Predicting fuel consumption for different types of bulk carriers

KAOUTAR HAJLI*[1], MIKAEL RÖNNQVIST[2], JEAN-FRANÇOIS CORDEAU[3], JEAN-FRANÇOIS AUDY[4], CAMÉLIA DADOUCHI[5], GURJEET WARYA[6], TRUNG NGO[7]

[1] Université Laval; kaoutar.hajli.1@ulaval.ca
[2] Université Laval; mikael.ronnqvist@gmc.ulaval.ca
[3] HEC Montréal; jean-francois.cordeau@hec.ca
[4] Université du Québec à Trois-Rivières; jean-francois.audy@uqtr.ca
[5] Polytechnique Montréal; camelia.dadouchi@polymtl.ca
[6] True North Marine; gwarya@tnmservices.com
[7] True North Marine; tngo@tnmservices.com

Abstract: Predicting the fuel consumption of ships is a key element for weather routing to achieve cost savings, operational efficiency, and compliance with IMO regulations on emissions. This study is based on data from over 1500 ships, and proposes two approaches to predict the fuel consumption of bulk carriers based on historical voyages, weather data and the physical characteristics of ships. The first approach uses a Multiple Linear Regression model while the second approach employs a mathematical model that integrates additional route information not accounted for by the machine learning model. Results show that both models accurately predict fuel consumption, with more than 90% of voyages achieving an average error below 0.02 metric ton per nautical mile. Furthermore, while both models can predict fuel consumption for an entire route, the second model uses the entire route information to reduce prediction errors.

Keywords: Weather routing; Bulk carriers; Fuel consumption; Machine learning.

1. Introduction

The effects of climate change are becoming more apparent, such as rising average and extreme temperatures, altered precipitation patterns, and increased frequency of severe weather events. While maritime transportation is generally considered a cleaner mode of transportation compared to air, rail or road transportation, it is also facing increasing pressure to reduce its CO_2 emissions.

Since 2018, the International Maritime Organization (IMO) has adopted a strategy to reduce greenhouse gas (GHG) emissions from ships at its Marine Environment Protection Committee. This strategy aims to decrease the carbon intensity of international maritime transportation by at least 40% by 2030, compared to levels in 2008, and to lower total annual GHG emissions by at least 50% by 2050, also compared to 2008 levels (IMO, 2018). As part of this strategy, a new regulation has been imposed, since January 2020, reducing the limit of sulfur content in fuel oil from 3.5% to 0.5%. The IMO's GHG strategy establishes a timeline and framework for deciding when to consider various policies, and includes a list of potential short-, medium-, and long-term policy measures. The proposed measures can be divided into three categories: technological measures, operational measures, and market-based measures.

Operational measures have drawn significant attention due to their relatively low cost and immediate applicability. One such measure is the Ship Energy Efficiency Management Plan (SEEMP), established under the MARPOL Annex VI regulations (MEPC, 2022), encourages vessel operators to optimize their operational strategies to reduce fuel consumption and emissions. One widely endorsed practice within this framework is the adoption of weather routing techniques and slow steaming strategies. By adjusting sailing speeds and navigating through favorable meteorological and oceanographic conditions, ships can reduce fuel use while maintaining voyage schedules and safety standards.

Simonsen et al. (2015) describe Ship Weather Routing (SWR) as the process of determining the optimal course and speed for a ship's voyage based on various factors including nautical charts, forecasted sea conditions, the captain's expertise, and the specific characteristics of the ship. Weather routing can help identify the most efficient speed for a voyage, which can minimize voyage costs, fuel consumption, improve energy efficiency, and ensure the safety of the ship, crew, and cargo, particularly for bulk carriers and other commercial vessels that undertake long, fuel-intensive voyages. The use of weather routing has been shown to reduce fuel consumption by up to 3% (Armstrong, 2013). A recent study (Ormevik et al., 2023) enhances this understanding by approaching the problem of scheduling on predetermined multi-stop routes, with a focus on speed optimization under variable weather conditions. They introduced a model that factors in wave direction, wave period, and wave height, resulting in an average reduction of 4.5% in fuel consumption across different routes and weather scenarios.

Despite these promising advances, the effective implementation of voyage optimization strategies remains heavily dependent on the accuracy of fuel consumption prediction models. Fuel accounts for a significant portion of a ship's operating cost, often more than 50%, making accurate forecasting not only an environmental imperative but also a financial one. Furthermore, as carbon pricing and emission trading systems become more integrated into maritime regulation, accurate estimation of emissions becomes critical for compliance and strategic planning. Nevertheless, key limitations persist in existing models such as the reliance on single-vessel data, exclusion of key environmental factors like ocean currents, and limited consideration of vessel heterogeneity, which limit their generalizability and operational relevance.

The aim of this study is to fill these gaps by proposing predictive models of fuel consumption that are not reliant on data from a single ship, but rather on a large dataset comprising 1513 ships of various physical characteristics, with a specific focus on bulk carriers. The proposed models provide a more robust and generalized approach, making them suitable for integration into weather routing systems to improve decision-making processes. By incorporating a wide range of vessels, this research offers a broader applicability to different ship types and operational scenarios. The core contribution of this study lies in the development of a multiple linear regression (MLR) model for predicting ship fuel consumption, utilizing a combination of historical voyage data, weather conditions, and vessel characteristics. This model represents a significant step forward in accurately predicting fuel consumption under varying operational and environmental conditions for different types of bulk carriers. In addition to the MLR model, a mathematical optimization model is introduced that accounts for total fuel consumption over the entire voyage, rather than just at specific points. This allows us to reduce errors by formulating constraints on the total error across the entire route, rather than just the fuel consumption measured at each noon report. The results show that the proposed models were able to predict the fuel consumption of more than 90% of voyages with an average error of less than 0.02 metric ton per nautical mile (MT/NM) for both Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

The structure of this paper is as follows. Section 2 presents a brief review of the relevant literature to contextualize the study. Section 3 describes the proposed methodology, including the problem formulation, data understanding and preprocessing, and the proposed models. Section 4 presents the case study. Section 5 presents the results obtained from the study, followed by a discussion of these findings in Section 6. Finally, Section 7 provides the conclusion of the paper.

2. Literature review

Effective fuel consumption management is a crucial part of this process, as fuel costs make up a significant portion of a ship's operating expenses (Bialystocki & Konovessis, 2016). Accurate forecasting of fuel consumption in different sea conditions requires detailed data on fuel usage. Analyzing this detailed data enables better prediction and management of fuel requirements, thereby reducing operating costs and emissions and improving overall vessel efficiency. Multiple studies support this focus and investigate ways to enhance prediction accuracy and thus reduce emissions. Merien-Paul et al. (2018) identify significant inaccuracies in traditional bottom-up methodologies for estimating fuel consumption and emissions, emphasizing the need for more reliable approaches based on in-situ data. Bottom-up methodologies estimate emissions by aggregating activity-specific data, such as engine load factors and average fuel consumption rates, to calculate total fuel use and emissions for vessels or fleets. However, these methods often fail to account for the complex interaction between environmental conditions and vessel-specific operational profiles, which limits their predictive reliability in dynamic routing scenarios.

To address this, several researchers have turned to machine learning and statistical techniques that incorporate real-time environmental and operational inputs. Chen et al. (2024) demonstrate the potential of Artificial Neural Networks (ANN) for predicting emissions from harbour vessels, achieving up to a 62% improvement in accuracy over bottom-up methods by incorporating meteorological such as wind speed, wave height, ambient temperature, and atmospheric pressure, as well as ship-specific factors such as GPS location, vessel speed and shaft power. Similarly, Uyanık et al. (2020) utilize statistical models, including Multiple Linear Regression and Ridge Regression, to predict commercial fuel consumption, highlighting the impact of variables such as engine performance and weather. Yang et al. (2024) explored the application of machine learning models, including tree-based ensembles like Random Forest and XGBoost, to predict fuel oil consumption of an ocean-going container ship. Their models incorporated voyage variables such as speed, displacement, and trim, alongside meteorological features such as wind, wave, sea current, sea water salinity and sea water temperature. While their approach demonstrated improved accuracy over simpler baselines, they emphasized the sensitivity of model performance to the availability and quality of meteorological inputs. In addition, their study was limited to a single vessel, which may restrict the generalizability of the findings. In a related study, Zhang et al. (2024) developed deep learning models, including Long Short-Term Memory (LSTM) networks, to predict ship fuel consumption using high-frequency sensor data from a Kamsarmax bulk carrier. Their models achieved high accuracy in capturing complex nonlinear dependencies. However, these models often lack interpretability, making it

challenging to extract the marginal effects of features. Additionally, the models generally require large, high-quality datasets from similar vessels, compromising transferability to different ships or operational profiles.

Voyage optimization and weather routing have also been explored as effective strategies to reduce fuel consumption. Zis et al. (2020) conducted a literature review on weather routing and voyage optimization in maritime shipping. The authors provided a comprehensive overview of the main methodological approaches used in the field and identified the key disciplines that address this problem. The review demonstrates that the majority of the studies in this area are focused on either minimizing fuel consumption or reducing sailing time. Among the 40 relevant works reviewed in this literature survey, only five studies (Hagiwara, 1989; Hagiwara & Spaans, 1987; Kobayashi et al., 2011; Prpić-Oršić et al., 2014; Szlapczynska, 2015) have taken into account ocean currents in addition to wind and waves when optimizing weather routing in maritime shipping even though currents have been shown to make a significant contribution to fuel savings in real-world routing scenarios (Godet et al., 2024). Furthermore, while some studies (Kepaptsoglou et al., 2015; Wang et al., 2018) have examined the impact of weather routing on fuel consumption using case studies involving multiple vessels, the majority of the studies reviewed in this paper have only used data from a single vessel, which may restrict the applicability and generalizability of the findings, specifically to other types of ships.

Across these studies, several key limitations emerge. First, most rely on restricted datasets, often derived from single vessels or limited operating conditions, constraining generalization across ship types and routes. Second, a persistent trade-off between interpretability and accuracy is observed: while deep learning and ensemble models excel at capturing nonlinear effects, their "black-box" nature limits their usefulness for operational insights. Finally, few of these approaches integrate all major environmental resistance factors simultaneously—particularly ocean currents—or enforce voyage-level consistency between predicted and actual fuel use, which is important for operational optimization and realistic routing applications.

