\max}, \\
&P_{\min} \leq P(t) \leq P_{\max}, \\
&(x(t), y(t)) \in \mathcal{R}, \\
&t_f \in [t_{\text{earliest}}, t_{\text{latest}}], \\
&E(t) \leq E_{\max},
\end{aligned}
\tag{2.1}
$$

Where C_f is the fuel cost, C_o the cost to run the operation, C_p the cost of any delays, t_0 the departure time, t_f the arrival time, V the sailing speed, P the engine power, $\mathcal{R}$ the sailing route and E the emissions. In this thesis the subject of interest is to minimise C_f under the ETA constraint while sailing on a fixed $\mathcal{R}$, therefore Equation 2.1 simplifies to,

$$
\begin{aligned}
&\min_{P(t)} J = \int_{t_0}^{t_f} \Big(C_f(V(t), P(t)) \Big)\, dt \\
&\text{s.t.} \\
&V_{\min} \leq V(t) \leq V_{\max}, \\
&P_{\min} \leq P(t) \leq P_{\max}, \\
&(x(t), y(t)) \in \mathcal{R}, \\
&t_f \in [t_{\text{earliest}}, t_{\text{latest}}],
\end{aligned}
\tag{2.2}
$$

To solve this problem, it is required to develop a mathematical model of the ship and simulate how it sails along $\mathcal{R}$.

2.1.1 Voyage description

Since the objective is only on minimising the fuel consumption, the ship is simplified to a point mass point described with a general state vector $\mathbf{S}$ defined by its coordinates (latitude and longitude) and timestamp. Furthermore, the voyage is divided into legs to simplify the optimisation problem. The state of the ship, when it has reached the $k-th$ waypoint within the j-th leg, is described as,

$$
\mathbf{S}_{j,k} = [x_{j,k}, y_{j,k}, t_{j,k}], \tag{2.3}
$$

Where $x_{j,k}$ and $y_{j,k}$ are the spatial coordinates, and $t_{j,k} \geq t_0$ is the timestamp. A first problem arises: real sailing happens in real-time, and computes operate in discrete time; therefore, the digital description of the voyage needs to be digital; for that, a zero-order hold is used; that is, the ship's state is only updated at the next visited waypoint.

$$\mathbf{S}_{j,k}(t) = \mathbf{S}_{j,k}, \quad \text{for } t \in [t_{j,k}, t_{j,k+1}) \tag{2.4}$$

It would be convenient if the state times $t_{j,k}$ correspond to full hours to simplify the interpolation of the metocean conditions as it reduces the order of the interpolation from tri-linear to bi-linear as the state time would match that of the hindcast data (see Section 3.2). Because the resolution of the hindcast database is $1h$, all the waypoints are collocated as,

$$t_{j+1,k} = \lfloor t_{j,k} + 1h \rfloor \tag{2.5}$$

This simplification introduces the challenge of identifying the position the ship is in $\mathcal{R}$ becomes more complicated. The problem of finding the sailing time between two adjacent waypoints $(x_{j,k}, y_{j,k})$ and $(x_{j,k+1}, y_{j,k+1})$ is straightforward, as the Haversine distance directly gives the distance,

$$\Delta d_{j,k} = 2 \cdot r \cdot \arcsin\left(\sqrt{\left(\sin^2\left(\frac{y_{j,k+1} - y_{j,k}}{2}\right) + \cos(y_{j,k}) \cdot \cos(y_{j,k+1}) \cdot \sin^2\left(\frac{x_{j,k+1} - x_{j,k}}{2}\right)\right)}\right) \tag{2.6}$$

and the sailing time,

$$\begin{aligned} \Delta t_{j,k} &= \frac{\Delta d_{j,k}}{V_{j,k}} \\ t_{j,k+1} &= t_{j,k} + \Delta t_{j,k} \end{aligned} \tag{2.7}$$

The inverse problem of finding the position $(x_{j,k+1}, y_{j,k+1})$ after the ship has sailed with speed $V_{j,k}$ with duration $\Delta t_{j,k}$ is more difficult because it requires a function that follows the curvature of $\mathcal{R}$, that has inputs $(x_{j,k}, y_{j,k})$ and $\Delta d_{j,k}$ and returns the new position. Such a function for a general $\mathcal{R}$ does not exist in a closed analytical form. The problem is approximated using a linear parametrisation of a finely discretised version of $\mathcal{R}$ to address this. This approach replaces the continuous curve with a series of discrete waypoints, where the distances between consecutive waypoints are approximated using the Euclidean distance. The solution involves approximating a starting point $(\hat{x}_{j,k}, \hat{y}_{j,k})$ within the discrete sequence and iteratively accumulating distances along the path until the ship has covered at least a total distance of $\Delta d_{j,k}$. Adjusting the resulting position $(\hat{x}_{j,k+1}, \hat{y}_{j,k+1})$ is achieved by using a correction factor.

2.1.2 Objective and constraint functions

The objective function from Equation 2.2 is also simplified by the zero-order hold of Equation 2.4. The discretised objective function is given by,

$$M_{fuel} = \sum_{j=1}^{n} f_j \left(P_j, W_{j,1:m_j}\right) = \sum_{j=1}^{n_j} \sum_{k=1}^{m_j-1} m_{\text{fuel}} \left(P_j, W_{j,k}\right) \cdot \Delta t_{j,k}, \tag{2.8}$$

where n is the total number of legs the $\mathcal{R}$ was segmented, m_j denotes the total number of waypoints within the j-th leg, and $W_{j,k}$ are the metocean conditions at state $\mathbf{S}_{j,k}$ given by,

$$W_{j,k} = \left[H_{s(j,k)}, \alpha_{H_s(j,k)}, T_{z(j,k)}, V_{\text{wind}(j,k)}, \alpha_{\text{wind}(j,k)}, V_{\text{current}(j,k)}, \alpha_{\text{current}(j,k)}\right] \tag{2.9}$$

The optimisation constraints that ensure the power setting P_j and speed V are within operational limits, and the total travel time aligns with the ETA are defined as,

$$C(P, W) = \begin{cases} P_{\text{min}} \leq P_j \leq P_{\text{max}}, \\ V_{\text{min}} \leq V(P_j, W_{j,k}) \leq V_{\text{max}}, \\ 0.99 \leq \dfrac{\sum_{j=1}^{n} \sum_{k=1}^{m_j-1} \Delta t_{j,k}}{\text{ETA}} \leq 1.01. \end{cases} \tag{2.10}$$

The power allocation $\mathbf{P}$ represents a set of valid engine power allocation values for the voyage,

$$\mathbf{P} = [P_1, P_2, \ldots, P_n], \tag{2.11}$$

2.1.3 Dynamic programming with parallel scenarios

To optimise $\mathbf{P}$, the cost function representing the fuel consumption for the j-th Equation 2.8 needs to be minimum. The problem is efficiently determining each P_j so that the fuel consumption of each leg M_j is minimum and the ship still arrives in time.

$$M_j \left(P_j, W_{j,1:m_j}\right) = \sum_{k=1}^{m_j-1} m_{\text{fuel}} \left(P_j, W_{j,k}\right) \cdot \Delta t_{j,k}, \tag{2.12}$$

The values P_j that minimise Equation 2.8 are determined using the modified description of the voyage so that each leg becomes decoupled. Figure 2.2 illustrates the traditional exact coupling, where each leg is sequentially coupled by connecting each leg's last waypoint with the next's first waypoint. However, this exact coupling requires sequential computation of each leg's start time based on the end of the previous leg, limiting computational efficiency. To address this, a parallel scenario-based

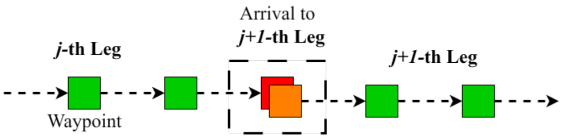


Figure 2.2: Exact coupling between consecutive legs.

DP approach is proposed, allowing each leg to be optimised independently within a specified time interval, as shown in Fig. 2.3.

For each leg $j+1$, the nominal departure time $t^{(0)}_{j+1,1}$ is calculated as:

$$t^{(0)}_{j+1,1} = \frac{d_{1\to j}}{V_{\text{average}}}, \tag{2.13}$$

Where $d_{1\to j}$ is the cumulative distance up to the end of the j-th leg, and V_{average} is the average speed needed to meet the ETA. For each leg, a set of parallel scenarios with different departure times is generated within this interval:

$$t_{j+1,1} \in \left[t^{(0)}_{j+1,1} - \frac{\Delta T_{j+1}}{2},\ t^{(0)}_{j+1,1} + \frac{\Delta T_{j+1}}{2}\right], \tag{2.14}$$

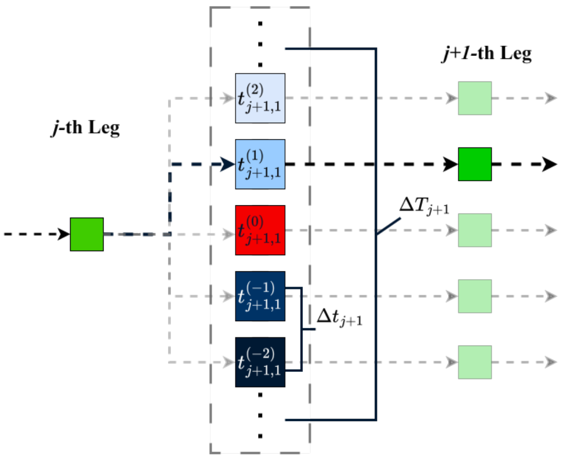


Figure 2.3: Parallel scenarios for optimising power allocation across voyage legs.

Where ΔT_{j+1} is the allowable time window for leg $j+1$. Each scenario explores various power settings from a discrete set $\mathbf{\Gamma} = [\Gamma_1, \Gamma_2, \ldots, \Gamma_r]$, covering the range $P_{min} \leq P \leq P_{max}$. Only scenarios that fit within the time interval defined are considered valid, and each scenario is paired with all possible power settings to simulate feasible sub-voyages, as shown in Fig. 2.3. The DP optimisation minimises fuel consumption across all valid scenarios using the recursive Bellman equation:

$$F_j = \min_{P_j} \left(M_j(P_j, W_{j,1:m_j}) + F_{j-1} \right), \tag{2.15}$$

Where F_j is the cumulative fuel consumption up to leg j, and $M_j(P_j, W_{j,1:m_j})$ represents the fuel consumption of leg j under power setting P_j and metocean conditions $W_{j,1:m_j}$.

