A fuel consumption prediction model for ships based on historical voyages and meteorological data

Kaoutar Hajli*[a], Mikael Rönnqvist[a], Camélia Dadouchi[b], Jean-François Audy[c], Jean-François Cordeau[c], Gurjeet Warya[e], Trung Ngo[e]

[a]*Department of Mechanical Engineering, Université Laval, Québec, Canada;* [b]*Department of Mathematical and Industrial Engineering, Polytechnique Montréal, Montréal, Canada;* [c]*Department of Management, Université du Québec à Trois-Rivières, Trois-Rivières, Canada;* [d]*Department of Logistics and Operations Management, HEC Montréal, Montréal, Canada;* [e]*True North Marine, Montréal, Canada;*

A fuel consumption prediction model for ships based on historical voyages and meteorological data

Predicting the fuel consumption of a ship during a voyage is a challenging task, given the internal and external factors that influence it. This challenge has gained crucial importance in light of the regulations imposed by the International Maritime Organization, which aim to reduce greenhouse gas emissions from ships. The objective of this study is to develop a fuel consumption prediction model using data collected from bulk carriers. These predictions are to be used as input for a ship routing tool. We propose a predictive model of fuel consumption for these bulk carriers using a multiple linear regression model considering the propeller rotational speed and the speed loss due to wind, waves and currents. The results show that the estimated fuel consumption of the studied bulk carriers is strongly affected by the engine setting and the meteorological conditions. The developed model can predict fuel consumption accurately for more than 80% of the voyages of the dataset with a mean absolute error and a root of the mean squared error lower than 0.01 metric ton per nautical mile, and a mean absolute percentage error of less than 15%, making it useful for ship routing purposes.

Keywords: green routing; ship energy efficiency; fuel consumption prediction; artificial intelligence; machine learning.

1. Introduction

Maritime transport is a key commercial area which accounts for more than 80 % of international freight transport. This sector is under high international pressure to reduce greenhouse gas emissions. As of April 2018, the International Maritime Organization (IMO) has adopted an initial strategy (Joung et al. 2020) regarding reducing greenhouse gas (GHG) emissions from vessels. This strategy aims, firstly, to reduce CO2 emissions per transport activity by at least 40% by 2030 compared to 2008, and secondly, to reduce the total volume of annual GHG emissions by at least 50% by 2050 compared to 2008. In the same perspective, and since January 2020, the IMO has imposed a new regulation that targets the reduction of sulfur oxide (SOx) emissions by lowering the limit of sulfur content in fuel oil used by vessels from 3.5% to 0.5%. This and other future IMO regulations have a significant impact on shipping costs and further motivate the optimization of the performance of vessels.

In order to control and optimize their performance, ships often rely on systems that provide weather routing. Zis et al. (2020) defined ship weather routing as the decision-making process that aims at selecting the optimal route in a given voyage. The optimality of the selected route depends on the chosen objectives, which can be the minimization of costs, voyage duration, or the reduction of potential delays and risks, all taking into account the forecasted meteorological conditions. Since fuel consumption represents about two-thirds of the cost of a voyage and more than a quarter of the overall operating costs of a vessel (Stopford 2008), its integration into the objectives to be optimized is crucial.

In the existing literature, ship fuel consumption (SFC) models have been widely studied and classified into three distinct categories: the white box model (WBM), the black box model (BBM), and the gray box model (GBM). The WBM is based on the mechanism analysis, which includes the statics and dynamics of the vessel and can be divided into floating state and hydrodynamics analysis. It computes the resistances to which the ships are subjected during sailing and then converts them into fuel consumption according to the relationship between the ship and engine-propeller. For example, Tillig and Ringsberg (2019) presented a quasi-static simulation model based on empirical methods and standardized numerical hull and propeller series, which solves the force and moment balance for four degrees of freedom (surge, drift, yaw and heel). The article showcases the model's practicality through case studies involving two ships: a tanker and a PCTC, operating on a route in the Baltic Sea. Additionally, Fan et al. (2020) relied on the Monte Carlo simulation method to simulate the energy efficiency by using fuel consumption data of the studied bulk carrier ship to verify the model. Constructing white box models requires a good knowledge in the field and access to technical details, which are frequently challenging to obtain.