In previous research (Hajli et al., 2024), we applied the MLR model to predict daily fuel consumption for a set of sister ships. The model used engine revolutions per minute raised to the third power (RPM^3) as the primary operational input, along with weather resistance factors—wind, wave, and current—expressed as speed loss coefficients. While this study provided accurate predictions for the vessels considered, it was limited to a single ship type and did not account for vessel load or enable voyage-level consistency in fuel consumption. In the current study, we address the gaps previously mentioned by modeling fuel consumption per nautical mile, leveraging the empirical relationship between this metric and the square of the speed over ground (SOG^2) (Kee et al., 2020). The same environmental resistance factors are included, and we additionally incorporate vessel load condition (laden or ballast) to account for variations in operational status. Furthermore, rather than focusing on a single vessel type, we consider a diverse set of 1513 bulk carriers. These vessels are grouped into three behaviorally distinct clusters using k-means clustering based on observed RPM and fuel consumption per day. Within each cluster, a separate MLR model is developed, enabling more context-specific predictions while maintaining model interpretability and consistency across vessel types. In addition to the MLR models, a complementary mathematical optimization model is introduced that incorporates the constraint that the total predicted fuel consumption for each voyage must equal the actual total fuel consumed. This approach moves beyond pointwise predictions to consider total fuel use holistically over entire voyages, thereby managing errors more effectively and enabling a more continuous and comprehensive optimization of fuel efficiency throughout the operational profile of the vessel.

Direct quantitative comparisons with other studies using metrics such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) are challenging due to differences in vessels, datasets, and input variables. However, conceptually, our approach extends previous methods by integrating operational, environmental, and vessel-specific factors across a large, diverse fleet, incorporating voyage-level constraints, and using cluster-specific modeling to provide a more generalizable framework for fuel consumption prediction and optimization.

Overall, this literature synthesis highlights the need for models that combine operational and environmental factors across diverse vessels to support more accurate and generalizable fuel consumption predictions, which this study aims to address.

3. Materials and Methods

3.1. Problem formulation

Accurately predicting fuel consumption is a significant challenge in the shipping industry, as it is influenced by a variety of internal and external factors. Our study is based on data collected from so-called noon reports sent by the captains of ships every 24 hours with a set of information on the state of the ship. In addition, our data contains weather forecasts, provided by an independent organization, on meteorological conditions, reported every 6 hours for the next seven days. Our objective is to use this information to develop models that take into account not only the speed of the vessel but

also the weather conditions, and the different physical characteristics of the ships. These models are essential because they enable more precise forecasting of fuel consumption, which can lead to significant cost savings and operational efficiencies.

3.2. General data context

The present study is based on historical voyages for the development of fuel consumption prediction models. At each noon report of the voyage of a given vessel, we have data on the identification of the vessel, its location, its SOG, the RPM of the engine and the remaining level of each type of fuel: Intermediate Fuel Oil (IFO) and Marine gas oil (MGO). The fuel consumption considered in this study specifically corresponds to the main propulsion system.

In addition, our data contains the direction and the speed/height of the wind, waves and currents at each point of the passage. The resulting speed loss for each meteorological component is calculated based on its added resistance, as described in the methodology of Kim et al. (2017). This added resistance is combined with the ship's baseline resistance, which is based on its performance in calm water. By calculating the total resistance under sea conditions, the necessary power for the ship to maintain its speed is determined. Finally, the achievable speed in these conditions is compared with the ship's speed in calm water to get the speed loss. The speed loss will combine the meteorological information, specifically the speed or height of wind, waves, and current, the speed and the type of vessel into a single value that we will call wind/wave/current factor (Figure 1). Furthermore, our industrial partner provided us with information regarding each bulk carrier, including IMO number, deadweight tonnage, year built, gross tonnage, summer draft, vessel size, engine type, propeller type, shipyard where built, engine model, length overall, and maximum width.

3.3. Data preprocessing

The dataset used in this study consists of noon reports manually recorded by ship captains, a process susceptible to various sources of error and inconsistency. Given the nature of these reports and the risk of human error, data quality can be compromised by inaccuracies, discrepancies and gaps. To ensure the robustness of our analysis despite these problems, data preprocessing is essential.

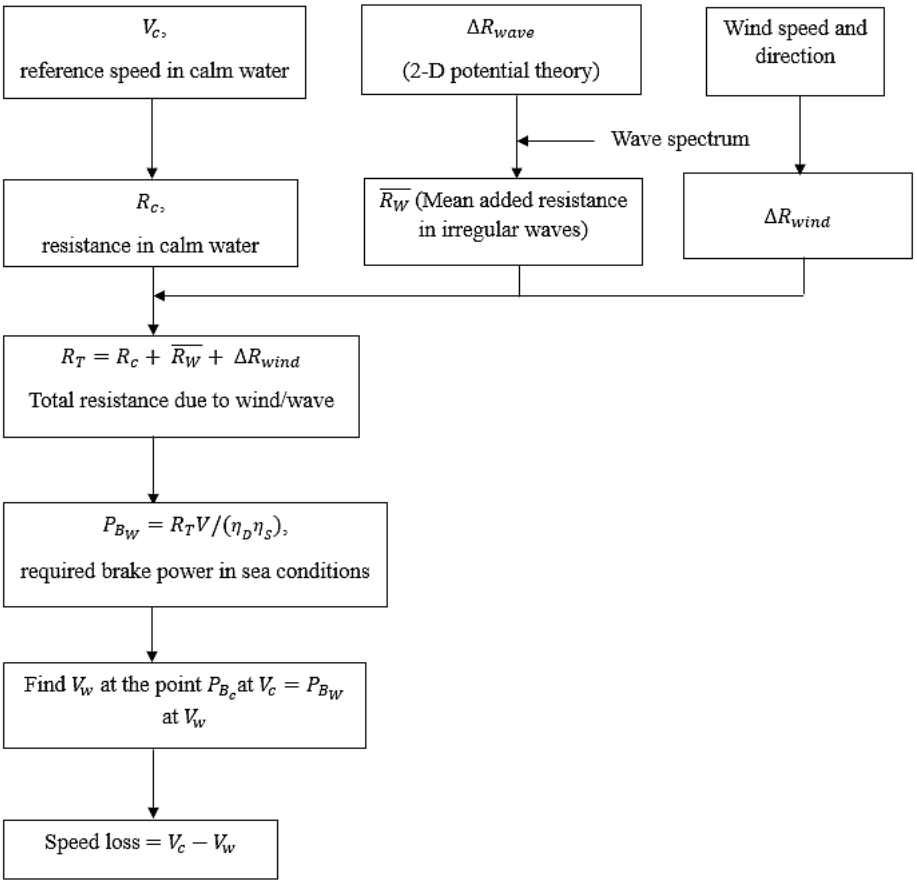


Figure 1: Ship speed loss estimation. Adapted from Kim et al. (2017).

3.3.1. Outlier removal

Data cleaning is a crucial step that affects our data analysis and prediction models. In this step, we first removed the null values of fuel consumption since ships always consume fuel during their voyages. For each engine, there are minimum and maximum values for the RPM provided by the builder that must be respected in order not to deteriorate it. Therefore, non-compliant RPM values should be removed.

To systematically identify and remove outliers, we employed the Interquartile Range (IQR) method, a robust statistical technique. This approach calculates the IQR as the difference between the third quartile (Q3) and the first quartile (Q1) of the dataset. Data points lying outside the interval $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$ are considered outliers and were removed accordingly. The IQR method is preferred due to its non-parametric nature and resilience to skewed data distributions, making it well-suited for detecting anomalies in maritime operational data.

3.3.2. Feature scaling

Given the diverse ranges and distributions of features influencing fuel consumption, it is essential to normalize the data to ensure that all variables contribute proportionally to the modeling process. To achieve this, we applied the RobustScaler technique, which scales features based on the median and the IQR, thereby reducing the impact of outliers and skewed distributions. Specifically, each feature x is transformed as follows:

$$\mathrm{x}_{\mathrm{scaled}} = \frac{\mathrm{x} - \mathrm{median(x)}}{IQR(x)} \quad (1)$$

This method centers data around the median and scales it according to the IQR, making it more robust in the presence of anomalies or non-normal feature distributions.

3.3.3. Feature engineering

At the feature engineering phase, we transformed our data to prepare it for the modelling step. To achieve this, we computed new features. Firstly, we computed the fuel consumption per nautical mile and per day based on the remaining levels of each fuel type (IFO and MGO) and the distance sailed. Secondly, considering that our speed loss data from wind, waves, and currents was available every 6 hours, and our fuel reports were per 24 hours, we averaged the weather information over each 24-hour period to align the data temporally. This approach was chosen because disaggregating daily fuel consumption into 6-hour intervals would introduce artificial noise, whereas averaging the weather data preserves the integrity of the original fuel measurements while capturing the overall environmental impact on daily vessel performance.

Moreover, since our study is based on developing a model that involves several bulk carriers, we clustered the vessels. Despite the fact that our study is focused on bulk carriers and that they have the same basic design and purpose, they do not have the same behavior during sailing and under different meteorological conditions. To address this heterogeneity and enhance the predictive accuracy of the proposed models, we applied an unsupervised clustering technique to group vessels exhibiting similar behavior.

Specifically, we employed the k-means clustering algorithm using three key operational indicators: daily fuel consumption and engine revolutions per minute (RPM), and speed over ground (SOG). By combining these variables, the clustering distinguishes vessels not only by design but also by how they operate under real-world conditions. Mathematically, the k-means algorithm partitions the dataset by minimizing the within-cluster sum of squared distances:

$$\arg\min_{C} \sum_{i=1}^{k} \sum_{x \in C_i} \left|\left| x - \mu_i \right|\right|^2, \quad (2)$$

where C_i denotes the set of points in cluster i, x is a data point, and μ_i is the centroid of cluster i. The algorithm iteratively assigns data points to clusters and updates centroids until convergence.

Following clustering, feature selection was performed to identify the most relevant predictors for modeling fuel consumption per nautical mile. Specifically, Pearson's correlation coefficient was used to evaluate the strength and direction of linear relationships between candidate features and the target variable. This step ensured the retention of features that significantly contribute to the predictive performance while reducing model complexity and enhancing interpretability.