2.2 Voyage segmentation

In Section 2.1.1 $\mathcal{R}$ was assumed to be divided in n legs. Segmenting the voyage is beneficial because solving for optimal instantaneous values of $P(t)$ is impossible. In this section, two data-driven methods for segmentation are compared to generate legs based on expected environmental conditions along $\mathcal{R}$ while sailing between times t_0 and t_f and other operational requirements.

Metocean conditions significantly impact a vessel's fuel consumption and sailing speed. Standard sailing strategies voluntarily reduce V in adverse metocean conditions to catch up with sailing. The objective is to determine the optimal segmentation based on metocean conditions and then use the DP solver to guide the navigation. Therefore, this research proposes a segmentation approach named the metocean score-based pruned exact linear time (MS-PELT) algorithm, which determines the optimal number of legs. For benchmarking purposes, the MS-PELT algorithm is compared with a multivariate time series clustering method known as toeplitz inverse covariance-based clustering (TICC).

2.2.1 MS-PELT algorithm

The MS-PELT algorithm comprises four steps, as illustrated in Fig. 2.4. In the first step, the target average speed, V_{average}, needed to meet the expected time of arrival (ETA) based on the voyage distance, is calculated. Next, Monte Carlo simulations generate multiple reference voyages featuring different speed profiles that all meet the average speed requirement within a specified speed range $[V_{\min}, V_{\max}]$.
For each reference voyage, the metocean conditions along $\mathcal{R}$ are obtained at each waypoint, including wind, wave, and current characteristics. A metocean score is computed for each waypoint, capturing the combined effects of environmental factors. The metocean score at the k-th waypoint for the q-th reference voyage is defined as:

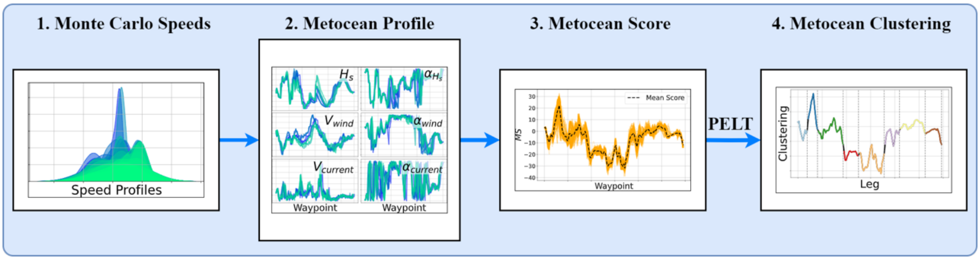


Figure 2.4: Workflow of the MS-PELT Voyage Segmentation Method.

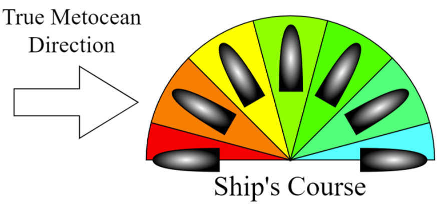
$$\begin{aligned}\mathrm{MS}_{q,k} = &\ \beta\left(\alpha_{H_s(q,k)}\right)\cdot\iota\left(H_{s(q,k)}\right) + \\ &\ \beta\left(\alpha_{\mathrm{wind}(q,k)}\right)\cdot\iota\left(V_{\mathrm{wind}(q,k)}\right) + \\ &\ \beta\left(\alpha_{\mathrm{current}(q,k)}\right)\cdot\iota\left(V_{\mathrm{current}(q,k)}\right), \end{aligned} \tag{2.16}$$

Where H_s is the significant wave height, V_{wind} and V_{current} are the wind and current speeds, and α_{H_s}, α_{wind}, and $\alpha_{\mathrm{current}}$ are the respective direction factors. Each component reflects the relative impact of each metocean variable at the waypoint. The functions $\beta(\cdot)$ and $\iota(\cdot)$ correspond to the direction score (Fig. 2.5) and the intensity score (Fig. 2.6), respectively.

The overall metocean score for waypoint k is obtained by averaging the metocean scores across all reference voyages:

$$\mathrm{MS}_k = \frac{1}{Q}\sum_{q=1}^{Q}\mathrm{MS}_{q,k}, \tag{2.17}$$

where Q is the total number of reference voyages. The MS-PELT algorithm then



Metocean direction (°)	β_i
165-180	-5
135-165	-3
105-135	-1
75-105	0
45-75	1
15-45	3
0-15	5

Figure 2.5: Metocean Direction Score $\beta(\alpha)$.

segments $\mathcal{R}$ by identifying a feasible sequence segmentation $\boldsymbol{\tau}$ based on the metocean scores, optimising the balance between fitting accuracy and complexity using the following objective function:

$$\boldsymbol{\tau} = \arg\min_{\boldsymbol{\tau}} \left(\sum_{j=1}^{b+1} \phi\left(\mathrm{MS}_{\tau_{j-1}+1:\tau_j}\right) + \gamma b \right), \tag{2.18}$$

Where ϕ represents the cost function of each segment, γ is a penalty factor to avoid overfitting, b is the number of change points.

V_{wind} (m/s)	0 - 0.2	0.2 - 1.5	1.5 - 3.3	3.3 - 5.4	5.4 - 7.9	7.9 - 10.7	10.7 - 13.8	13.8 - 17.1	17.1 - 20.7	20.7 - 24.4	24.4 - 28.4	28.4 - 32.6	32.6 +
Beaufort scale (ι_{wind})	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
H_s (m)	0 - 0.1	0.1 - 0.5	0.5 - 1.25	1.25 - 2.5	2.5 - 4	4 - 6	6 - 9	9 - 14	14+				
Douglas scale (ι_{H_s})	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**				
$V_{current}$ (m/s)	0 - 0.1	0.1 - 0.2	0.2 - 0.5	0.5 - 1	1 - 1.5	1.5 +							
Current scale ($\iota_{current}$)	**0**	**2**	**4**	**6**	**8**	**10**							

Figure 2.6: Metocean Intensity Score $\iota(H_s)$.

2.2.2 TICC algorithm

The TICC algorithm (Hallac, Vare, et al., 2017) is used as a comparison method for voyage segmentation. Unlike MS-PELT, which scores metocean conditions based on ensemble values, TICC directly clusters multivariate time series subsequences of metocean conditions. For each waypoint, TICC generates an average matrix of metocean variables overall reference voyages:

$$\mathbf{W} = \begin{bmatrix} \overline{H_s}(1) & \overline{H_s}(2) & \dots & \overline{H_s}(a) \\ \overline{\alpha_{H_s}}(1) & \overline{\alpha_{H_s}}(2) & \dots & \overline{\alpha_{H_s}}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\alpha_{\mathrm{current}}}(1) & \overline{\alpha_{\mathrm{current}}}(2) & \dots & \overline{\alpha_{\mathrm{current}}}(a) \end{bmatrix}, \tag{2.19}$$

Where a is the number of waypoints along $\mathcal{R}$. The time series data are divided into fixed-length subsequences, each represented by a matrix of size $s \times l$, where s is the number of metocean variables and l is the length of the subsequence.

Instead of clustering individual waypoints, TICC clusters subsequences to account for temporal dependencies. Each cluster is associated with an inverse covariance matrix forming a Markov random field that encodes structural patterns across segments. TICC aims to find optimal segment assignments by solving the following:

$$\boldsymbol{\Theta}, \boldsymbol{\sigma} = \arg\min_{\boldsymbol{\Theta},\boldsymbol{\sigma}} \sum_{j=1}^{n} \left(\lambda \|\theta_j\|_1 + \sum_{\mathbf{W}_k \in \sigma_j} \big(-\log L(\mathbf{W}_k; \theta_j) + \Omega \mathbf{1}\left[\sigma_{k-1} \neq \sigma_j\right] \big) \right), \tag{2.20}$$

Where $\boldsymbol{\Theta} = \{\theta_1, \ldots, \theta_n\}$ are the inverse covariance matrices for each cluster, $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_n\}$ denotes the segment assignments, λ and Ω control sparsity and temporal consistency, $L(\mathbf{W}_k; \theta_j)$ is the likelihood function, and $\mathbf{1}[\cdot]$ is the indicator function.

2.3 Optimisation of the double-ended ferry

Double-ended ferries present unique optimisation challenges due to their dual-engine configuration, which requires simultaneously determining the optimum of both engine powers. This setup requires separate throttle controls for each engine, typically labelled as the "bow" and "stern" engines, based on their position relative to the vessel's current heading. Managing the double-ended ferry's fuel consumption is more complex as both engines must be adjusted to ensure efficient operation, regardless of the vessel's direction. Unlike conventional vessels, where a single propulsion direction simplifies fuel management, double-ended ferries must dynamically switch engine roles depending on the travel direction, adding additional complexity to fuel optimisation.

The optimisation approach for the double-ended ferry is based on a decision support system (DSS) framework, as shown in Fig. 2.7. The DSS framework utilises prior, current, and hindcast data information specific to each trip, incorporating the following essential inputs:

- **$\mathcal{R}$:** Provides historical sailing waypoints, including longitudes, latitudes, heading angles, and similar trip characteristics, such as speed V_g and fuel consumption.

- **MetOcean data:** Obtained from the weather forecasting database. Once interpolated to match the ship's $\mathcal{R}$ and schedule, the data ensures that metocean conditions are accurately represented.

- **Initial operational guess:** A "pre-assumed" optimal operation parameter set that directs the optimiser towards a search space containing likely stationary points, enhancing the possibility of reducing fuel consumption.