The second category concerns the BBM. This type of models is based on data analysis and uses different methodologies of statistical and machine learning models. For instance, Wang and Ji (2018) presented a novel ship fuel consumption prediction model, which utilizes the least absolute shrinkage and selection operator (LASSO) regression algorithm. This model incorporates 20 input variables, including main-engine status, cargo weight, ship draft, sea-states, weather conditions, and other relevant factors, sourced from data obtained from container ships. Hu et al. (2019) also employed machine learning techniques, namely the Back-Propagation Neural Network (BPNN) and Gaussian Process Regression (GPR), to predict fuel consumption using two datasets from one container ship. The inputs considered in their models include ship shaft, speed, average draft, trim, and weather data. Tarelko and Rudzki (2020) explored the application of artificial neural network (ANN) techniques to model ship speed and fuel consumption, considering a comprehensive set of inputs including speed, displacement, wind force, wind wave height, swell height, sea current factor, and trim.

The last type of models is the GBM which combines aspects from the previously discussed WBM and BBM. The WBM and BBM can be combined either by introducing a WBM-based BBM for relationships that cannot be expressed as equations by the WBM or by introducing a BBM-based WBM model that can be used to check whether the output of the BBM is consistent. Many studies have been conducted on the optimization of ship fuel consumption using GBM as it is the case for Weiqiang and Honggui (2013), who used gray system theory to develop a model for predicting the diesel consumption of diesel generator sets. They analyzed the patterns in fuel consumption as running time varies and compared the predicted diesel consumption obtained from the model with the actual measurements to assess the accuracy of the model. In the same context, Yuan and Nian (2018) developed a Gaussian process metamodel to predict the ship fuel consumption considering the effects of operational conditions such as speed and trim as well as the impacts of weather conditions such as wind and wave effects. By incorporating these factors into their predictive model, they aimed to accurately estimate ship fuel consumption under different scenarios. Coraddu et al. (2017) have also introduced two distinct Gray Box Models (GBMs) that effectively leverage both mechanistic knowledge of the underlying physical principles and available measurements. These GBMs served as predictive models for fuel consumption, employing data specifically obtained from a Handymax chemical/product tanker. Based on these models, the authors proposed a new strategy for optimizing the trim of a vessel.

A recent literature review conducted by Fan et al. (2022) studied articles published between 2001 and 2021 that investigated ship fuel consumption (SFC) models. The review identified a gap in the generalizability of existing SFC models, which are often tailored to the specific characteristics and voyage data of individual ships, limiting their applicability to other vessels. Among the 24 articles analyzed, only Tran (2021), Işıklı et al. (2020), and Fan et al. (2020) focused on bulk carriers, which account for approximately 34.4% of the world's merchant fleet in terms of gross tonnage according to Equasis (2020). The majority of the remaining articles primarily addressed container ships, representing about 18 % of the world's merchant fleet in terms of gross tonnage. Furthermore, out of the three articles that considered bulk carriers, only Fan et al. (2020) incorporated meteorological factors, known to have a significant impact on fuel consumption. In the same literature review, various studies exploring fuel consumption prediction models which are based on regression analysis were examined. Gkerekos et al.

(2019) employed a diverse set of machine learning regression algorithms, relying on various measurements such as the propeller rotational speed, sea state, wind speed, propeller slip, draft, sea current, wind direction, and sea direction for daily fuel consumption prediction. Simonsen et al. (2018) estimated the hourly fuel consumption of cruise ships navigating Norwegian waters, based on Automatic Identification System (AIS) data and technical ship particulars, including service speed, total power, and the number of engines. Kim et al. (2021) developed models based on Artificial Neural Network (ANN) and Multiple Linear Regression (MLR), utilizing an extensive set of features such as the speed over the ground, speed through water, relative wind speed and direction, mean draught, trim, displacement for fuel consumption prediction. Farag and Ölçer (2020) introduces a model combining Artificial Neural Network (ANN) and Multi-regression (M.R) to predict ship fuel consumption across varying sea environments, utilizing inputs like ship's speed, seawater depth, wave parameters, swell parameters, and current.

In the present study, we develop an SFC model based on data from nine sister bulk carriers, sharing the same design and construction features, which presents an opportunity to address this gap in the generalizability of SFC models. Utilizing data from multiple sister bulk carriers, with a focus on the to the propeller rotational speed and the weather factors, significantly enhances the generalizability of the proposed SFC model and proves advantageous in cases where detailed travel information is limited. By capturing the common characteristics and operational patterns among these vessels, the model can effectively encompass a broader range of ships, thereby offering greater generalizability, particularly for a type of ship that has received limited research attention. Additionally, the inclusion of multiple sister ships ensures a representative sample, enhancing the robustness of the model. The derived knowledge and insights from this approach can transcend individual vessels, facilitating the application of the model to similar ships, even those that were not originally part of the dataset. This broader applicability increases the practical utility of the developed SFC model and promotes its adoption in the maritime industry.