Finally, further categorization was applied to our data set. Given the variable sailing durations of the collected voyages, we chose to classify them into short, medium and long voyages. This categorization was important for the modeling and evaluation phases, as it allowed us to perform stratified cross-validation, ensuring that each fold

contained a representative distribution of voyage lengths, which improves the robustness and generalizability of our predictive models.

3.4. Prediction models

3.4.1. Multiple linear regression model

Compared to our previous work (Hajli et al., 2024), the modeling approach in this study incorporates several key enhancements to better capture operational variability and improve prediction accuracy. First, fuel consumption is modeled per nautical mile, using the square of speed over ground (SOG^2) as the primary predictor, which explicitly accounts for the non-linear relationship between vessel speed and fuel use. Second, in addition to environmental resistance factors (wind, wave, and current) the model integrates vessel load condition (laden or ballast) to reflect differences in operational status. Third, rather than estimating a single model for all vessels, the fleet is grouped into behaviorally distinct clusters using k-means clustering based on observed RPM and daily fuel consumption, and separate MLR model is estimated for each cluster.

For each cluster c, the MLR model is expressed as follows:

$$f_c = \alpha_{0,c} + \alpha_{1,c}\, SOG_c^2 + \alpha_{2,c} WindFactor_c + \alpha_{3,c} WaveFactor_c + \alpha_{4,c} CurrentFactor_c + \alpha_{5c} Load_c\,, \tag{3}$$

where f_c represents the fuel consumption per nautical mile for cluster c, SOG_c denotes the speed over ground, and $WindFactor_c$, $WaveFactor_c$, and $CurrentFactor_c$ quantify the average speed loss caused by, respectively, wind, waves, and currents over 24 hours within each cluster c and calculated based on the methodology proposed in Kim et al. (2017). $Load_c$ is a binary indicator of the vessel's load condition (laden or ballast). The coefficients $\alpha_{0,c}$, $\alpha_{1,c}$, $\alpha_{2,c}$, $\alpha_{3,c}$, $\alpha_{4,c}$ and $\alpha_{5,c}$ are regression coefficients specific to each cluster, capturing the influence of each variable on fuel consumption. These coefficients are estimated by minimizing the residual sum of squares between the actual and predicted fuel consumption values.

3.4.2. Mathematical model

In our study, we propose a second fuel consumption prediction model utilizing the least squares methodology. This model not only targets minimizing the error in fuel consumption per nautical mile in noon reports, as it is the case in the previous model, but also aims to enhance the accuracy of total fuel consumption for each voyage. The focus on total consumption is due to the fact that the consumption reported in noon reports can be subject to inaccuracies. By focusing on the entire voyage, for which the remaining fuel should tend towards zero, we reduce the impact of these reporting errors and improve the overall reliability of our fuel consumption forecasts.

Let V_c be the set of voyages conducted by vessels in cluster c, I_c be set of noon reports corresponding to these voyages, $I_{v,c} \subset I_c$ be the subset of noon reports concerning voyage $v \in V_c$ and n_c be the number of data points (noon reports) belonging to cluster c. Each noon report $i \in I_{v,c}$ contains information regarding fuel consumption f_i, speed over ground SOG_i, distance sailed since the previous noon report d_i and the load condition $Load_i$ of the vessel. Moreover, meteorological conditions are accounted for through factors representing the average speed reduction caused by wind (WindFactor), waves (WaveFactor), and currents (CurrentFactor) over a 24-hour period. We also define variables $\alpha_{0,c}$, $\alpha_{1,c}$, $\alpha_{2,c}$, $\alpha_{3,c}$, $\alpha_{4,c}$ and $\alpha_{5,c}$ to denote the weights corresponding to the mentioned parameters. Finally, a large positive constant β is introduced to balance the errors between the total fuel consumption over the entire voyage and the fuel consumption measured at each individual noon report.

With these definitions in place, the mathematical model can be formulated as follows, and is solved separately for each cluster c:

$$Min \sum_{i \in I_c} [f_i - [\alpha_{0,c} + \alpha_{1,c}\, SOG_i^2 + \alpha_{2,c} WindFactor_i + \alpha_{3,c} WaveFactor_i + \alpha_{4,c} CurrentFactor_i + \alpha_{5,c} Load_i\,]]^2 + \beta \sum_{v \in V_c} p_v^2 \tag{4}$$

subject to

$$\sum_{i \in I_{v,c}} d_i\, f_i = \sum_{i \in I_{v,c}} d_i\, [\alpha_{0,c} + \alpha_{1,c}\, SOG_i^2 + \alpha_{2,c} WindFactor_i + \alpha_{3,c} WaveFactor_i + \alpha_{4,c} CurrentFactor_i + \alpha_{5,c} Load_i] + p_v \qquad \forall v \in V_C \tag{5}$$

This model is designed with two objectives. First, it minimizes the squared difference between the estimated and actual fuel consumption for each voyage. This ensures the model's predictions closely align with real-world data. Second, the model optimizes a slack variable introduced within the constraint. This variable p_v represents the deviation between the total estimated and actual fuel consumption for each voyage. Minimizing the squared value of this slack variable multiplied by β encourages the model to focus more on finding solutions where the estimated total fuel consumption closely matches the actual values.

3.5. Evaluation

To ensure robust assessment and validation of our models, we employed a stratified 5-fold cross-validation technique. This technique is particularly useful for data sets with classes of different proportions. It involves dividing the dataset into five subsets of approximately equal size. Each subset is then stratified, preserving the proportion of classes in the original dataset, so that each fold contains approximately the same percentage of each class as the original dataset. The technique repeats the model evaluation on all five folds. For each iteration, one fold is selected as the test set and set aside. The model is then trained on the remaining four folds, which constitute the training data set. The model is finally evaluated on the set-aside test set and its performance recorded. The average model performance over the five iterations is calculated and used as an estimate of overall model performance on new data.

Moreover, within each fold, we calculated the mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) as standard performance metrics (Eq. (6), (7), and (8) respectively), where y_i represents the actual values, $\hat{y}_i$ denotes the predicted values and n is the total number of observations. Subsequently, we assessed the performance of each model initially on the entire dataset and then on a per-voyage basis. This methodological approach provided a comprehensive understanding of the errors of the models at the voyage level.

$$\mathrm{MAE} = \frac{1}{n}\sum_i |y_i - \hat{y}_i| \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2} \tag{7}$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_i \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{8}$$

4. Case study

In contrast to the previous case study by Hajli et al. (2024), which analyzed fuel consumption for a small and homogeneous set of nine sister bulk carriers with similar technical characteristics, the present dataset is considerably larger and more diverse. The data originates from our industrial partner, who collects it from ship captains via noon reports. The initial dataset includes information on 1,513 bulk carriers of four different sizes and five different engine models and covers voyages between 2014 and 2022. It includes data on a total of 5,842 voyages and contains 72,125 data

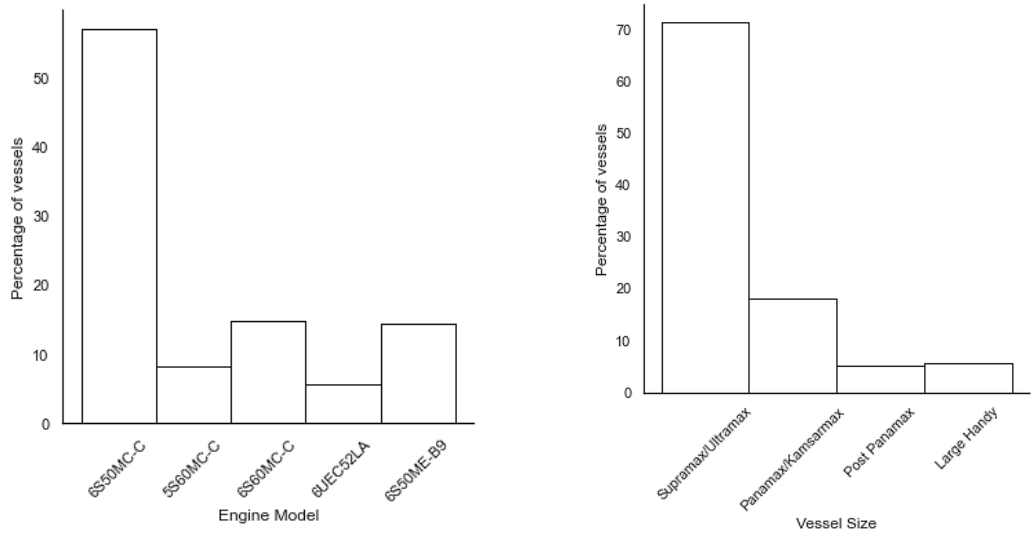


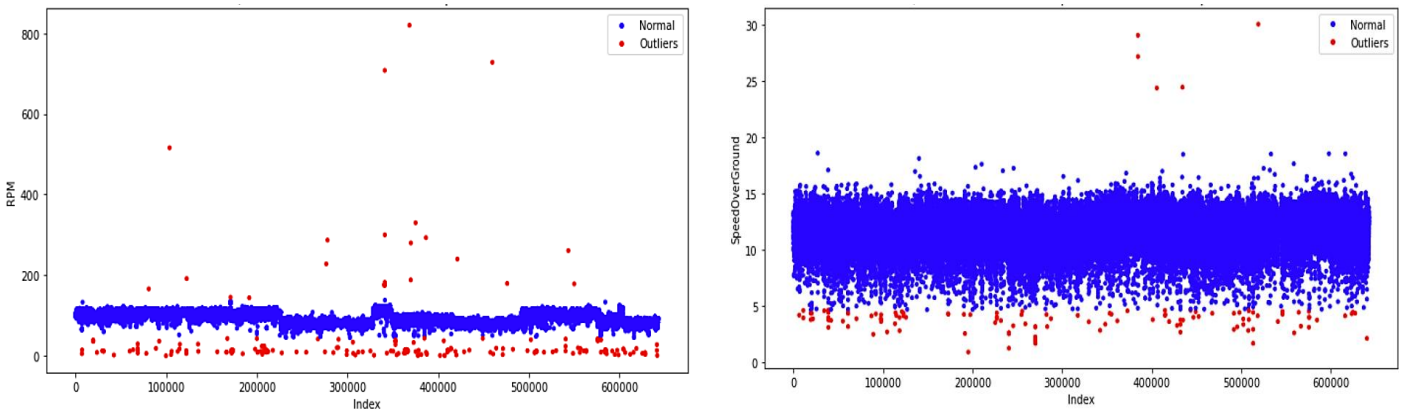
Figure 2: (a) The distribution of engine models across voyages. (b) The distribution of ship sizes across voyages.

points. Figure 2 shows the distribution of engine models and ships sizes across the voyages in our dataset. The vessel size categories range from smaller Large Handy carriers, typically carrying between 30,000 and 40,000 metric tons of cargo, to medium-sized Supramax and Ultramax vessels with capacities of 50,000 to 65,000 metric tons, and Panamax and Kamsarmax vessels carrying up to approximately 82,000 metric tons. Larger Post-Panamax ships exceed the original Panama Canal limits, accommodating loads of over 90,000 metric tons. The engine models present in the dataset include various commonly used marine diesel engines recognized for their efficiency and reliability in bulk carrier propulsion.