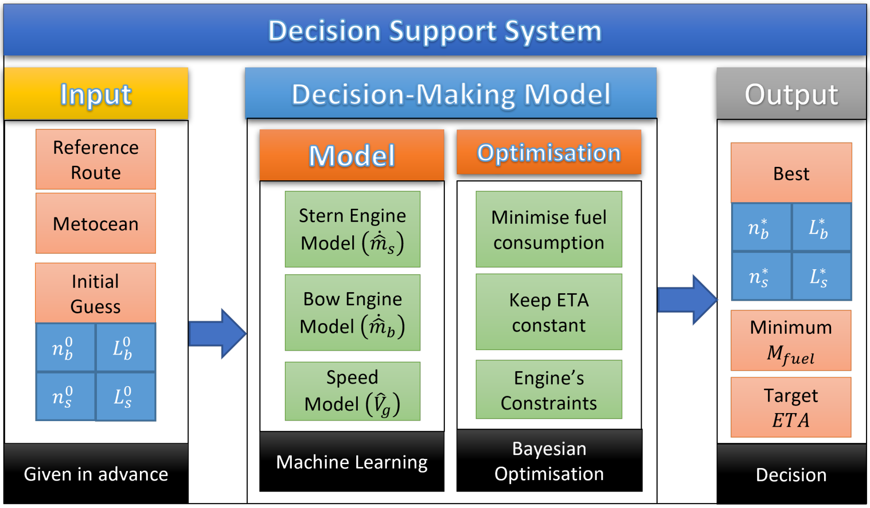


Figure 2.7: Decision Support System (DSS) workflow for optimising power allocation used in Paper II.

Given that the ferry operates on very short routes, typically without significant variations in metocean conditions over a single trip, segmentation into multiple legs is not employed. Instead, the entire journey is treated as a single leg, simplifying the optimisation step to determine a single engine parameter for each engine and the load distribution. Note that this simplifies the objective function of the double-ended ferry voyage to just Equation 2.8.

2.3.1 Power ratio

To evaluate and control power distribution between the bow and stern engines, the power allocation is quantified using a parameter referred to as the power ratio, R_p. This ratio, defined as follows, reflects the proportion of power allocated to the stern engine relative to the total power:

$$R_p = \frac{P_{\text{stern}}}{P_{\text{bow}} + P_{\text{stern}}} \tag{2.21}$$

Where P_{stern} and P_{bow} represent the engine power at the stern and bow engines, respectively. The power ratio R_p ranges from 0 to 1, where $R_p = 0$ indicates full power allocation to the bow engine, and $R_p = 1$ signifies full power allocation to the stern engine. The mean power over timer ratio $\overline{R_p}$ for each trip is given by:

$$\overline{R_p} = \frac{1}{n} \sum_{i=1}^{n} R_p^{(i)} \tag{2.22}$$

Where n denotes the total number of measurements taken within a trip, and $R_p^{(i)}$ is the instantaneous power ratio recorded at each sampling point. The dominant hypothesis explored was that a higher $\overline{R_p}$ corresponded to a decrease in fuel consumption.

2.3.2 Bayesian optimisation

To validate the power ratio hypothesis, the DSS optimisation framework aimed to identify optimal values for the power ratio R_p that minimise the fuel consumption for the journey while maintaining the ETA constraints. Bayesian optimisation is employed to solve this problem. Bayesian optimisation (BO) is a probabilistic model-based optimisation technique primarily used for optimising expensive-to-evaluate functions. This method is useful when the function evaluation is hard or time-consuming. Bayesian optimisation balances exploration and exploitation of the search space by performing the optimisation on a surrogate model and deciding the next points to evaluate through an acquisition function.

Given an objective function $f : \mathcal{X} \rightarrow \mathbb{R}$ where $\mathcal{X}$ is the input space, bayesian optimisation seeks to maximise f by iteratively selecting points in $\mathcal{X}$ to evaluate. It builds a probabilistic model of f and uses this model to decide where to consider next, focusing on areas of high uncertainty or high expected improvement.

$$f(\xi) \sim \mathcal{GP}(\mu(\xi), k(\xi_i, \xi_j)) \tag{2.23}$$

Based on previous evaluations, the GP model used in Bayesian optimisation estimates the objective function. It is defined by a covariance function, $\Sigma_{i,j}$, to represent correlations between sampled data points ξ_i and ξ_j in the search space:

$$\Sigma_{i,j} = k(\xi_i, \xi_j) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{\|\xi_i - \xi_j\|}{\rho} \right)^{\nu} K_\nu \left(\sqrt{2\nu} \frac{\|\xi_i - \xi_j\|}{\rho} \right) \tag{2.24}$$

The GP model iteratively estimates the mean $\mu(\xi_{k+1})$ and variance $\sigma^2(\xi_{k+1})$ of the fuel consumption function, facilitating exploration and exploitation through an acquisition function. The negative lower confidence bound (NLCB) acquisition function is utilised to guide the selection of subsequent parameter configurations:

$$\xi_{k+1}^* = \arg\max_{\xi_{k+1}} u(\xi_{k+1}) = \arg\max_{\xi_{k+1}} \left(\mu(\xi_{k+1}) - \beta \cdot \sigma(\xi_{k+1})\right) \tag{2.25}$$

Where β controls the balance between exploration and exploitation, encouraging sampling in regions of high uncertainty and potential improvement in fuel efficiency. The BO optimisation framework follows an iterative approach:

1. Select initial points and evaluate the objective function at these points.

2. Fit a GP surrogate model to the data.
3. Use the acquisition function to find the next point to evaluate.
4. Evaluate the objective function at this new point and update the data.
5. Update the surrogate model with the new data and repeat the process.

Through this DSS framework, the optimal power allocation between the bow and stern engines is determined for each trip.

CHAPTER 3

Full-scale measurements for data–driven modelling

While the previous chapter described the optimisation of M_{fuel}, it deliberately omitted the models required for m_{fuel} and V, required to evaluate the objective function (Equation 2.8) under the different $\mathbf{P}$. Due to the availability of high-frequency data for two ships, a data-driven modelling approach is adopted in this thesis, and machine learning regression models are trained on full-scale sensor data. These ML models establish predictive Input/Output relationships for ship performance as shown in Figure 3.1.

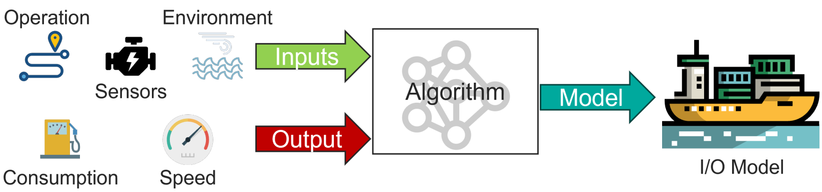


Figure 3.1: Machine Learning I/O models.

For model training, two datasets are used: The first dataset, collected from a double-ended ferry, provides the basis for models used in Papers I and II. The second dataset, based on a chemical tanker, is used in Paper III. The structure of this section is organised as follows: first, the datasets used in each paper are introduced; second, the methodology for data processing is presented; and finally, the machine learning algorithm, along with the hyperparameter tuning process, is described.

3.1 Full-scale ship measurements

The dataset for papers I and II comes from one year of operational data of *M/S Uraniborg*, a RoRo passenger ferry operating between Ven and Landskrona, Sweden. Key specifications for *M/S Uraniborg* are summarised in Table 3.1. The ferry's on-board energy management system, provided by BlueFlow (Ventrafiken, 2021), records a variety of operational parameters, including engine load, engine power, speed over ground (V), propeller speed (RPM_{prop}), and fuel consumption rate (m_{fuel}).

Table 3.1: Specifications and Operational Data of *M/S Uraniborg* (Papers I and II)

Parameter	**Value**
Length	49.95 m
Beam	12.00 m
Draught	2.85 m
Max Power	709 kW (each engine)
Engine Rotation Speed	1600 rpm
Service Speed	11.5 knots
Data Collection Period	Jan 2021 - Jan 2022
Data Frequency	1 sample per minute

The dataset for paper III, in contrast, is derived from three years of operational data of a chemical tanker sailing between ports in European waters across the Baltic Sea, North Sea and the English Channel. Key specifications for the ship are provided in Table 3.2.

Table 3.2: Specifications and Operational Data of the Chemical Tanker (Paper III)

Parameter	**Value**
Length Between Perpendiculars	138.22 m
Breadth	23.76 m
Design Draft	9.27 m
Displacement	25174 m^3
Maximum Continuous Rating (MCR)	7200 kW
Service Speed	14 knots
Data Collection Period	Nov 2020 - Mar 2024
Data Frequency	1 sample per minute

The full-scale data from the vessels was downsampled from 1-minute to 10-minute intervals to reduce noise while preserving relevant trends. Then, steady-state filters based on first derivative thresholds were applied to the ship's engine power, speed, and speed, effectively removing transients and manoeuvres. These signals were further smoothed using a second-order Savitzky-Golay filter.

3.2 Feature engineering

Feature engineering involves selecting, transforming, and creating new input features from raw data to train machine learning models. This thesis applies domain knowledge to choose the relevant signals from the measurement data. In this context, domain knowledge refers to understanding naval architecture, aerodynamics, and the physical principles governing how ships move through water. The models of interest are for speed prediction V, which is necessary for estimating the ETA and the fuel consumption rate m_{fuel}. The objective is to identify features χ_i so that,

$$V = V(\chi_{fuel}) \tag{3.1}$$

$$m_{fuel} = m_{fuel}(\chi_{fuel}) \tag{3.2}$$

From principles of naval architecture, it is known that internal combustion engines (ICE) combined with screw propellers are the most commonly used propulsion systems in ships (Latarche, 2021). ICE systems are preferred over alternative propulsion methods due to their efficiency and suitability for long-distance operations (Farnsworth, 2022). This propulsion setup typically consists of a main engine and propeller arrangement designed to provide thrust and propel forward the vessel. The thrust power, P_T, represents the power transferred from the propeller to the surrounding water and is defined as,

$$P_T = T \cdot V \propto P_{eng} \tag{3.3}$$

This relation indicates that speed is directly influenced by the engine power P. Similarly, the fuel consumption, m_{fuel}, required to produce the necessary thrust power is given by:

$$m_{\text{fuel}} = \frac{\text{SFOC} \cdot P_T}{\eta_0 \cdot \eta_R \cdot \eta_s} \tag{3.4}$$

Where η_0, η_R, and η_s are the efficiencies for the open propeller, relative rotation, and shaft, respectively, while SFOC stands for specific fuel oil consumption. Additionally, it is known that power is proportional to engine speed (RPM),

$$P_T \propto f(RPM_{eng}) \tag{3.5}$$

$$P_T \propto f(RPM_{eng}, \varphi_{eng}) \tag{3.6}$$

The physical quantities presented in these equations are compared to the available measurements, with P_{eng}, RPM_{eng}, and V representing significant features for regression modelling. Other operational parameters, such as the ship draught T_m, are also included, providing the model with information about the loading condition.