Furthermore, it is important to underline that our study goes beyond the conventional consideration of the impact of weather conditions on ship fuel consumption. Our study integrates meteorological factors contributing to ship speed loss, particularly those resulting from wind waves and currents, which therefore takes into account the magnitude/speed of each component, as well as their respective directions, enabling a more comprehensive understanding of the influences on fuel consumption. The scope of our study provides valuable information to the marine industry, offering a comprehensive perspective that goes beyond the usual focus on the magnitude/speed, and direction of weather components.

For this purpose, a large dataset has been prepared, analyzed and used into a multiple linear regression (MLR) model for fuel consumption prediction. These predictions take into account the meteorological forecasts and the historical routes to ensure practical applicability and yield a high accuracy. The results show that the proposed model is able to predict the fuel consumption based on the propeller rotational speed, and the weather conditions of more than 80% of the voyages of the studied bulk carriers with a MAE and a RMSE lower than 10^{-2} metric ton per nautical mile.

The rest of the paper is structured as follows. Section 2 describes the materials and methods employed in this study, including the problem formulation, data preprocessing, and the prediction model. The results obtained from applying this methodology to our dataset are presented in Section 3, followed by a discussion of these results in Section 4. Finally, the conclusion of the study is provided in Section 5.

2. Materials and methods

This section first introduces a general formulation of the problem and then provides the details of the data acquisition, the preprocessing of the data, and the selection of variables for the modeling. Then, details are provided on the model and the solution proposed for this study. A graphical representation of the developed methodology is presented in Figure 1.

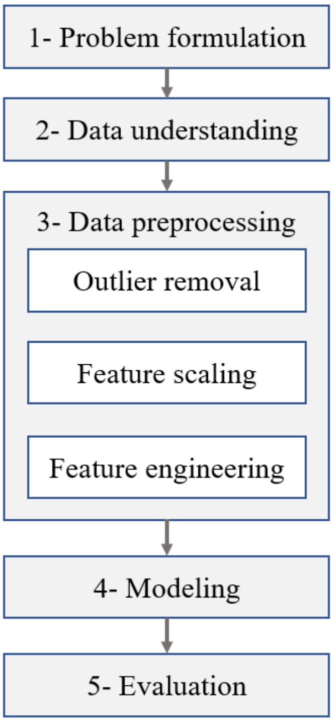

Figure 1: The research methodology overview.

2.1. Problem formulation

This study aims to estimate the fuel consumption of bulk carriers, which is affected by several external and internal factors, including main and auxiliary engines, geometry of ship hull, propeller design, and other parameters. There is a classical relationship between the sailing speed and the fuel consumption per time unit which is:

$$F(v) = \lambda v^n \quad (1)$$

$$\lambda > 0, n > 0,$$

where F represents the daily fuel consumption as a function of the speed v of the ship and λ is a positive scalar constant. The value of n in the formula is usually 3, but in practice, its value depends on the ship type and the speed (Bialystocki and Konovessis 2016). A recent study (Psaraftis and Lagouvardou 2023) reviewed various papers that question the validity of the cubic exponent in equation (1). These studies show that exponents lower than 3, and sometimes even lower than 2 or 1, seem to fit the data better. The study shows that these results, which violate the fundamental laws of hydrodynamics, are due to the fact that many models ignore the dependence between meteorological conditions and other factors, leading to misleading results when using regression analysis. The relationship (1) comes from the proportionality between power, speed, and resistance on one side and the proportionality between the power and the fuel consumption on the other side, as explained in MAN Energy Solutions (2018). It should be noted that our study is based on fuel consumption per distance rather than per time. This preference is explained by the fact that fuel consumption per distance gives a more complete perspective of a vessel's efficiency, taking into account the amount of fuel a vessel uses to cover a given distance, as well as the variability of sailing speeds, and provides a better assessment of a vessel's environmental footprint over the course of a voyage, making it a preferred metric for operational and environmental considerations in the marine industry. Since the speed of the ship is proportional to the propeller rotational speed measured in revolutions per minute (RPM) in calm water, the fuel consumption can be estimated as a function of the RPM too. Furthermore, and in order to approximate fuel consumption in a way that takes into account the actual performance of the vessels, it is essential to include information on weather conditions at sea such as wind, waves, and current (Wang and Meng 2012).