5. Results

5.1. Data preprocessing

Data preparation is a crucial phase in our study, serving as the foundation for accurate analysis and meaningful insights. In this process, we meticulously address various issues present in the raw data to ensure its quality and reliability. The initial data cleaning focused on missing values as well as outliers that significantly deviated from the expected range within each feature identified through data visualization as shown in Figure 3, where red dots represent outliers and blue ones correspond to normal data points. The initial data cleaning procedure impacted 12.34% of the original dataset. While this might seem substantial, it is important to contextualize this figure. Our data originates from manual reports provided by captains, introducing the potential for both missing values and manual input errors. Additionally, discrepancies in data may arise from non-reporting instances or inconsistencies in meteorological reports.



(**a**) (**b**)

Figure 3: Scatter plots of RPM values (a) and SOG (b) across data points (x-axis represents data point index).

To maintain the integrity of our analysis, outliers identified in both stages were removed from our dataset. This decision was made considering the adequacy of the remaining data for our modeling purposes, as well as a desire to avoid compromising the accuracy of our results through imputation methods for missing data. Table 1 provides an overview of key statistics pertaining to important features both before and after the data cleaning process, offering

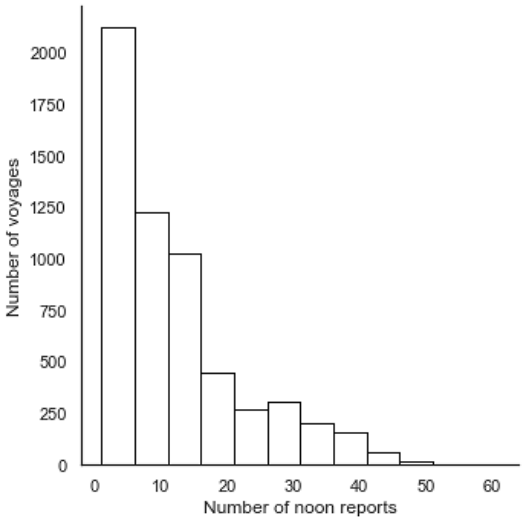


Figure 4: The distribution of the number of noon reports across voyages.

insight into the impact of our cleaning efforts on the dataset's characteristics. Notably, the removal of extreme and implausible values significantly reduced the variability and corrected unrealistic ranges in critical variables such as fuel consumption, wave height, and engine RPM, thereby ensuring a more reliable and representative dataset for modeling.

The distribution of noon reports across voyages shows that almost 71% of them have more than five noon reports (Figure 4). Based on this distribution, we have chosen to classify our data into short, medium, and long voyages in such a way that short voyages have fewer than 10 noon reports, medium voyages have between 10 and 20 noon reports, and long voyages have more than 20 noon reports.

Table 1. Descriptive statistics of the most significant features before and after removing outliers.

	Before removing outliers				After removing outliers			
	Mean	**Standard Deviation**	**Mini-mum**	**Maxi-mum**	**Mean**	**Standard Deviation**	**Mini-mum**	**Maximum**
Fuel Consumption (Mt/nautical mile)	0.096	0.184	-48.035	25.074	0.096	0.020	0.050	0.283
Speed Over Ground (knots)	11.637	1.547	0.181	79.700	11.605	1.399	4.676	18.536
RPM	92.104	15.163	0.080	820.000	92.058	13.233	45.000	136.200
Wind Speed (Beaufort)	3.716	1.491	0.000	26.000	3.857	1.356	0.000	10.000
Wind Factor (knots)	-0.085	0.186	-3.280	2.000	-0.084	0.170	-1.282	0.180
Wave Height (m)	2.697	99.656	0.010	9999.000	1.792	1.015	0.010	6.180
Wave Factor (knots)	-0.308	0.622	-99.000	2.990	-0.339	0.585	-4.110	1.910
Current Speed (m/s)	0.547	0.467	0.019	5.251	0.560	0.466	0.019	4.102
Current Factor (knots)	-0.085	0.499	-44.000	64.000	-0.091	0.480	-2.752	2.650

The classification of voyages resulted in 55.7% being short voyages, 24.4% being medium voyages, and 19.9% being long voyages, as shown in Figure 5. The reason for this classification is its usefulness when splitting the data for the stratified cross-validation method presented in Section 3.5.

Vessel behavior was analyzed using k-means clustering applied to key operational variables: fuel consumption per nautical mile, SOG, and RPM, as detailed in Section 2.3.3. Figure 6 presents the relationship between fuel consumption per nautical mile and SOG across vessel size categories, with the three clusters identified within each category. In developing these clusters, we also tested additional inputs such as vessel size and engine model, but the most robust and meaningful separation of operational regimes was achieved without them. Based on tests using the Silhouette coefficient, which measures cluster cohesion and separation, three clusters were identified as the most suitable configuration. This clustering (Figure 6) reveals distinct operational patterns that reflect variability in vessel performance and fuel efficiency among vessels. The green cluster corresponds to the most fuel-efficient operating regime, characterized by relatively low fuel consumption at moderate speeds. These conditions are typically associated with optimized sailing performance. In contrast, the red cluster represents less efficient behavior, where vessels operate at lower speeds but exhibit higher fuel consumption. This pattern may reflect suboptimal conditions such as increased hull resistance or unfavorable sea states. The blue cluster lies between these two extremes, capturing a more balanced

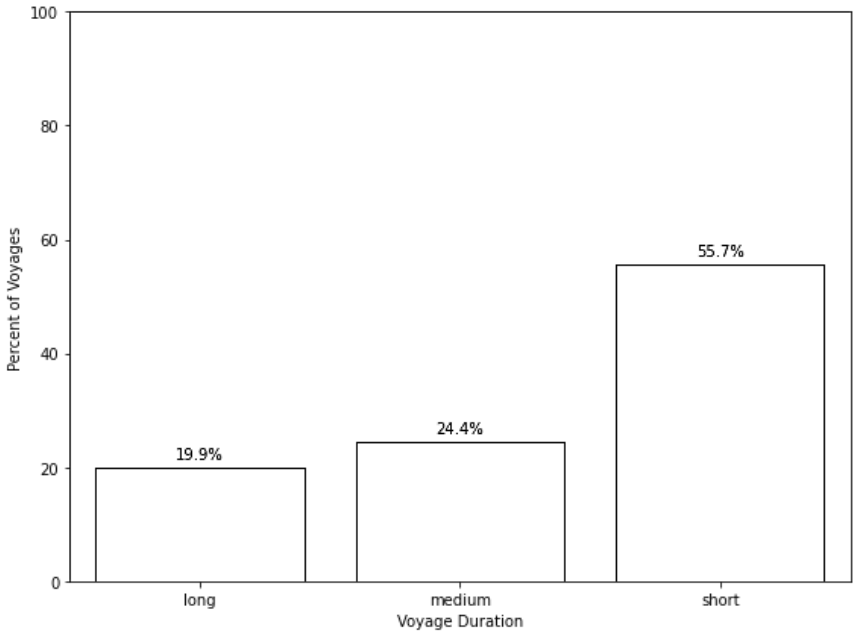


Figure 5: The distribution of voyage duration across voyages.

operational profile with moderate consumption and speed levels. These operational regimes are observed across all vessel size categories, though their distribution varies in range and density. Figure 7 complements this by showing the percentage distribution of noon reports across clusters, providing insights into the relative frequency of each behavioral mode in the dataset.

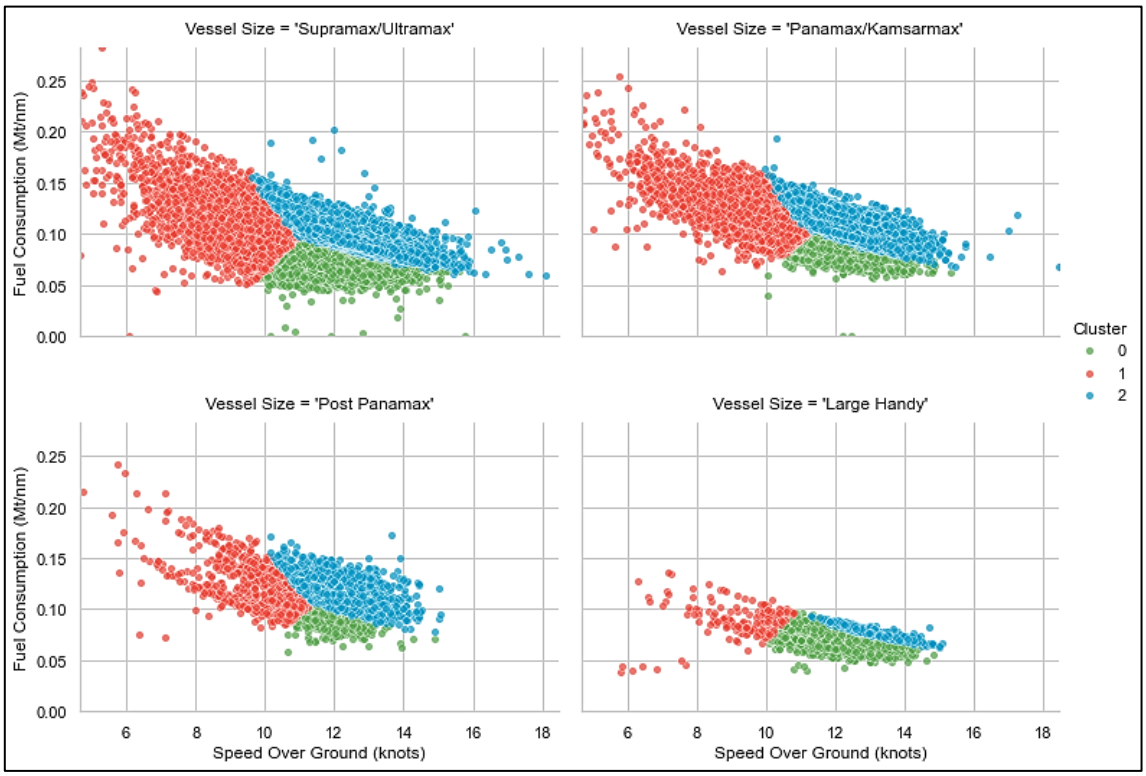


Figure 6: Fuel Consumption per nautical mile vs SOG by Vessel Size and Cluster.