It is also well established that metocean conditions significantly affect V and m_{fuel}

Lang et al., 2022; H. Wang, 2020. Metocean data are extracted from multiple sources: wind speed (V_{wind}), wind direction (φ_{wind}), significant wave height (H_s), wave direction (φ_{wave}), and wave period (T_s) are extracted from the ERA5 reanalysis dataset, while sea current speed ($V_{current}$), and sea current direction ($\varphi_{current}$) are obtained from the Global Ocean Physics Analysis and Forecast. These datasets have spatial resolutions of 0.5° × 0.5° and $1h$ temporal resolution (Copernicus Climate Change Service (C3S), 2023; Hersbach et al., 2020), and 0.083° × 0.083° with 30m resolution (Mercator Océan International, 2024) with trilinear interpolation used for data integration.

3.3 Data-driven ship performance modelling

With available data, ML algorithms were used to map models for V and m_{fuel}. The XGBoost regression algorithm was selected in this study due to its effectiveness in modelling complex, non-linear relationships (Lang et al., 2022).

$$m_{\text{fuel}} = f_{\text{fuel}}(P, V, \text{RPM}, T_{\text{mean}}, H_s, \alpha_{H_s}, T_z, V_{\text{wind}}, \alpha_{\text{wind}}, V_{\text{current}}, \alpha_{\text{current}}) \tag{3.7}$$

$$V = f_V(P, \text{RPM}, T_{\text{mean}}, H_s, \alpha_{H_s}, T_z, V_{\text{wind}}, \alpha_{\text{wind}}, V_{\text{current}}, \alpha_{\text{current}}) \tag{3.8}$$

To achieve the best performance in the prediction models, careful tuning of the model hyperparameters is required. As mentioned before, Bayesian optimisation is a common tool used to determine the optimal combination of the parameters (Lang et al., 2022). In the case of XGBoost, the list of hyperparameters to tune is shown in Table 3.3. The scope of the hyperparameter tuning is to minimise the prediction error (Mean Squared Error – MSE) given by,

$$\text{MSE} = \sum(\hat{f}_i - f_i)^2 \tag{3.9}$$

Where f_i is the prediction objective.

Table 3.3: List of Hyperparameters and Search Bounds

Hyperparameter	**Bounds**	**Hyperparameter**	**Bounds**
Number of estimators	<20000	Min Child Weight	1-5
Learning Rate	0.01-0.3	Gamma	0-5
Max Depth	5-15	Regularisation Alpha	0-0.5
Sub Sample	0.8-1	Regularisation Lambda	0.1-5
Column Sample by Tree	0.7-1		

CHAPTER 4

Summary of appended papers

The research led to three papers corresponding to two case study ships. Common factors between all research publications and an overview of the research methods, questions, and results are shown in Figure 4.1.

	Ship Type	Data Volumen	Machine Learning	Power Optimisation	Energy Efficiency
Paper I	Double-Ended Ferry	1 Year.	Multiple Algorithms		35% Potential Savings*
Paper II	Double-Ended Ferry	1 Year.	XG Boost	Bayesian Optimisation	15% Actual Savings
Paper III	Chemical Tanker	3 Years.	XG Boost PELT TICC	Dynamic Programming	15% Potential Savings

Figure 4.1: Relationships between different research subjects.

The rest of this section provides a concise summary of the key findings from each paper. First, the scope and contributions of each paper are introduced. Next, a brief overview of the methods used is presented, followed by the main results and discussion.

4.1 Paper I

4.1.1 Summary of Paper I

This first paper describes how data analytics and ML regression reduce fuel consumption in a double-ended ferry. It investigates the most efficient loading conditions for the ferry's engines from a data analytics point of view to reduce fuel consumption. A new variable, R_P, is introduced to describe the loading condition of the vessel stern's engine with respect to the combined engine utilisation. It was discovered that a high R_P had a high potential for fuel consumption reduction with up to 35% fuel savings achievable with changes in the operation.

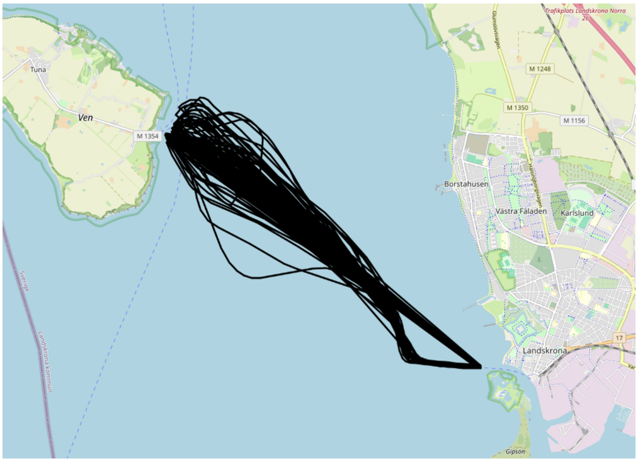


Figure 4.2: Sailing region of the case study double-ended ferry.

4.1.2 Scope and contribution

This study analyses the operational data of a case study of a double-ended ferry to determine insights for potential improvements in its operations in light of reducing fuel consumption. Double-ended ferries are a sustainable alternative to bridges or tunnels in congested urban and coastal environments. These vessels perform very short voyages multiple times a day. Therefore, any savings achieved in their operation are amplified by the frequency of their trips. Figure 4.2 shows the sailing region and typical routes of the case study ferry. Therefore, The pipeline for this paper is as follows:

- **Exploratory Data Analysis** is used to determine operational trends from sensor and environmental data to determine patterns in fuel consumption. In particular, the relative distribution of engine power allocation (between stern and bow) is called the Power Ratio (R_P).

- **Regression analysis** is used to quantify the impact of R_P on the total fuel consumption of the ferry.

- **Machine learning (ML)** is used to model the precise energy performance of the ferry based on operational conditions, metocean data and R_P.

4.1.3 Results and Discussion

As discussed in Section 2.3.1 R_P describes the contribution of each engine to the total propulsive power. The data was processed to determine the average $\overline{R}_P$ and the total fuel consumption for every trip. Then, exploratory data analysis revealed a consistent trend between fuel consumption and the power ratio R_P for each trip. Then, simple regression analysis showed that the power ratio clearly influenced total fuel consumption, as observed in Figure 4.3. The trendlines for both sailing directions correspond to the following equations,

$$M_{\text{west}} = 83.87 - 53.834 \cdot R_P \tag{4.1}$$

$$M_{\text{east}} = 84.455 - 53.888 \cdot R_P \tag{4.2}$$

These linear equations have their minimum at $R_P = 1$, suggesting that allocating more power to the stern engine results in lower fuel consumption.
The problem with simple linear regression analysis is that it does not capture the effects of metocean conditions and operational parameters on fuel consumption. Therefore, a more detailed regression analysis was introduced to describe the influence of R_P on fuel consumption. Among the various machine learning models tested, XG-Boost had the best performance (see Fig. 4.4) with an R^2 coefficient of 0.964.

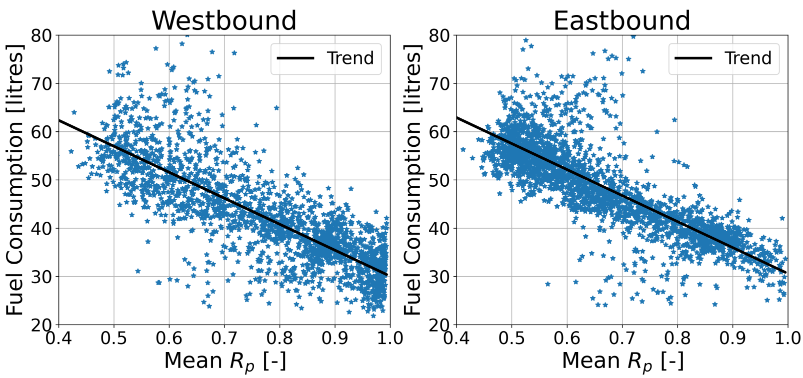


Figure 4.3: Fuel Consumption vs Power Ratio for each trip and direction in 2020.

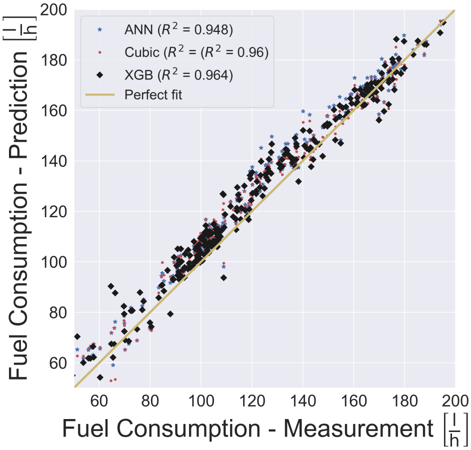
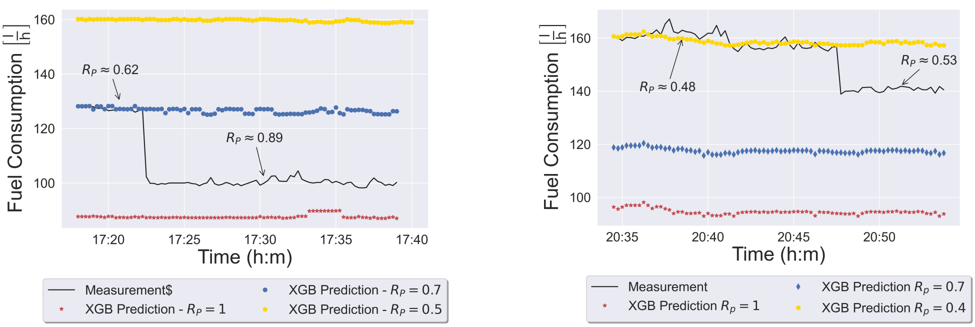


Figure 4.4: Performance of ML Algorithms.



(a) R_P Scenarios for an Eastbound Trip

(b) R_P Scenarios for a Westbound Trip

Figure 4.5: Comparison of the power allocation impact on two trips

Using the XGBoost model, different power allocation scenarios of R_P were simulated by manually setting the value R_p as shown in Figures 4.5a and 4.5b. The simulations confirmed that higher R_P resulted in substantial fuel savings, consistent with the linear regression analysis. In particular, a power ratio of $R_P = 1$ leads to fuel savings of up to 18% and 35% on eastbound and westbound trips, respectively, as summarised in Table 4.1.