In order to accurately quantify the influence of meteorological conditions, including waves, wind, and current, on the fuel consumption of vessels, it is essential to consider the concept of speed loss, as extensively studied by Kim et al. (2017). In our study, we define the speed loss resulting from the additional resistance attributed to wind, wave, and current as the wind factor, wave factor, and current factor, respectively. The calculation of these factors is based on the diagram shown in the flowchart of Figure 2, where the authors provided a procedure to estimate ship speed loss due to wind and irregular waves in specific sea conditions. It involves estimating the added resistance due to waves (ΔR_{wave}) and wind (ΔR_{wind}), as well as considering variables such as propulsion efficiency (η_D) and transmission efficiency (η_S). The total resistance due to wind and waves (R_T) is calculated by combining the predicted calm water resistance (R_C)

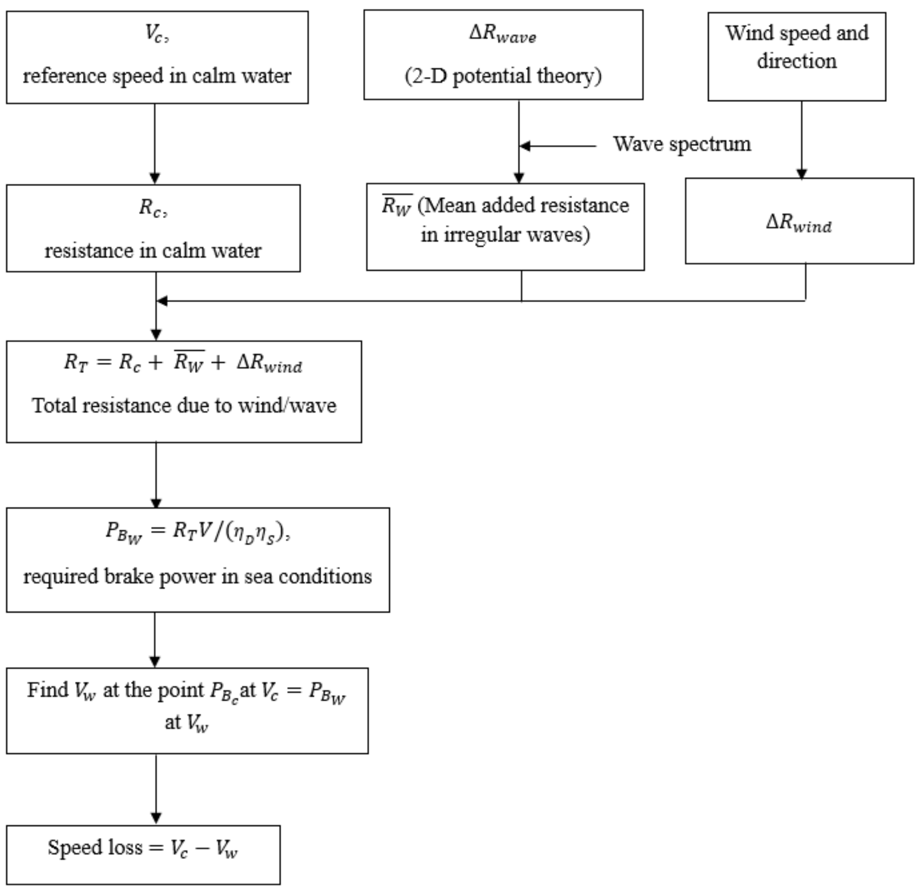

Figure 2: Ship speed loss estimation. Adapted from Kim et al. (2017).

with the estimated added resistances. Finally, the ship speed loss is determined by comparing the achievable ship speed in the specific sea conditions (V_w) with the reference ship speed in calm water (V_c).

The meteorological factors considered in our study hold significant importance as they encompass both the relative heading (see Figure 3) and the speed/height of wind, wave, and current. These factors are combined to yield a representative value for each ship type, and to account for its specific physical characteristics.