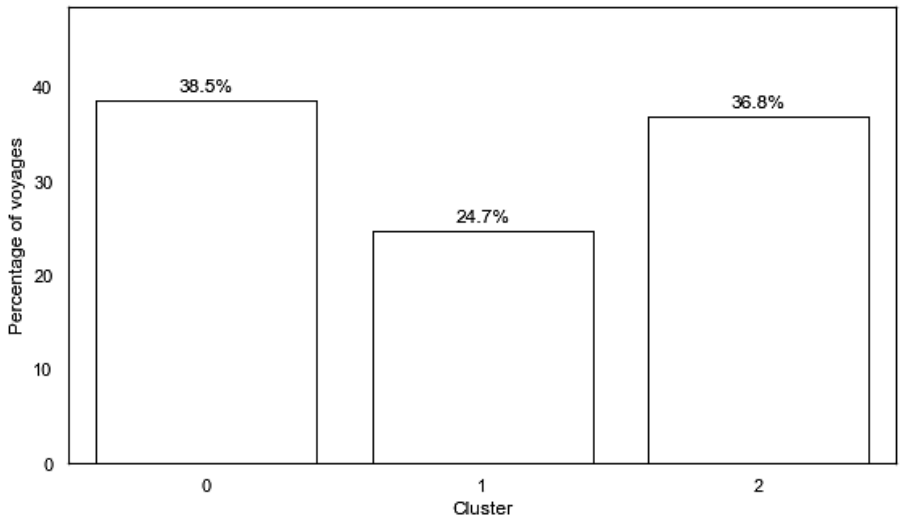


Figure 7: Percentage Distribution of Noon Reports across k-means Behavioral Clusters.

During the feature selection phase, we aimed to identify the most relevant predictors of fuel consumption per nautical mile. As shown in the correlation matrix (Figure 8), SOG was prioritized due to its significant association with fuel consumption. To account for environmental influences, we included the aggregated 24-hour weather factors for wind, waves, and current, which collectively capture the combined effects of direction and intensity (i.e., speed or height) of these external forces, as previously described. In addition, we incorporated the vessel's load condition (laden or ballast), which impacts fuel consumption per nautical mile. Given the available dataset, this selection represents the optimal combination of features for predictive performance.

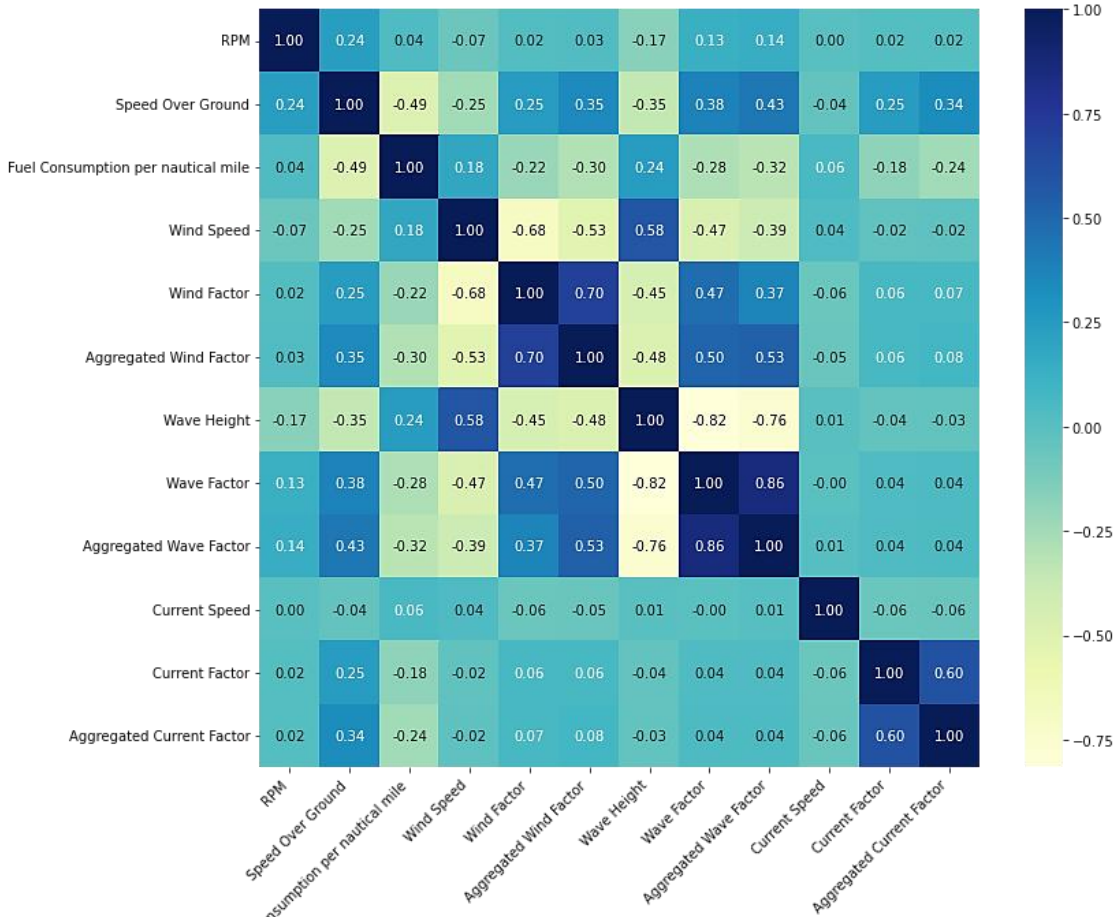


Figure 8: Correlation matrix

5.2. Modeling

The first model was trained for the different clusters of vessels, using MLR from the Scikit-learn library (Pedregosa et al. 2011). This process produced regression coefficients, detailed in Table 2. The second model was also developed and solved using AMPL (Fourer et al. 1990), a modeling language specifically designed for formulating and solving mathematical optimization problems. The regression coefficients resulting from this model are also presented in Table 2. Both models were evaluated using stratified 5-fold cross-validation. As shown in Figure 9, which illustrates one of the test folds, this ensures consistent distributions of voyage durations and clusters across folds.

Table 2. Regression coefficients of the developed models.

	MLR model			Mathematical model		
	Cluster 0	**Cluster 1**	**Cluster 2**	**Cluster 0**	**Cluster 1**	**Cluster 2**
Intercept	1.05×10^{-1}	1.48×10^{-1}	1.61×10^{-1}	8.03×10^{-2}	1.00×10^{-1}	1.14×10^{-1}
SOG²	-1.82×10^{-4}	-4.77×10^{-4}	-3.72×10^{-4}	-5.26×10^{-3}	-7.43×10^{-3}	-1.81×10^{-2}
Aggregated Wind Factor	-7.20×10^{-5}	-1.17×10^{-2}	-3.99×10^{-3}	1.37×10^{-3}	-1.56×10^{-3}	-1.93×10^{-3}
Aggregated Wave Factor	-1.65×10^{-3}	-2.85×10^{-3}	-1.55×10^{-3}	-9.27×10^{-4}	5.74×10^{-4}	-9.46×10^{-4}
Aggregated Current Factor	-1.66×10^{-3}	-4.16×10^{-3}	-1.59×10^{-3}	-5.98×10^{-4}	-8.85×10^{-4}	-2.64×10^{-3}
Load	2.25×10^{-3}	5.10×10^{-3}	-8.80×10^{-4}	1.91×10^{-3}	-8.85×10^{-4}	-2.80×10^{-3}

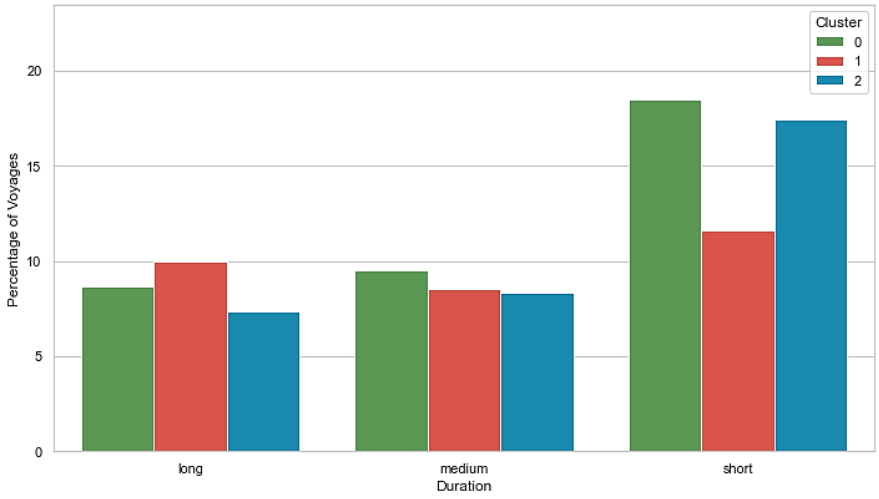


Figure 9: The distribution of voyages per clusters and voyage duration.