Table 4.1: Summary of the Selected Simulation Results

Direction	Measured M_{fuel} (litres)	Simulated M_{fuel} (litres)	Savings
Westbound	54.90	34.72	-35.8%
Eastbound	41.36	35.5	-18.45%

4.2 Paper II

4.2.1 Summary of Paper II

In Paper II, sensor data from the double-ended ferry is used to improve its operation through a DSS. A Bayesian Optimisation-based DSS is proposed to help determine the operational set-points of the ferry's engines in terms of R_P and engine speed while simultaneously satisfying an ETA constraint for the voyage. Through optimisation, the DSS can reduce the ship's fuel consumption by up to 40 % with no significant change in the ETA. These results were validated through full-scale experiments. The experiments demonstrated an average of 15 % fuel consumption reduction when the ship operates at a high R_P, confirming the efficacy of the DSS.

4.2.2 Scope and contribution

In paper I, the impact of R_P on the fuel consumption was analysed from a full voyage-to-voyage perspective, an approach that neither considers nuances specific to each voyage nor considers the waypoint-to-waypoint evolution of the fuel consumption as the ship sails. Therefore, this paper explores strategies to improve the double-ended ferry's optimal operational efficiency. It is accomplished through:

- **Machine Learning Modelling**. As the energy performance of double-ended ferries is rarely studied, Machine Learning Models (XGBoost) for ship speed and fuel consumption are trained from the available data.

- **Decision Support System (DSS)**. A Bayesian optimisation–based (BO) DSS was designed to determine the optimal R_P for each voyage.

- **Full Scale Experiments**. Experiments using a high R_P were proposed to and performed by the shipping company.

4.2.3 Results and Discussion

The DSS determines the optimal power ratio R_P and engine rotation speeds n_b and n_s that minimised the fuel consumption. In general, results agree that allocating more power to the stern engine reduced fuel consumption. The objective function was the total fuel consumption of the trip (M_{fuel} that was determined through the simulation of the voyage. XGBoost models were used for the fuel consumption rate of each engine and the ship's speed overground (Fig. 4.6).

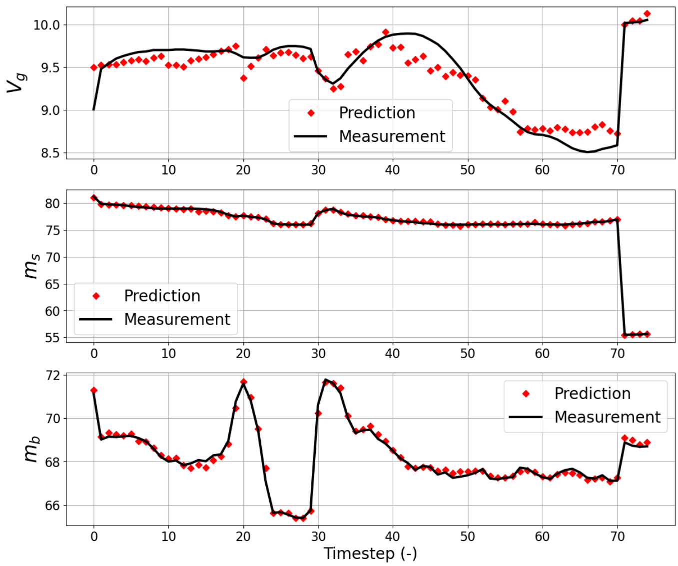


Figure 4.6: XGBoost model forecast for Trip W1.

BO was then used to optimise power allocation, achieving up to a 43% reduction in fuel consumption without affecting the ship's sailing time (the ETA). The optimised results, presented in Fig. 4.7, showed less than a 2% in sailing time deviation from the actual measurements (<1 min). Table 4.2 provides a comprehensive summary of the simulated voyages. The optimal set points n_b, L_b, n_s, and L_s are illustrated in Fig. 4.8. It can be noted that for the models used, the optimum R_P did not always correspond to 1 as observed in Table 4.2.

To test the DSS, full-scale tests were conducted between 19 and 22 August 2022 on board the ferry to examine the impact of power allocation strategies on fuel consumption. The Captain and 1st Mate alternated operating the ferry every other round trip to ensure comparable conditions. The Captain was instructed to maintain power predominantly on the stern thruster. At the same time, the 1st Mate, unaware of the experiment, operated as normal.

Round trips were timed at 1–2-hour intervals to minimise variations due to weather and traffic. Data analysis involved filtering out trips with irregular speeds, leaving 10 valid trips per operator for the comparison (Fig. 4.9). Reference data of the ferry operation from the same crew before the experiment was used for a formal comparison. It confirmed that fuel consumption was reduced under the improved power allocation.

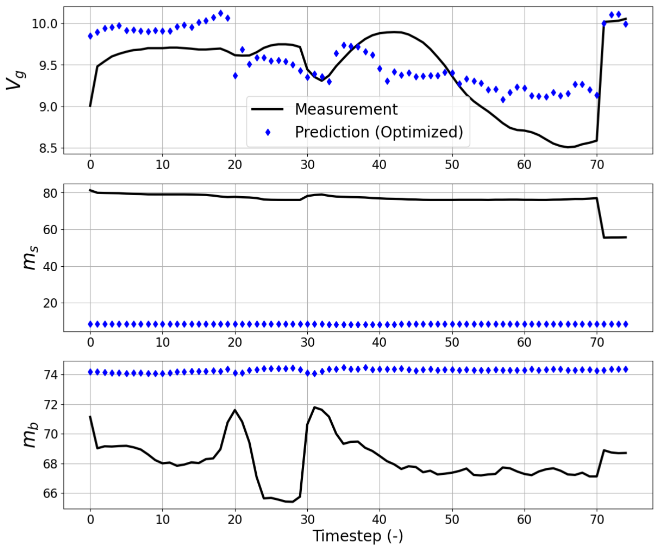


Figure 4.7: Optimised Operation for Trip W1.

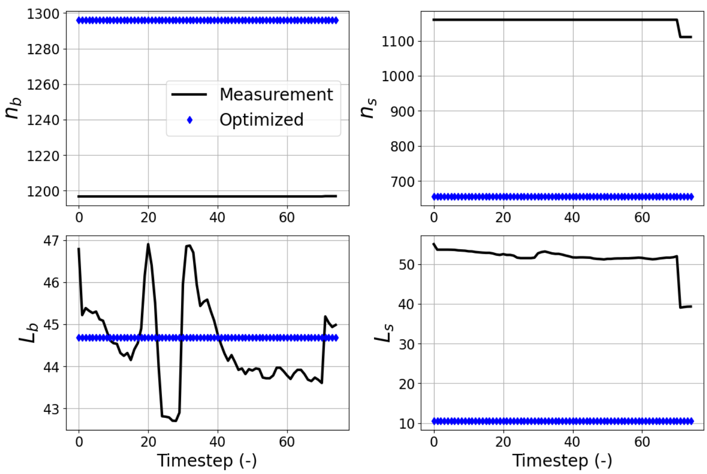


Figure 4.8: Optimised Operation Setpoints for Trip W1.

The experiment results are summarised in Table 4.3. The analysis showed an 18% reduction in fuel consumption for the Captain with respect to the operation by the first Mate, assumed to be the regular operation. For the reference period, the fuel consumption was reduced by 15% for the Captain and 3.5% for the First Mate, illustrating the advantages of high stern power allocation on the operation's efficiency.

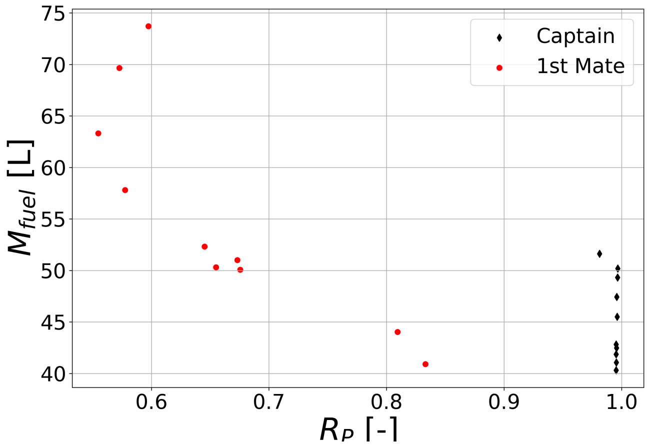


Figure 4.9: Observed trend in the full-scale experiments.

Table 4.2: Verification of DSS for historical trips based on performance simulations.

Direction	Case.	M_{fuel} $[l]$	R_p	M^*_{fuel} $[l]$	R'_p	V_g [%]	M^*_{fuel} [%]
Westbound	W1	45.12	0.472	25.45	0.898	+1.44 %	-43.59 %
Westbound	W2	40.08	0.549	25.21	0.947	+0.40 %	-37.10 %
Westbound	W3	39.78	0.480	35.62	0.523	+0.07 %	-10.46 %
Eastbound	E1	33.48	0.565	28.47	0.684	+1.39 %	-14.96 %
Eastbound	E2	31.93	0.649	23.11	0.958	+1.8 %	-27.62 %
Eastbound	E3	29.11	0.535	18.73	0.624	+0.44 %	-35.66 %

Table 4.3: Summary of full-scale test results.

Operator	Period	N-Trips	V_g $[kt]$	R_p	M_{fuel} $[l]$	% Change w.r.t. Operator	% Change w.r.t. Reference
1st Mate	Reference	10	9.3	0.7	57	-	-
Captain	Reference	10	9.3	0.8	53	-7%	-
1st Mate	Experiment	10	9.4	0.7	55	-	-3.5%
Captain	Experiment	10	9.5	1.0	45	-18%	-15%

4.3 Paper III

4.3.1 Summary of Paper III

This study introduces a two-stage methodology for optimising a fixed route voyage and, unlike Papers I and II, for a more conventional ship (a chemical tanker) sailing across European waters. For the vessel, three years of sailing routes are presented in Figure 4.10. The first stage consists of the optimal route segmentation through the MS-PELT algorithm to reduce computational overhead due to excessive segmentation. Then, a parallel scenario of Dynamic Programming is used to find the optimal power allocation for each leg of the voyage. The ship's energy efficiency is directly optimised compared to the unoptimised operation.