An additional factor to consider in our fuel consumption analysis is the propeller slip (Bayraktar and Sokukcu 2023). The ship's propeller, driven by the main engine and transmission equipment, operates according to variables such as RPM, running time and pitch value. These parameters are essential for calculating the theoretical distance the vessel should cover in one revolution of the propeller. However, real-life conditions introduce a series of external factors, including winds, currents, waves, the vessel's

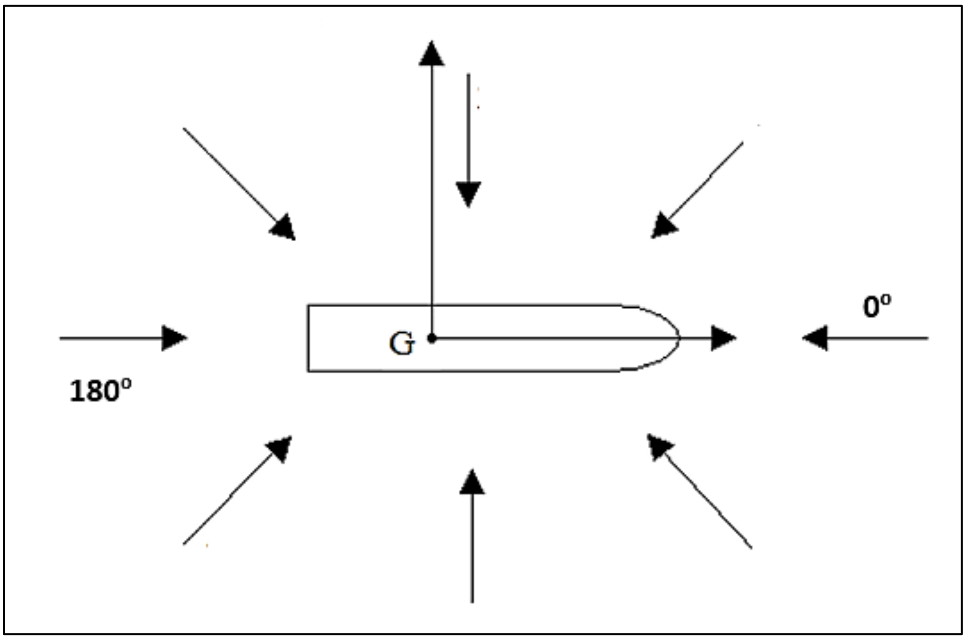

Figure 3: Illustration of weather direction relating to the heading of the vessel.

draught and hull fouling. These factors together contribute to the difference between the theoretical distance and the actual distance covered, which is the propeller slip.

2.2. Data understanding

Data on vessel performance during voyages is often collected either from so-called noon reports or from mounted sensors. In our study, we only deal with the noon reports reported by the captains every 24 hours. These reports have various information that can be divided into three categories:

- Point identification: this category concerns the identification of each voyage between a port of departure and a port of arrival by a unique ID in addition to the identification of the current and future points of the vessel's passage.
- Spatio-temporal context: this category contains information on the latitude and longitude of each point of the vessel's passage in addition to information on the date, time and distance sailed from the port of departure.
- Vessel status: in this category, the captains report the remaining level of each type of fuel (intermediate fuel oil and marine gas oil), the speed over ground, the propeller slip and the propeller RPM at each point of the passage.

In addition, our study benefits from comprehensive meteorological information received every 6 hours from our industrial partner. These meteorological reports provide information on the speed/magnitude and direction of wind, waves and currents, as well as precomputed speed loss values for each meteorological component, determined according to the methodology explained in the previous section. The noon reports and weather reports contain the information detailed in Table 1.

Table 1: Overview of the data

Parameter	Description
VoyageId	ID number given to each voyage between one origin and one destination
PointId	Unique ID number for each row of the database
NextPointId	The following point ID in the order of passage
PointType	The type of point (departure, arrival, noon report)
VesselName	Name of the vessel
Passage	The origin and destination of the voyage
VoyageType	Indicates whether the voyage is laden or ballast
Latitude	Latitude in degrees
Longitude	Longitude in degrees
Distance	The distance travelled since the last point (nautical miles)
Date	The time and date at which the point was crossed by the vessel
TimeElapsedFromPreviousNoon	Number of hours elapsed since the last noon report
DistanceFromPreviousNoon	Distance traveled since the last noon report
IfoRemaining	Quantity of intermediate fuel oil on board (metric tons) reported every 24 hours
MgoRemaining	Quantity of marine gas oil on board (metric tons) reported every 24 hours
RPM	Propeller rotational speed (revolutions per minute) reported every 24 hours
Slip	Propeller slip (%) reported every 24 hours
AverageSpeed	The speed over ground in knots
WindDirection	The direction of the wind reported every 6 hours
WindMagnitude	The speed of the wind in Beaufort reported every 6 hours
WindFactor	The speed loss caused by the wind in knots reported every 6 hours
WaveDirection	The direction of the waves
WaveMagnitude	The height of the waves in meters reported every 6 hours
WaveFactor	The speed loss caused by waves in knots reported every 6 hours
CurrentDirection	The direction of the current reported every 6 hours
CurrentMagnitude	The speed of the current in m/sec reported every 6 hours
CurrentFactor	The speed loss caused by the current in knots reported every 6 hours