5.3. Evaluation

As previously mentioned, the assessment of our model's performance was conducted using the stratified 5-fold cross-validation technique, based on MAE, RMSE, and MAPE. By applying this technique, we obtained these types of errors for each model, for each cluster, and for each fold of our dataset. To provide a representation of these errors at the

cluster level, we averaged the errors across the five folds for each cluster, for both the MLR model and the mathematical model, as shown in Table 3. The values allow for a comparative analysis of the two modeling approaches across clusters. Overall, the errors remain relatively low across all clusters, confirming the consistency of both models. Notably, the mathematical model performs slightly better in Cluster 0 in terms of MAPE (10.7% vs. 10.82%) but shows a marginally higher RMSE in Clusters 1 and 2. In general, the performance differences between the models are small, but the MLR model tends to yield slightly lower errors. These results suggest that both models are capable of capturing the underlying fuel consumption patterns, with minor variations depending on the cluster characteristics.

Table 3. A summary of the MAE, RMSE and MAPE errors of the developed models across the entire dataset.

	MLR model			**Mathematical model**		
	Cluster 0	**Cluster 1**	**Cluster 2**	**Cluster 0**	**Cluster 1**	**Cluster 2**
MAE(MT/NM)	8.29×10^{-3}	1.32×10^{-2}	8.49×10^{-3}	8.31×10^{-3}	1.41×10^{-2}	9.04×10^{-3}
RMSE (MT/NM)	1.02×10^{-2}	1.77×10^{-2}	1.07×10^{-2}	1.02×10^{-2}	1.93×10^{-2}	1.14×10^{-2}
MAPE (%)	10.82	12.24	8.29	10.7	12.53	8.90

After evaluating the overall performance of the models across various data subsets and their generalization to unseen data, we proceeded with a more detailed analysis. We calculated the MAE_v, RMSE_v and MAPE_v between the actual and predicted consumption of each voyage v, utilizing regression coefficients derived from training the models on the entire dataset (Table 2).

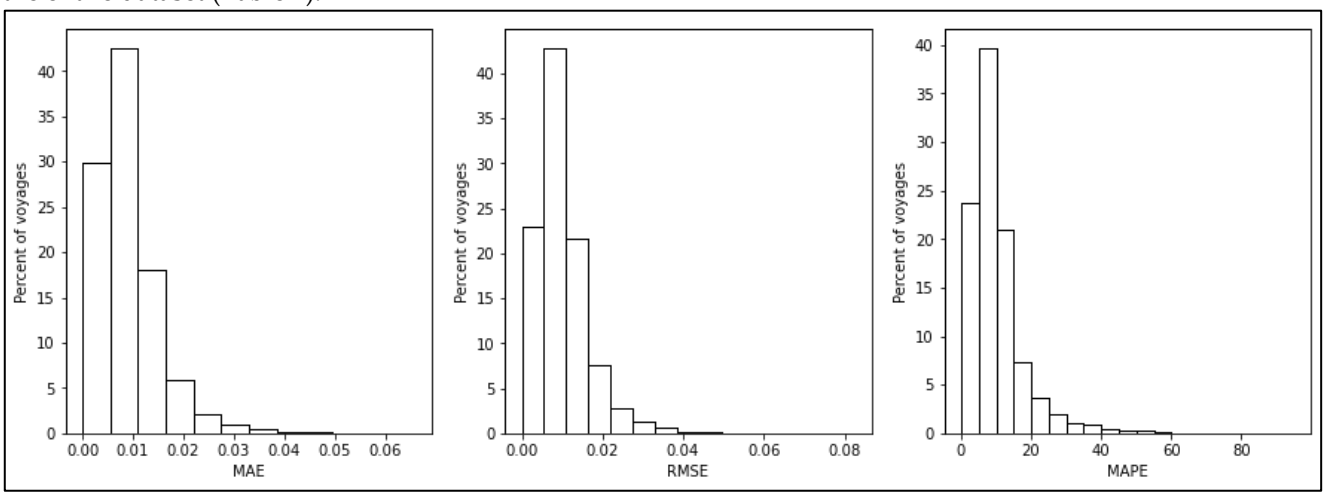


Figure 10: Distribution of errors across voyages for the MLR model.

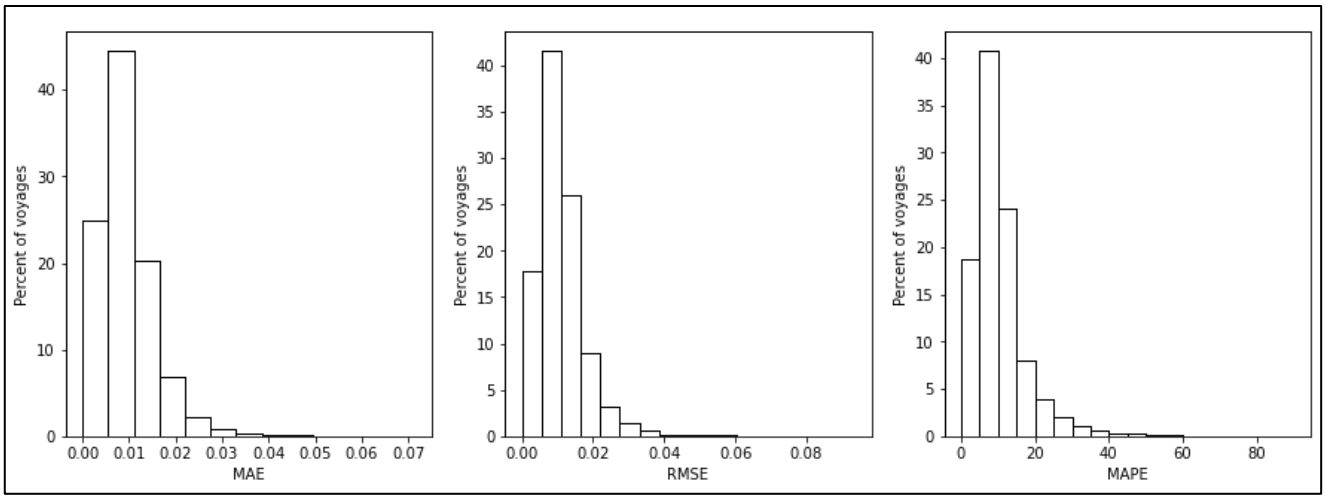


Figure 11: Distribution of errors across voyages for the mathematical model.

For the MLR model, 94.54% of the voyages have a MAE error below 0.02 MT/NM, 92.98% show an error level under the same threshold of 0.02 MT/NM, and 91.51% of the voyages have a MAPE lower than 20% (Figure 10). Meanwhile, the mathematical model yielded the following error rates: 94.80% of the voyages have a MAE error below 0.02 MT/NM, 91.98% present an RMSE level under the same threshold of 0.02 MT/NM, and 91.48% of the voyages have a MAPE lower than 20% (Figure 11).

To interpret feature contributions in the MLR model, we employed SHAP (SHapley Additive exPlanations) values, which quantify the marginal contribution of each feature i to the prediction relative to a baseline. Formally, for each observation:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S) \right],$$

where F is the set of all features, S is a subset of features excluding feature i, x_S represents the values of features in S for a given observation, and $f_S(x_S)$ denotes the model's output given only the features in subset S. Global feature importance is summarized by the mean absolute SHAP value:

$$I_i = \frac{1}{n} \sum_{j=1}^{n} |\phi_{ij}|,$$

where n is the number of observations and ϕ_{ij} is the SHAP value of feature i for observation j. Higher I_i values indicate stronger influence on predictions. Figure 12 presents the mean absolute SHAP values for each feature by category, highlighting the predominant influence of the squared term of speed over ground (SOG^2) on the prediction of fuel consumption per nautical mile, while demonstrating that the relative contributions of other features vary across categories.

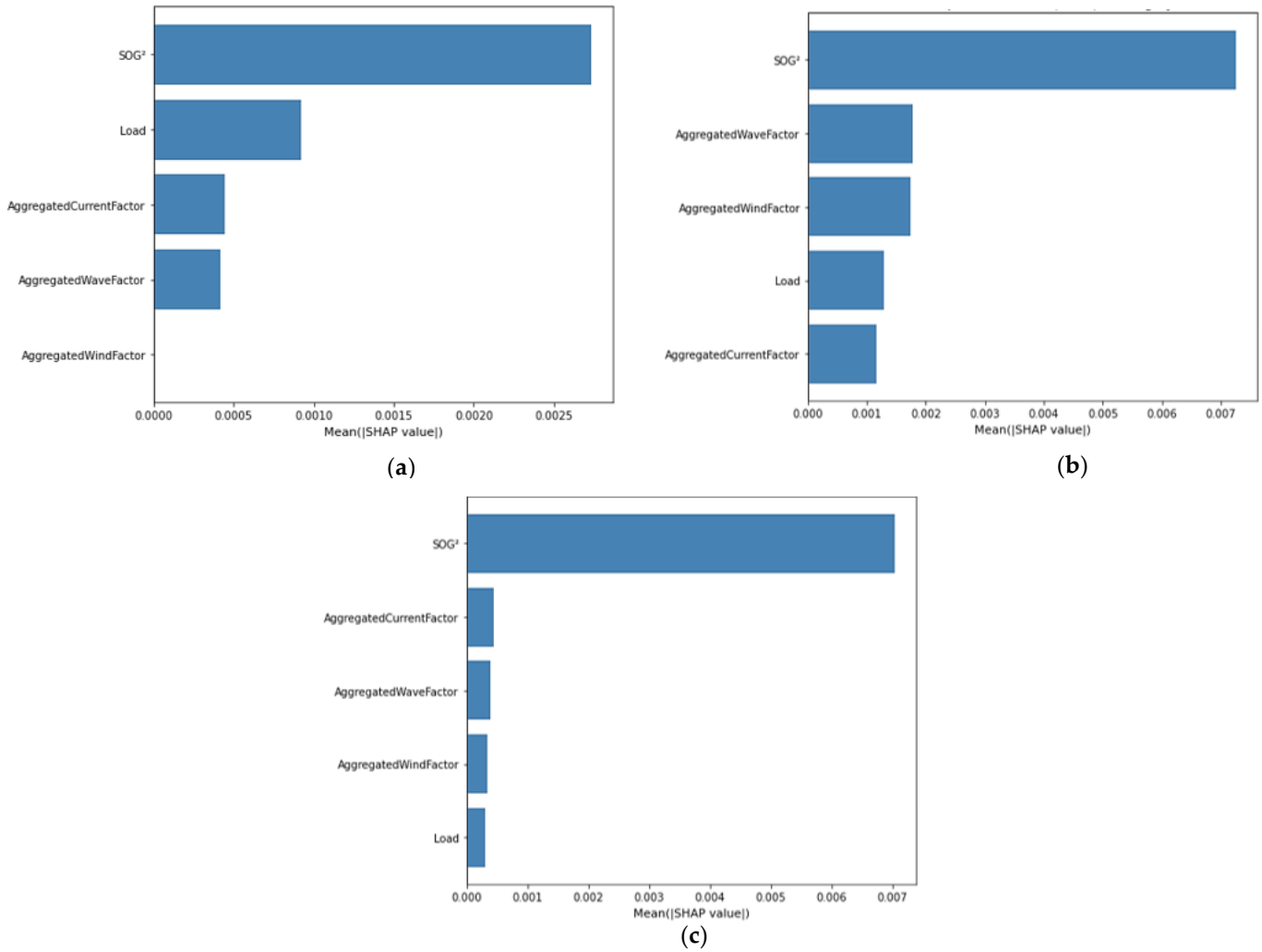


Figure 12: Feature Importance Across Three Clusters: (a) Cluster 0, (b) Cluster 1, and (c) Cluster 2, Based on Mean Absolute SHAP