4.3.2 Scope and contribution

Speed optimisation is the most common operation optimisation approach proposed in the literature. It has the caveats of subjecting the ship's engine to a fluctuating load to sustain the speed levels in varying metocean conditions. For that reason, the scope of Paper III was to improve the energy efficiency of a ship by directly optimising the engine power allocation while subjecting the overall voyage to an average speed or sailing time constraint. This is performed in two steps:

- **Route Segmentation:** A metocean-conditions–based score was defined for every waypoint along the route; the PELT algorithm is then employed to segment the route into legs, and the combined approach here is named MS-PELT. The MS-PELT algorithm optimises the number of power adjustments along the route, reducing the dimensionality of the power allocation process. The algorithm is benchmarked against the TICC clustering algorithm applied on expected metocean conditions during the voyage.

- **Parallel Scenario Dynamic Programming:** A simplification of the interconnection between the legs is used to pre-compute parallel sailing scenarios, generating an optimisation network. Dynamic programming is then used to find the optimal power allocation on the determined voyage segments. Here, the objective function is the total fuel consumption of the voyage M_{fuel} under a sailing time constraint (ETA). As in previous research, XG Boost-based models were used to model the fuel consumption rate and sailing speed.

4.3.3 Results and Discussion

The MS-PELT and TICC segmentation methods were applied to segment routes under varying metocean conditions. MS-PELT demonstrated the capability to produce larger and more consistent route segments. In contrast, TICC produced smaller and more frequent route segments than MS-PELT. This is a disadvantage as it encourages frequent engine power adjustments that could increase fuel consumption and operational complexity. The segmentation results are shown in Figures 4.11a and 4.11b. It is noteworthy that MS-PELT is more computationally efficient, completing in about 10 milliseconds compared to TICC's 50 seconds. This makes MS-PELT more suitable for real-time applications.

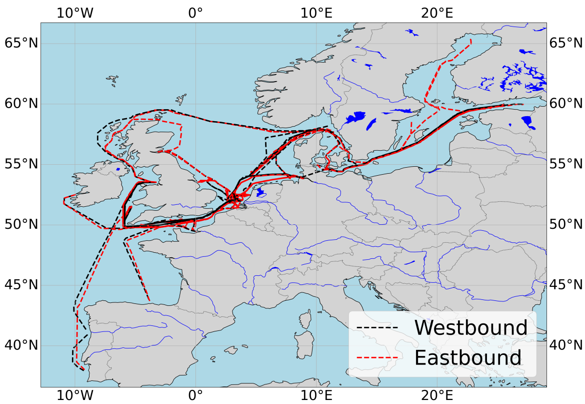
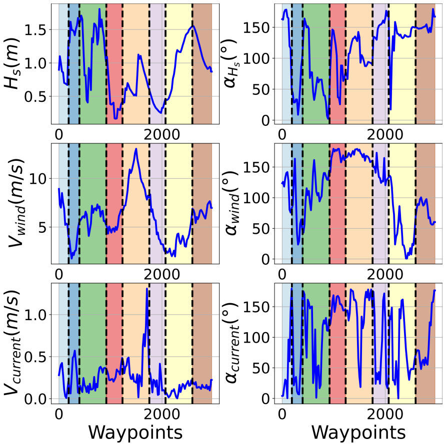
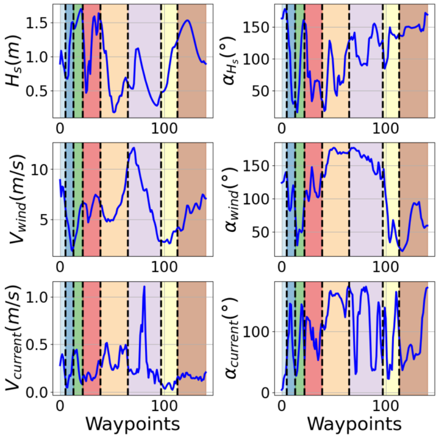


Figure 4.10: Sailing routes for the simulated voyages.

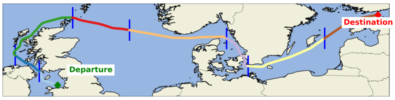
To evaluate the power allocation optimisation, eight case studies were simulated and summarised in Table 4.4. Two cases were selected to illustrate the combined segmentation–optimisation approach. Case 1 covers a route across the Baltic and North Seas (see Fig.4.12a). In this case, the vessel encountered harsher weather conditions—with wave heights reaching up to 3 meters—and the optimised strategy resulted in a 7.2% fuel saving with only a 17-minute delay on the ETA. The strategy prioritised higher power during calm metocean conditions and lower power in segments with adverse conditions, as observed in Figure 4.12b. On the other hand, Case 5 corresponds to a voyage across the English Channel (see Fig.4.13a). Applying the optimisation method led to a 14.5% fuel saving, with only a 7-minute difference in the ETA. Figure 4.13b shows the corresponding optimal power profile. The results summarised in Table 4.5 demonstrate the effectiveness of optimising power settings dynamically, with the MS-PELT method proving an efficient tool for achieving fuel and time savings.
The framework combines efficient data-driven segmentation with dynamic programming, yielding fuel savings and emission reductions of up to 14% on long voyages across European salt waters while incurring time delays of less than 1% of the ETA. The system's high computational efficiency enables real-time or near-real-time application. However, the approach relies heavily on a robust ship performance model, which remains a limitation.



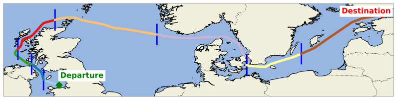
(a) PELT R1.



(b) TICC R1.

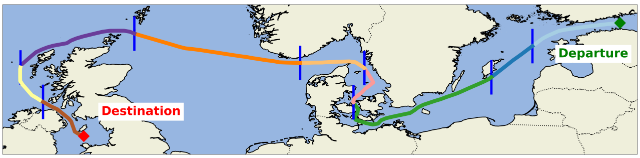


(c) PELT MAP R1.

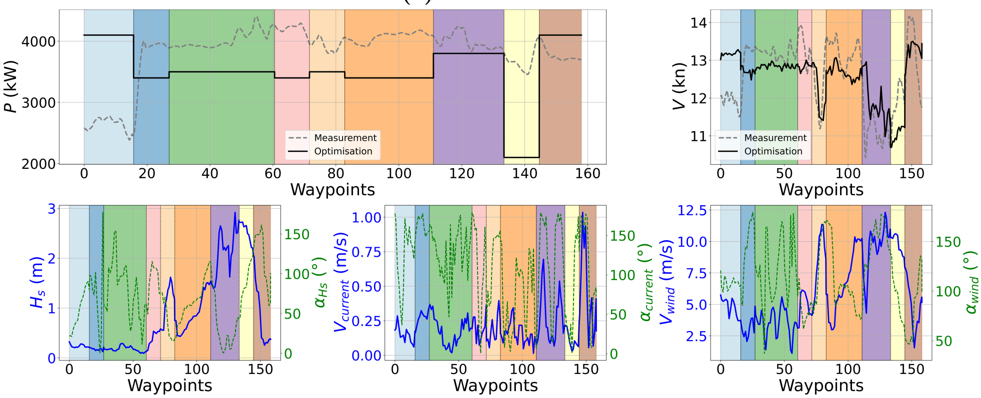


(d) TICC MAP R1.

Figure 4.11: Comparison between PELT and TICC for R1 and MAP R1.

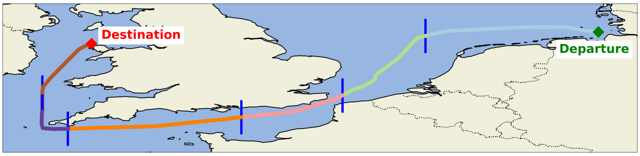


(a) Route Case 1.

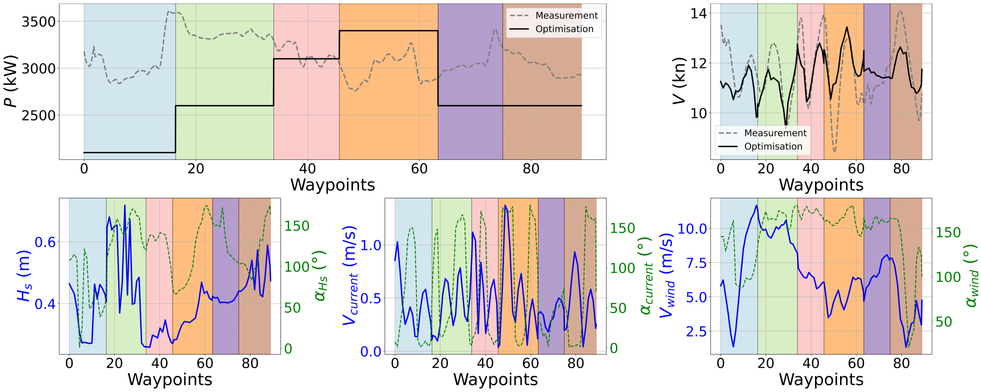


(b) Optimisation R1.

Figure 4.12: Case 1 Optimisation Results.



(a) Route Case 5.



(b) Optimisation R5.

Figure 4.13: Case 5 Optimisation Results.

Table 4.4: Summary of all voyages and GHG emissions optimisation.

Voyage ID	Sailing Area	Distance [km]	ETA [hours]	Time Delay [hours]	Emissions [ton]	Optimised Emissions [ton]
1	Baltic and North Sea	3388.04	145.50	0.28 (+0.19%)	383.16	355.56 (-7.2%)
2	Baltic and North Sea	3089.35	143.83	0.08 (+0.06%)	271.68	259.42 (-4.5%)
3	Baltic	1118.20	60.50	0.17 (+0.28%)	83.06	77.14 (-7.1%)
4	Baltic	1153.58	55.33	0.06 (+0.10%)	90.24	83.63 (-7.3%)
5	North Sea and English Channel	2043.59	108.50	0.18 (+0.17%)	223.41	190.32 (-14.8%)
6	North Sea and English Channel	1718.97	80.33	0.58 (+0.73%)	176.90	160.07 (-9.5%)
7	North Sea and English Channel	1712.24	78.00	-0.15 (-0.19%)	147.47	140.22 (-4.9%)
8	English Channel	980.23	47.00	0.13 (+0.27%)	69.10	62.93 (-8.9%)

Table 4.5: Summary Cases 1 and 5.