2.3.Data preprocessing

Noon reports represent a significant challenge as manual data entry has an impact on the quality of the developed models. The data quality could be low due to several reasons. One reason is that the remaining fuel meter is not accurate. Another is that the captain is not reviewing the current status in detail and may even use information from the previous day. Also, there may be contractual reasons to report inaccurate information to show better (or worse) performance than actual. Therefore, the pre-processing of the data is a crucial step for modeling.

2.3.1. Outlier removal

In this step, we attempt to detect outliers in the data without compromising the relevant information we need. To accomplish this, we start with a visual inspection of the available parameters by plotting the dataset values and examining them for any extreme values that deviate significantly from the majority of the data points. Additionally, considering the specific characteristics of the vessel, such as its type, we also take into account the predetermined minimum and maximum values of, for example, the RPM that the vessel is designed to operate within, in order to identify and handle any data points that fall outside of these predefined limits.

As a statistical method to detect outliers, we use the z-score which indicates the distance between a data point and the mean of the database divided by the standard deviation according to the following formula:

$$z = \frac{x - \mu}{\sigma}, \tag{2}$$

where x is the value of the data point, μ is the mean of the data set, and σ is the standard deviation. Outliers are defined as points whose z-score is smaller or greater than a given threshold.

2.3.2. Feature scaling

Since the parameters of the ships have different ranges and units, it is important to scale them in the preprocessing stage. One of the most common feature scaling techniques is Min-Max normalization. This technique consists in scaling the features so as to have them bounded between 0 and 1 by the following relation:

$$x_m = \frac{x - x_{min}}{x_{max} - x_{min}}\ , \tag{3}$$

where x_m is the normalized value of the actual variable x. The values x_{min} and x_{max} are the minimum and maximum observed values of the variable x, respectively. The processed data is then used to train and validate the model.

2.3.3. Feature engineering

Feature engineering is an important process for data-driven modeling that consists of transforming raw data into features that more accurately represent the problem underlying the predictive model. This process includes two important steps: feature generation, through which new features are generated from existing ones and feature selection, which consists in selecting the relevant features that will serve as inputs for our model.

Given the different frequencies of fuel consumption and meteorological reports, we have chosen to average the meteorological factors for each fuel consumption report by averaging them as shown in Figure 4. This operation aims at obtaining an average representation over 24 hours of the speed loss resulting from wind, waves and currents explained by the meteorological factors. At the end of this process, we have obtained three new features: averaged wind factor, averaged wave factor and averaged current factor.

During the feature selection phase, we use a correlation coefficient, in particular Pearson's correlation coefficient, to determine which features affect our target variable, which is fuel consumption per nautical mile. The Pearson correlation coefficient is a statistical measure that evaluates the strength of the linear relationship or association between two continuous variables. It provides indications of the strength (magnitude) and direction of this correlation, helping us to make informed decisions on feature selection.

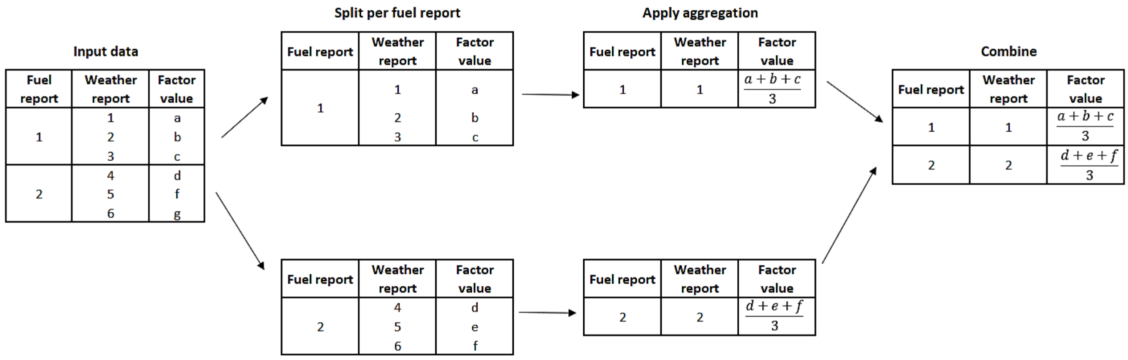

Figure 4: Example of data averaging in the case of three meteorological reports in each fuel consumption report.