Furthermore, we aimed to evaluate the total fuel consumption for each voyage rather than solely focusing on the average fuel consumption per nautical mile. Therefore, we analyzed the residuals of the models at the voyage level using the following relation:

$$Residual_v = \sum_{i \in I_v} d_i \, (f_i - \hat{y}_i) \quad , \tag{9}$$

where f_i is the recorded fuel consumption at the noon-report i of the subset of noon reports of the voyage v, I_v. The distance sailed since the previous noon report is denoted as d_i and the predicted fuel consumption at the noon report i as $\hat{y}_i$.

The distribution of residuals for the MLR model shows that 86% of the voyages have residuals concentrated between -50 and +50 metric tons, while the mathematical model presents 87% of the voyages for the same range of residuals with more concentration around zero (Figure 13).

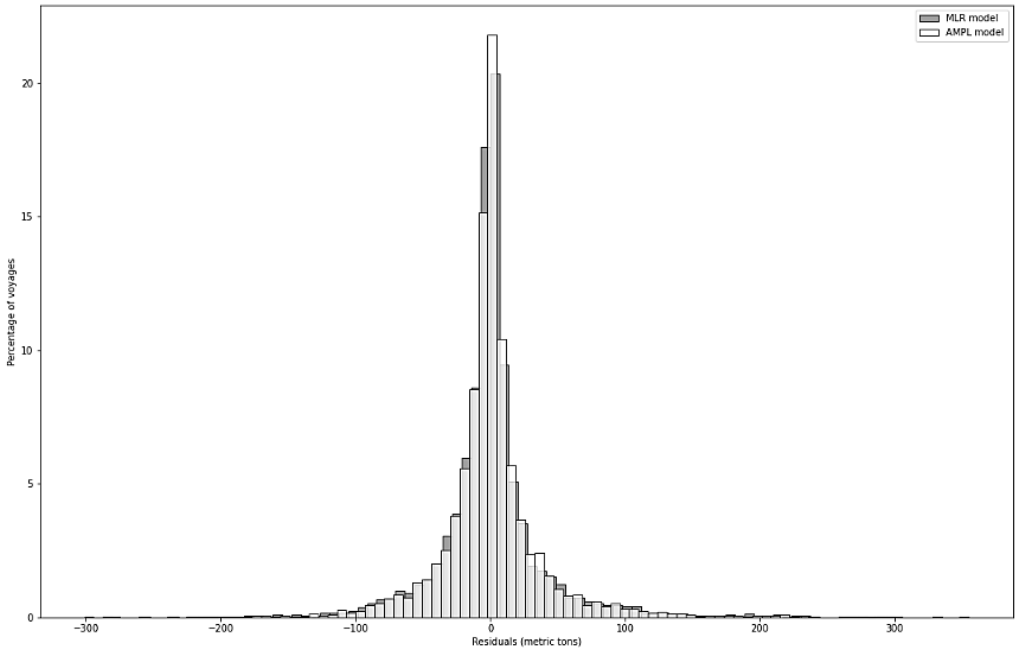


Figure 13: Comparison of the residuals' distribution between the actual and predicted total fuel consumption per voyage for the MLR and mathematical (AMPL) model.

Since we have clustered the dataset during data preprocessing stage, it is interesting to compare the results with and without this application. Indeed, the MLR model without clustering showed that 79.76% of voyages have an MAE lower than 0.02 MT/NM and 77.63% have an RMSE lower than the same threshold, while 78.05% of voyages have a MAPE lower than 20% (Figure 14). As for the mathematical model, it indicates that 80.54% of voyages have an MAE lower than 0.02 MT/NM, then 77.58% have their RMSE lower than the same threshold, and finally 77.83% of voyages have a MAPE lower than 20% (Figure 15). This suggests that while the mathematical model performs slightly better overall and demonstrates strong generalizability across varied data distributions, clustering remains an important consideration. Specifically, data clustering improves the accuracy of predictions for both models by ensuring that the data

is segmented into more homogeneous groups, enhancing the model's ability to learn and adapt to specific patterns within each cluster.

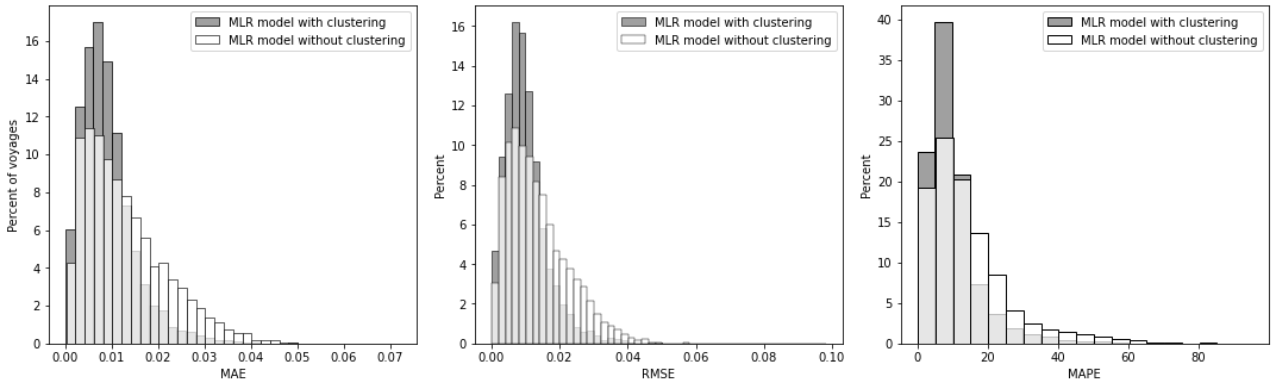


Figure 14: Comparison of the error distribution among voyages for the MLR model with and without the clustering.

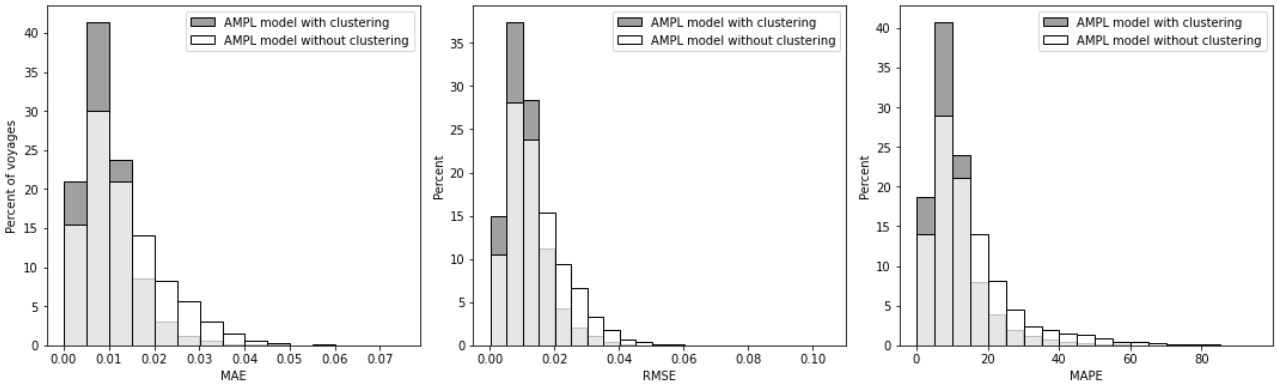


Figure 15: Comparison of the error distribution among voyages for the mathematical (labelled AMPL) model with and without the clustering.

6. Discussion

By training and testing our models on different folds, our error level did not exceed 13%, thus demonstrating the ability of our models to predict unseen data. Moving to the voyage level, the results showed that over 94% of voyages had errors of less than 0.02 MT/NM for both models, whereas our dataset had a minimum fuel consumption per nautical mile value of 0.05, a mean of 0.096, and a maximum of 0.283 (Table 1). It is noteworthy that at this level, the performance of the MLR model and the mathematical model is nearly the same.

Regarding the total fuel consumption per voyage, we observed that the residuals of both models were mostly concentrated around ± 50 metric tons while the majority of residuals fell within the range of approximately ± 200 metric tons. The distribution of errors was nearly the same, except that the mathematical model had more concentration around zero compared to the MLR model. This can be explained by the fact that the mathematical model has two objectives: the first is to minimize the error at the noon report level, which is the same as MLR, and the second is to minimize the error when it comes to total fuel consumption per voyage. The improved performance of the mathematical model in this regard highlights the importance of the weighting parameter β, which controls the trade-off between these two objectives. A large value of β increases the emphasis on matching the total voyage-level fuel consumption, while a smaller β focuses the optimization more on minimizing local errors at the noon report level.

Another crucial aspect to consider is the impact of clustering. When we compare the models, a significant rise in errors is noticed when clustering is not performed. This arises from the fact that vessels present varying behaviors during voyages, influenced not only by external factors but also by their intrinsic physical attributes. By incorporating clustering into our analysis, we enable the predictive models to account for distinct operational patterns emerging from the behavioral data. This approach allows for more precise estimation of fuel consumption, as the models are calibrated to the typical behavior of each cluster—capturing differences in vessel speed profiles, environmental conditions, and loading status that would otherwise be averaged out in a unified model.