	Case 1	**Case 5**
Actual fuel consumption	121.6 tons	46.8 tons
Fuel saving	8.76 tons (-7.2%)	2.5 tons (-14.5%)
Emissions Reduction	28.00 tons (-7.2%)	43.00 tons (-14.5%)
Actual sailing time	5d 21h 25m	3d 06h 00m
Time delay	17m (+0.21%)	-7m (-0.15%)
Segmentation time	30 ms	30 ms
Optimisation time	90 s	90 s

CHAPTER 5

Conclusions

This thesis demonstrates the development and application of an integrated power allocation optimisation framework to improve energy efficiency in short-sea shipping. It combines a data-driven machine-learning approach for ship performance modelling with a power allocation-based optimisation strategy, offering a more practical alternative by reducing the need for frequent engine adjustments. Furthermore, a voyage segmentation algorithm is introduced, ensuring that sailing routes are dynamically segmented based on metocean conditions while minimising the number of segments to maintain computational efficiency. The main findings and conclusions are presented below.

Data analytics
Data analysis and simple regression were powerful tools used throughout all papers, particularly in Paper I, where the data analysis led to identifying an operational trend in the operation of the ferry, potentially leading to a 35% reduction in fuel consumption in Paper II. These data analysis tools were fundamental for the building of Machine Learning models for the data processing steps required for the correct training and validation of the different algorithms evaluated.

AI-based simulations
The AI-based voyage simulation framework used in Papers II and III was developed in several stages: (1) The position interpolation along the route presented no inconveniences, (2) Accurately determining the waypoint-to-waypoint performance presented many difficulties, in particular with estimating vessel speed from engine power and

metocean data. It was addressed using the XGBoost algorithm, as discussed in Chapter 3, resulting in models exhibiting varying levels of generalisation: the speed prediction models showed $R^2 = 0.95 - 0.96$ during training but demonstrated lower reliability in validation ($R^2 = 0.6 - 0.8$). In contrast, the fuel consumption models achieved $R^2 = 0.99$ in training and validation. This discrepancy demonstrates that precise speed prediction remains challenging even when using machine learning. Nevertheless, these models facilitated the simulation of the optimisation control variable and enabled a streamlined simulation of the voyage under different control inputs. Ultimately, they were successfully used to assess the ship's performance across various scenarios and to support an efficient optimisation objective function.

Power allocation optimisation
A power allocation-based optimisation strategy was introduced to enhance voyage efficiency. This approach was explored in two ways: In Paper II, it was implemented using a Bayesian Optimisation DSS, optimising power allocation based on the performance of a simulated double-ended ferry. In Paper III, a novel two-stage power allocation framework was developed, employing dynamic programming to achieve more efficient and adaptive power allocation.
The two-stage power allocation framework accomplished the optimal segmentation of the ship's route while considering uncertain metocean conditions and the optimal power allocation for each leg as presented in Chapter 2. The computationally efficient MS-PELT algorithm was developed and compared with the established multivariate temporal clustering TICC for route segmentation. It illustrates that the MS-PELT algorithm was not only faster at determining ship segments compared to TICC ($10ms$ vs $90s$) but also identifies longer segments for each sea state, reducing the overall number of segments and the corresponding overhead when solving the optimisation using a discrete solver. Then, based on the determined segments, a simplified interconnection approach is introduced to parallelise the optimisation, leading to the Parallel Dynamic Programming algorithm. The framework allows to determine the optimal power allocation for each leg of the voyage.

Fuel consumption savings
The BO-DSS successfully optimised the fuel consumption of the ferry. By using this approach, it was concluded that up to 40% fuel consumption savings were possible when the ferry was operated as proposed. These results were validated in experiments where the ship operated at nearly $R_P = 1$ (see Section 2.3.1). Under these conditions, the ship demonstrated improved efficiency in real scenarios, resulting in average fuel consumption savings of 15%. Similarly, the combined two-stage framework demonstrated that fuel and emissions reduction of 4 to 15% are feasible. These results provide improvements in the operations of short-sea shipping.

CHAPTER 6

Future work

Future efforts will enhance the models using new data processing methods and advanced machine learning algorithms, including developing physics-informed machine learning models. These improvements optimise voyage power allocation, especially when working with limited data. Additionally, the current model's limitations could be addressed by incorporating factors such as biofouling, which can impact fuel efficiency over extended periods. Further improvements could involve developing adaptive learning techniques to manage varying data quality, including low-resolution or noisy datasets. One approach could be the application of transfer learning, where models pre-trained on high-quality data can be adapted to operate with sparser or lower-quality datasets. Another area to look into is the implementation of reinforcement learning (RL) algorithms to solve the general voyage optimisation problem. These algorithms would allow the system to learn optimal power allocation policies through continuous interaction with the operational environment.

Optimising power allocation could also extend beyond fuel efficiency to consider the alignment of arrival times with port slot availability. These optimisation strategies could minimise idle time at port, reduce congestion, and improve fuel efficiency by avoiding last-minute speed changes to meet scheduled arrival times. Lastly, integrating fuel costs and potential penalties for late or early arrivals into the optimisation framework could provide a more economically optimal solution, offering valuable insights for the shipping industry.

References

Abebe, M., Shin, Y., Noh, Y., Lee, S., & Lee, I. (2020). Machine learning approaches for ship speed prediction towards energy efficient shipping. *Applied Sciences*, *10*(7), 2325. https://doi.org/10.3390/app10072325

Bahrami, N., & Siadatmousavi, S. M. (2024). Ship voyage optimisation considering environmental forces using the iterative dijkstra's algorithm. *Ships and Offshore Structures*, *19*(8), 1173–1180. https://doi.org/10.1080/17445302.2023.2231200

Bassam, A. M., Phillips, A. B., Turnock, S. R., & Wilson, P. A. (2022). Ship speed prediction based on machine learning for efficient shipping operation. *Ocean Engineering*, *245*, 110449. https://doi.org/10.1016/j.oceaneng.2021.110449

Beşikçi, E. B., Arslan, O., Turan, O., & Ölçer, A. I. (2016). An artificial neural network based decision support system for energy efficient ship operations. *Computers & Operations Research*, *66*, 393–401. https://doi.org/10.1016/j.cor.2015.04.004

Brooks, M. R., Sanchez, R. J., & Wilmsmeier, G. (2014). Developing Short Sea Shipping in South America: Looking beyond Traditional Perspectives. *Ocean Yearbook Online*, *28*(1), 495–525. https://doi.org/10.1163/22116001-02801018

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*(368), 829–836.

Comi, A., & Polimeni, A. (2020). Assessing the Potential of Short Sea Shipping and the Benefits in Terms of External Costs: Application to the Mediterranean Basin. *Sustainability*, *12*(13), 5383. https://doi.org/10.3390/su12135383

Copernicus Climate Change Service (C3S). (2023). Era5 hourly data on pressure levels from 1940 to present [DOI: 10.24381/cds.bd0915c6 (Accessed on 14/11/2024)].

Dai, X., Sheng, K., & Shu, F. (2022). Ship power load forecasting based on PSO-SVM. *Mathematical Biosciences and Engineering*, *19*(5), 4547–4567. https://doi.org/10.3934/mbe.2022210

Douet, M., & Cappuccilli, J. F. (2011). A review of Short Sea Shipping policy in the European Union. *Journal of Transport Geography*, *19*(4), 968–976. https://doi.org/10.1016/j.jtrangeo.2011.03.001

Du, W., Li, Y., Shi, J., Sun, B., Wang, C., & Zhu, B. (2023). Applying an improved particle swarm optimization algorithm to ship energy saving. *Energy*, *263*, 126080. https://doi.org/10.1016/j.energy.2022.126080

Du, Y., Meng, Q., Wang, S., & Kuang, H. (2019). Two-phase optimal solutions for ship speed and trim optimization over a voyage using voyage report data. *Transportation Research Part B: Methodological*, *122*, 88–114.

Fadda, P., Fancello, G., Mancini, S., Pani, C., & Serra, P. (2020). Design and optimisation of an innovative two-hub-and-spoke network for the Mediterranean short-sea-shipping market. *Computers & Industrial Engineering*, *149*, 106847. https://doi.org/10.1016/j.cie.2020.106847

Fan, A., Yang, J., Yang, L., Liu, W., & Vladimir, N. (2022). Joint optimisation for improving ship energy efficiency considering speed and trim control. *Transportation Research Part D: Transport and Environment*, *113*, 103527.

Farnsworth, A. (2022). Long live the internal combustion engine [Accessed on 14-Nov-2024]. https://www.wartsila.com/insights/article/long-live-the-internal-combustion-engine

Guericke, S., & Tierney, K. (2015). Liner shipping cargo allocation with service levels and speed optimization. *Transportation Research Part E: Logistics and Transportation Review*, *84*, 40–60. https://doi.org/10.1016/j.tre.2015.10.002

Hallac, D., Leskovec, J., & Boyd, S. (2017). Toeplitz inverse covariance-based clustering of multivariate time series data. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 215–223. https://doi.org/10.1145/3097983.3098066

Hallac, D., Vare, S., Boyd, P. S., & Leskovec, J. (2017). Toeplitz inverse covariance-based clustering of multivariate time series data. *CoRR*, *abs/1706.03161*. http://arxiv.org/abs/1706.03161

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hollenbach, K. (1998). Estimating resistance and propulsion for single-screw and twin-screw ships. *Schiffstechnik/Ship Technology Research*, *45*(2), 72–76.

Holtrop, J., & Mennen, G. (1982). An approximate power prediction method. *International Shipbuilding Progress*, *29*, 166–170.

International Maritime Organization. (2009). *Second imo ghg study 2009*.