This methodology also plays a crucial role in preventing multicollinearity between our input variables, ensuring that our results remain unbiased and reliable.

2.4.Prediction model

Since our aim is to develop a predictive model that will be integrated into a weather routing system and, consequently, to have weights for each input to predict fuel consumption per distance unit, we are using regression analysis in our study.

Regression analysis is among the most common statistical methods used for fuel consumption prediction in the literature (Fan et al. 2022). The multiple linear regression (MLR) is used to estimate the regression coefficients $\alpha_0, \alpha_1, \ldots, \alpha_k$, of equation (4), corresponding to each of the selected inputs in the feature selection step by minimizing the sum of squared residuals (SSR) according to the formula (5).

$$y_i = \alpha_0 + \alpha_1 x_{1i} + \cdots + \alpha_k x_{ki} + \epsilon_i \quad (4)$$

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \sum_{i=1}^{n} (y_i - \alpha_0 - \sum_{j=1}^{k} \alpha_j x_{ji})^2 \quad (5)$$

Here, y_i is the *i*-th observed value of the dependent variable y, $x_{1i}, x_{2i}, \ldots, x_{ki}$ are the *i*-th observed values of independent variables $x_1, x_2, \ldots, x_k$, ϵ_i is the residual term, n is the sample size and k is the number of independent variables.

2.5.Evaluation

In order to evaluate the accuracy of our predictions and to validate the model, we relied on a set of performance measures often used in the literature, namely the mean absolute error (MAE), the root of the mean squared error (RMSE), and the mean absolute percentage error (MAPE). The MAE provides an indication of the average absolute deviation between predicted and actual value, the RMSE penalizes larger errors more heavily, providing insight into the overall spread of the errors, while the MAPE measures the relative error in percentage terms and evaluates the model's performance in relation to the magnitude of the data. Therefore, using all three metrics provides a comprehensive assessment of the model's performance from different perspectives.

We also employed the k-fold cross-validation (CV), a commonly used technique for model evaluation (Krstajic et al. 2014). This technique involves partitioning the data set into *k* folds, where each fold contains an equal-sized sample from the original data set. Among the *k* folds, *k*-1 folds are utilized for model training, while the remaining fold is used for model validation. This process is repeated *k* times, and the performance measures, including MAE, RMSE, and MAPE are calculated by averaging the results obtained over all the folds. The formulas for each of these performance measures are as follows:

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

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

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

where $\hat{y}_i$ denotes the forecast value for the i-th observation y_i.

3. Results

In this section, we present the results of each step of the proposed methodology (Figure 1) , providing a comprehensive overview of the outcomes and insights gained throughout the evaluation process.

3.1.Case study

In our study, the prediction of fuel consumption has been performed on a set of nine sister bulk carriers having the same characteristics (see Table 2). The dataset provided by our industrial partner concerns 601 voyages of various durations (see Figure 5) sailed between 2019 and 2021 with a total of 4337 noon reports (see Figure 6).

Table 2: Main characteristics of bulk carriers

Size	Handymax
Engine	MAN-B&W
Deadweight (metric tons)	40,481
Gross tonnage (gt)	24,725
Length overall (m)	176.65

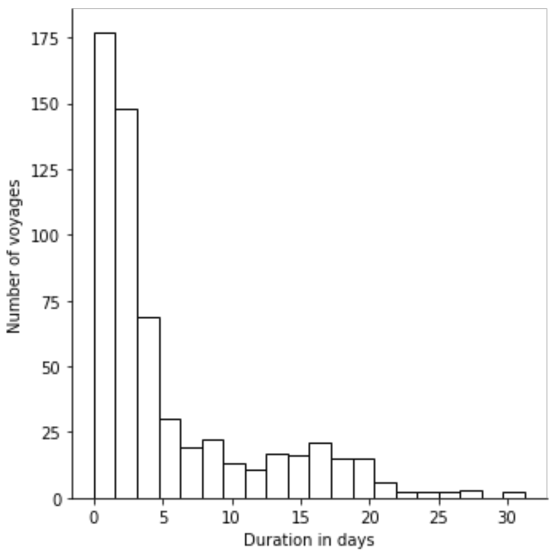

Figure 6: The distribution of the durations of the case study voyages (in days).