A comparison of the coefficients between the two modeling approaches reveals consistent trends but distinct sensitivities. The mathematical models exhibit larger absolute values for SOG^2 and the aggregated weather factors, indicating a stronger response of fuel consumption to speed and hydrodynamic resistance, whereas the MLR models show smoother, data-driven variations. The MLR coefficients remain more sensitive to the empirical data structure and the inverted sign convention of the weather factors, while the mathematical formulation preserves physical directionality through its voyage-level optimization constraint.

Across clusters, these patterns align with expected operational regimes. The negative SOG^2 coefficients in both models confirm that higher vessel speeds are associated with lower fuel consumption per nautical mile, reflecting improved engine efficiency and shorter voyage duration per unit distance. This inverse relationship captures the practical observation (Figure 6) that, within normal operating speeds, the increase in hourly fuel rate is more than compensated by the greater distance covered per hour. Weather-related coefficients (wind, wave, and current) are consistently negative due to the inverted data convention, signifying that more adverse conditions lead to higher fuel use. Their magnitudes diminish from Cluster 0 to Cluster 2, indicating reduced environmental influence under more stable or optimized operations. Variations in load coefficients highlight differing operational behaviors across clusters. Overall, the mathematical model emphasizes physically coherent relationships, while the MLR model reproduce the same tendencies with smoother empirical sensitivity.

The proposed models successfully predicted fuel consumption for the majority of voyages, despite the observed errors. However, it is important to acknowledge the complexity of factors influencing fuel consumption for bulk carriers, some of which may not be captured in our input data. Moreover, the process of collecting data manually is inherently susceptible to various biases and inaccuracies. Captains may record data inconsistently or subjectively, influenced by factors such as workload, environmental conditions, and individual judgment. Additionally, contractual obligations or incentives tied to reported performance metrics can introduce further biases, potentially skewing the accuracy of the data. These limitations in data quality and integrity pose significant challenges for predictive modeling efforts, as inaccuracies or biases in the input data can propagate through the modeling process, leading to errors in predictions. The analysis is limited by its focus on bulk carriers, which may not take into account fuel consumption patterns specific to other vessel types. This limitation highlights the need to expand the dataset to include a variety of vessel types in order to achieve a more comprehensive analysis. In addition, comparison of our models with others is difficult due to differences in input data and methodologies, impacting on the ability to directly assess and validate model performance in different contexts.

7. Conclusion

Maritime shipping is a sector that presents many challenges due to the number of actors involved and the complexity of the operations that are interconnected. The environmental side adds to these challenges and requires more attention in view of the new IMO regulations and the targets it has set.

Controlling and estimating fuel consumption is one of the important keys to reducing emissions from this sector. The present study proposed two approaches for the prediction of fuel consumption for bulk carriers. In both approaches, meteorological conditions were added as an important factor affecting the performance and the energy efficiency of ships during sailing. We first proposed a MLR model that considers bulk carriers with different physical characteristics and thus different behaviors during voyages. Then, we introduced a new approach by incorporating considerations on the total fuel consumption of bulk carriers per voyage. By proposing both models, we aim to provide an generalizable response to the various needs and challenges encountered in forecasting bulk carrier fuel consumption in different operational contexts and ship categories.

The training and testing were performed on a large dataset, which allowed better learning of the variations of the fuel consumption under different meteorological conditions. The results show that both the mathematical model and the MLR model accurately predicted fuel consumption for over 90% of voyages. A comparison between the two models indicated that the mathematical model slightly outperformed the MLR model in predicting total fuel consumption. While this difference was marginal in the current analysis, it suggests that the mathematical model could potentially offer better performance under different conditions, such as changes in the dataset, improvements in data quality, or variations in the factors influencing fuel consumption. Moreover, we explored the impact of clustering data before modeling, highlighting the importance of this step in enabling our models to capture the diverse patterns and operational behaviors of bulk carriers observed during voyages.

In conclusion, while our models have shown promising performance in predicting fuel consumption for bulk carriers, there are avenues for further enhancement. By incorporating additional inputs that better explain the factors influencing fuel consumption and by transitioning from manual data collection to more advanced methods, such as

automated sensors, we can improve the accuracy and reliability of our predictions. Additionally, exploring more advanced modeling techniques could offer insights into more complex relationships within the data. However, in our case, a lower level of model complexity was necessary to integrate the predictions into an optimization tool for bulk carriers, where the final weights of the inputs were required. Hence, striking a balance between model complexity and practical applicability was crucial in achieving our objectives.

These models have significant potential for practical application across multiple aspects of maritime routing decision-making. While trained on a specific dataset, the modeling approach can be generalized to other datasets with similar features, enabling broader applicability across different fleets and voyage profiles. Specifically, they can be integrated into weather routing systems, which are crucial for optimizing fuel consumption and minimizing operational costs during voyages. Furthermore, the models can be applied in fleet management systems, where they can provide insights on fuel efficiency across vessels or clusters of behavior, supporting performance benchmarking, operational optimization, and informed decisions on maintenance scheduling or retrofitting. In port operations, the models can predict fuel consumption during approach and departure phases, thereby improving port call scheduling. Computationally efficient, these models can be deployed on existing shipboard systems, providing real-time predictions with minimal infrastructure changes, making them ideal for operational environments that require timely decision-making.

8. Acknowledgments

The authors express their gratitude to True North Marine, for supplying the data and verifying the results presented in this article, and acknowledge the support of PROMPT and MITACS, which funded this project. The authors also thank two anonymous referees for their valuable comments, which helped improve the quality and clarity of this paper.

References

Armstrong, V. N. (2013). Vessel optimisation for low carbon shipping. Ocean Engineering, 73, 195-207.

Bialystocki, N., & Konovessis, D. (2016). On the estimation of ship's fuel consumption and speed curve: A statistical approach. Journal of Ocean Engineering and Science, 1(2), 157-166.

Chen, Z. S., Lam, J. S. L., & Xiao, Z. (2024). Prediction of harbour vessel emissions based on machine learning approach. Transportation Research Part D: Transport and Environment, 131, 104214.

Fourer, R., Gay, D. M., & Kernighan, B. W. (1990). A modeling language for mathematical programming. Management Science, 36(5), 519-554.

Godet, A., Wallner, L. J. M., Panagakos, G., & Barfod, M. B. (2024). Developing correction factors for weather's influence on the energy efficiency indicators of container ships using model-based machine learning. Ocean & Coastal Management, *258*, 107390.

Hagiwara, H., & Spaans, J. A. (1987). Practical weather routing of sail-assisted motor vessels. The Journal of Navigation, 40(1), 96-119.

Hagiwara, H. (1989). Weather routing of (sail-assisted) motor vessels. Ph. D. thesis, Delft Univ. Tech.

Hajli, K., Rönnqvist, M., Dadouchi, C., Audy, J. F., Cordeau, J. F., Warya, G., & Ngo, T. (2024). A fuel consumption prediction model for ships based on historical voyages and meteorological data. Journal of Marine Engineering & Technology, 1-12.

International Maritime Organization (IMO). (2018). Initial IMO strategy on reduction of GHG emissions from ships. https://www.imo.org/en/MediaCentre/HotTopics/pages/reducing-greenhouse-gas-emissions-from-ships.aspx

Kepaptsoglou, K., Fountas, G., & Karlaftis, M. G. (2015). Weather impact on containership routing in closed seas: A chance-constraint optimization approach. Transportation Research Part C: Emerging Technologies, 55, 139-155.

Kee, K. K., Simon, B. Y. L., & Renco, K. H. Y. (2018). Prediction of ship fuel consumption and speed curve by using statistical method. Journal of Computer Science & Computational Mathematics, 8(2), 19-24.

Kim, M., Hizir, O., Turan, O., Day, S., & Incecik, A. (2017). Estimation of added resistance and ship speed loss in a seaway. Ocean Engineering, 141, 465-476.

Kobayashi, E., Asajima, T., & Sueyoshi, N. (2017). 19. Advanced Navigation Route Optimization for an Oceangoing Vessel. Methods and Algorithms in Navigation: Marine Navigation and Safety of Sea Transportation, 149.

MEPC, R. (2022). 2022 Guidelines for the development of a Ship Energy Efficiency Management Plan (SEEMP).

Merien-Paul, R. H., Enshaei, H., & Jayasinghe, S. G. (2018). In-situ data vs. bottom-up approaches in estimations of marine fuel consumptions and emissions. Transportation Research Part D: Transport and Environment, 62, 619-632.

Ormevik, A. B., Fagerholt, K., Meisel, F., & Sandvik, E. (2023). A high-fidelity approach to modeling weather-dependent fuel consumption on ship routes with speed optimization. Maritime Transport Research, 5, 100096.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. Journal of machine learning research, 12, 2825-2830.

Prpić-Oršić, J., Vettor, R., Faltinsen, O. M., & Guedes Soares, C. (2015). Influence of ship routes on fuel consumption and CO2 emission. Maritime Technology and Engineering. Taylor & Francis Group, London.

Psaraftis, H. N., & Lagouvardou, S. (2023). Ship speed vs power or fuel consumption: Are laws of physics still valid? Regression analysis pitfalls and misguided policy implications. Cleaner Logistics and Supply Chain, 7, 100111.

Simonsen, M. H., Larsson, E., Mao, W., & Ringsberg, J. W. (2015). State-of-the-art within ship weather routing. In International Conference on Offshore Mechanics and Arctic Engineering (Vol. 56499, p. V003T02A053). American Society of Mechanical Engineers.

Szlapczynska, J. (2015). Multi-objective weather routing with customised criteria and constraints. The Journal of Navigation, 68(2), 338-354.

Uyanık, T., Karatuğ, Ç., & Arslanoğlu, Y. (2020). Machine learning approach to ship fuel consumption: A case of container vessel. Transportation Research Part D: Transport and Environment, 84, 102389.

Wang, S., Ji, B., Zhao, J., Liu, W., & Xu, T. (2018). Predicting ship fuel consumption based on LASSO regression. Transportation Research Part D: Transport and Environment, 65, 817-824.

Yang, H., Sun, Z., Han, P., & Ma, M. (2024). Data-driven prediction of ship fuel oil consumption based on machine learning models considering meteorological factors. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, 238(3), 483-502.

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.

Zis, T. P., Psaraftis, H. N., & Ding, L. (2020). Ship weather routing: A taxonomy and survey. Ocean Engineering, 213, 107697.