Karagiannidis, P., & Themelis, N. (2021). Data-driven modelling of ship propulsion and the effect of data pre-processing on the prediction of ship fuel consumption and speed loss. *Ocean Engineering*, *222*, 108616. https://doi.org/10.1016/j.oceaneng.2021.108616

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Kim, Y.-R., Jung, M., & Park, J.-B. (2021). Development of a fuel consumption prediction model based on machine learning using ship in-service data. *Journal of Marine Science and Engineering*, *9*(2), 137. https://doi.org/10.3390/jmse9020137

Lang, X., & Mao, W. (2020). A semi-empirical model for ship speed loss prediction at head sea and its validation by full-scale measurements. *Ocean Engineering*, *209*, 107494. https://doi.org/10.1016/j.oceaneng.2020.107494

Lang, X., & Mao, W. (2021). A practical speed loss prediction model at arbitrary wave heading for ship voyage optimization. *Journal of Marine Science and Application*, *20*(3), 410–425. https://doi.org/10.1007/s11804-021-00224-z

Lang, X., Wu, D., & Mao, W. (2022). Comparison of supervised machine learning methods to predict ship propulsion power at sea. *Ocean Engineering*, *245*, 110387. https://doi.org/10.1016/j.oceaneng.2021.110387

Lang, X., Wu, D., & Mao, W. (2024). Physics-informed machine learning models for ship speed prediction. *Expert Systems with Applications*, *238*, 121877. https://doi.org/10.1016/j.eswa.2023.121877

Latarche, M. (2021). Pounder's Marine Diesel Engines and Gas Turbines. Butterworth-Heinemann.

Laurie, A., Anderlini, E., Dietz, J., & Thomas, G. (2021). Machine learning for shaft power prediction and analysis of fouling related performance deterioration. *Ocean Engineering*, *234*, 108886. https://doi.org/10.1016/j.oceaneng.2021.108886

Lee, S.-M., Roh, M.-I., Kim, K.-S., Jung, H., & Park, J. J. (2018). Method for a simultaneous determination of the path and the speed for ship route planning problems. *Ocean Engineering*, *157*, 301–312. https://doi.org/https://doi.org/10.1016/j.oceaneng.2018.03.068

Li, X., Sun, B., Jin, J., & Ding, J. (2022). Speed optimization of container ship considering route segmentation and weather data loading: Turning point-time segmentation method. *Journal of Marine Science and Engineering*, *10*(12), 1835. https://doi.org/10.3390/jmse10121835

Li, X., Sun, B., Jin, J., & Ding, J. (2023). Ship speed optimization method combining fisher optimal segmentation principle. *Applied Ocean Research*, *140*, 103743. https://doi.org/10.1016/j.apor.2023.103743

Li, Z., Wang, K., Hua, Y., Liu, X., Ma, R., Wang, Z., & Huang, L. (2024). Ga-lstm and nsga-iii based collaborative optimization of ship energy efficiency for low-carbon shipping. *Ocean Engineering*, *312*, 119190. https://doi.org/10.1016/j.oceaneng.2024.119190

Ma, D., Ma, W., Jin, S., & Ma, X. (2020). Method for simultaneously optimizing ship route and speed with emission control areas. *Ocean Engineering*, *202*, 107170. https://doi.org/https://doi.org/10.1016/j.oceaneng.2020.107170

Malte Mittendorf, M., Nielsen, U. D., Bingham, H., & Dietz, J. (2023). Assessment of added resistance estimates based on monitoring data from a fleet of container vessels. *Ocean engineering*, *272*, 113892.

Mao, W., Li, Z., Ogeman, V., & Ringsberg, J. (2015). A regression and beam theory based approach for fatigue assessment of containership structures including bending and torsion contributions. *Marine Structures*, *41*, 244–266.

Mao, W., Rychlik, I., Wallin, J., & Storhaug, G. (2016). Statistical models for the speed prediction of a container ship. *Ocean Engineering*, *126*, 152–162. https://doi.org/10.1016/j.oceaneng.2016.08.033

Mercator Océan International. (2024). Global ocean physics analysis and forecast [Copernicus Marine Service].

Mulligan, R. F., & Lombardo, G. A. (2006). Short sea shipping: — Alleviating the environmental impact of economic growth. *WMU Journal of Maritime Affairs*, *5*(2), 181–194. https://doi.org/10.1007/BF03195103

Papadimitriou, S., Lyridis, D. V., Koliousis, I. G., Tsioumas, V., Sdoukopoulos, E., & Stavroulakis, P. J. (2018). *The dynamics of short sea shipping: New practices and trends*. Palgrave Macmillan.

Parkes, A., Sobey, A., & Hudson, D. (2018). Physics-based shaft power prediction for large merchant ships using neural networks. *Ocean Engineering*, *166*, 92–104. https://doi.org/10.1016/j.oceaneng.2018.07.060

Qi, X., & Song, D.-P. (2012). Minimizing fuel emissions by optimizing vessel schedules in liner shipping with uncertain port times. *Transportation Research Part E: Logistics and Transportation Review*, *48*(4), 863–880. https://doi.org/10.1016/j.tre.2012.02.001

Raza, Z. (2020). Effects of regulation-driven green innovations on short sea shipping's environmental and economic performance. *Transportation Research Part D: Transport and Environment*, *84*, 102340. https://doi.org/10.1016/j.trd.2020.102340

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639.

Schirmann, M., Gose, J., & Collette, M. (2023). A comparison of physics-informed data-driven modeling architectures for ship motion predictions. *Ocean engineering*, *286*, 115608.

Shang, C., Fu, L., Bao, X., Xiao, H., Xu, X., & Hu, Q. (2024). Dynamic joint optimization of power generation and voyage scheduling in ship power system based on deep reinforcement learning. *Electric Power Systems Research*, *229*, 110165.

Sørensen, A. J. (2013). *Marine cybernetics: Modelling and control of marine systems.* NTNU.

Sørensen, A. J., Fossen, T. I., & Berge, S. P. (1997). Marine dynamic positioning systems. *Control Engineering Practice*, *5*(4), 481–486.

Szlapczynska, J., & Szlapczynski, R. (2019). Preference-based evolutionary multi-objective optimization in ship weather routing. *Applied Soft Computing*, *84*, 105742. https://doi.org/https://doi.org/10.1016/j.asoc.2019.105742

Tzortzis, G., & Sakalis, G. (2021). A dynamic ship speed optimization method with time horizon segmentation. *Ocean Engineering*, *226*, 108840. https://doi.org/https://doi.org/10.1016/j.oceaneng.2021.108840

Ventrafiken. (2021, March). Datainsamling för artificiellt intelligens kan minska färjors energiförbrukning [Accessed: 2024-11-27]. https://ventrafiken.se/en/2021/03/08/pressmeddelande/

Vergara, D., Alexandersson, M., Lang, X., & Mao, W. (2023). A machine learning-based bayesian decision support system for efficient navigation of double-ended ferries [In Press]. *Journal of Ocean Engineering and Science*. https://doi.org/10.1016/j.joes.2023.11.002

Wang, H. (2020). *Development of voyage optimization algorithms for sustainable shipping and their impact to ship design* (Publication No. 4726) [Doctoral dissertation, Department of Mechanics and Maritime Sciences, Chalmers University of Technology] [Telephone +46 (0)31-772 1000].

Wang, H., Lang, X., & Mao, W. (2021). Voyage optimization combining genetic algorithm and dynamic programming for fuel/emissions reduction. *Transportation Research Part D: Transport and Environment*, *90*, 102670. https://doi.org/10.1016/j.trd.2020.102670

Wang, H., Lang, X., Mao, W., Zhang, D., & Storhaug, G. (2020). Effectiveness of 2D optimization algorithms considering voluntary speed reduction under uncertain metocean conditions. *Ocean Engineering*, *200*, 107063.

Wang, H., Mao, W., & Eriksson, L. (2019). A three-dimensional dijkstra's algorithm for multi-objective ship voyage optimization. *Ocean Engineering*, *186*, 106131. https://doi.org/10.1016/j.oceaneng.2019.106131

Wang, K., Li, J., Huang, L., Ma, R., Jiang, X., Yuan, Y., Mwero, N. A., Negenborn, R. R., Sun, P., & Yan, X. (2020). A novel method for joint optimization of the sailing route and speed considering multiple environmental factors for more energy efficient shipping. *Ocean Engineering*, *216*, 107591. https://doi.org/10.1016/j.oceaneng.2020.107591

Wang, K., Yan, X., Yuan, Y., Jiang, X., Lodewijks, G., & Negenborn, R. R. (2017). Study on route division for ship energy efficiency optimization based on big environment data. *Proceedings of the 4th International Conference on*

Transportation Information and Safety (ICTIS 2017), 111–116. https://doi.org/10.1109/ICTIS.2017.8047752

Wang, K., Xu, H., Li, J., Huang, L., Ma, R., Jiang, X., Yuan, Y., Mwero, N. A., Sun, P., Negenborn, R. R., & Yan, X. (2021). A novel dynamical collaborative optimization method of ship energy consumption based on a spatial and temporal distribution analysis of voyage data. *Applied Ocean Research*, *112*, 102657. https://doi.org/https://doi.org/10.1016/j.apor.2021.102657

Wang, S., & Meng, Q. (2012). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E: Logistics and Transportation Review*, *48*(3), 701–714. https://doi.org/10.1016/j.tre.2011.12.003

Wu, W.-M. (2020). The optimal speed in container shipping: Theory and empirical evidence. *Transportation Research Part E: Logistics and Transportation Review*, *136*, 101903. https://doi.org/10.1016/j.tre.2020.101903

Yan, X. P., Wang, K., Yuan, Y. P., Jiang, X. L., & Negenborn, R. R. (2018). Energy-efficient shipping: An application of big data analysis for optimizing engine speed of inland ships considering multiple environmental factors. *Ocean Engineering*, *169*, 457–468. https://doi.org/10.1016/j.oceaneng.2018.08.050

Yu, Y., Zhang, H., Mu, Z., Li, Y., Sun, Y., & Liu, J. (2024). Trim and engine power joint optimization of a ship based on minimum energy consumption over a whole voyage [Information pending]. *Unknown*, *Unknown*, Unknown.

Zaccone, R., Ottaviani, E., Figari, M., & Altosole, M. (2018). Ship voyage optimization for safe and energy-efficient navigation: A dynamic programming approach. *Ocean Engineering*, *153*, 215–224. https://doi.org/10.1016/j.oceaneng.2018.01.100

Zhang, C., Zhang, D., Zhang, M., Zhang, J., & Mao, W. (2022). A three-dimensional ant colony algorithm for multi-objective ice routing of a ship in the arctic area. *Ocean Engineering*, *266*, 113241. https://doi.org/https://doi.org/10.1016/j.oceaneng.2022.113241

Zhang, M., Tsoulakos, N., Kujala, P., & Hirdaris, S. (2024). A deep learning method for the prediction of ship fuel consumption in real operational conditions. *Engineering Applications of Artificial Intelligence*, *130*, 107425. https://doi.org/10.1016/j.engappai.2023.107425