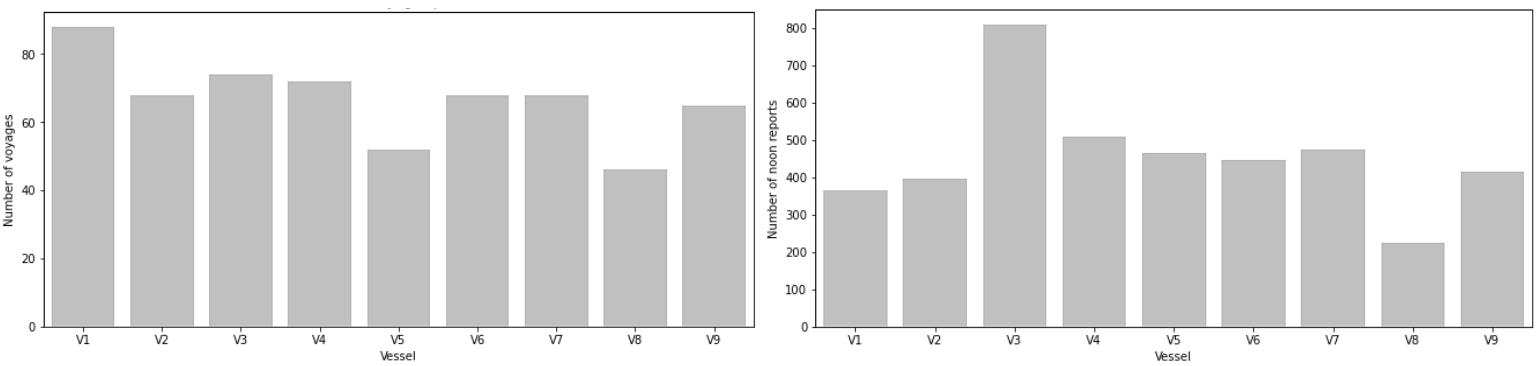

Figure 5:The number of voyages (Left) and the number of noon reports (Right) for each bulk carrier.

Furthermore, for our analysis, we define a set of scenarios corresponding to the fuel consumption prediction model with different features as inputs, as summarized in Table 3. The first scenario corresponds to the cubic relationship between fuel consumption and the RPM. The second scenario extends this cubic relationship by incorporating meteorological factors. This scenario takes into account the meteorological factors available at each noon report, hence at the time of the report. To reflect the weather conditions over the past 24 hours before the noon report, we define scenario 3, which considers the averaged weather factors over 24 hours while maintaining the cubic RPM speed. We also define two additional scenarios, 4 and 5. Scenario 4 is similar to scenario 2 but includes the propeller slip feature along with meteorological factors at noon reports. Scenario 5 is the same as scenario 3, including the slip and thus considering the averaged weather factors.

Table 1: Scenarios for predicting fuel consumption with various combinations of input features.

	RPM^3	Wind Factor	Averaged Wind Factor	Wave Factor	Averaged Wave Factor	Current Factor	Averaged current Factor	Slip
Scenario 1	X							
Scenario 2	X	X		X		X		
Scenario 3	X		X		X		X	
Scenario 4	X	X		X		X		X
Scenario 5	X		X		X		X	X

3.2. Data preprocessing

At the outlier removal stage, we employed a two-step process consisting of visual inspection and z-score method to identify and remove outliers, as explained in Section 2.3.1. The visual inspection involved thoroughly examining the dataset for any unusual or extreme values, while the z-score method was applied to calculate the z-scores for each data point and identify those with z-scores above a predefined threshold of 3 or below -3. In total, 5.68% of the data points were identified as outliers and removed from the dataset. The removal of these outliers resulted in changes in the distribution of the data, which can be observed in Table 4. This table presents descriptive statistics of the features that were identified as having the most impact on fuel consumption, both before and after the outlier removal process.

Table 2: Descriptive statistics of the most significant features before and after removing outliers

Feature	Before outliers removal				After outliers removal			
	Mean	**Standard Deviation**	**Minimum**	**Maximum**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
FuelConsumption (Mt/nautical mile)	0.068	1.073	-49.705	56.785	0.066	0.010	0.022	0.113
SpeedOverGround (knots)	11.479	2.372	1.020	127.522	11.577	1.291	6.410	15.979
RPM	81.093	8.670	7.200	470.000	81.226	5.881	55.520	96.000
Slip (%)	1.936	6.097	-105.60	96.000	7.288	8.197	-22.000	38.650
Wind Speed (Beaufort)	3.964	1.366	0.500	9.000	3.888	1.371	0.500	8.000
Wind Factor (knots)	-0.101	0.190	-4.000	0.000	-0.089	0.158	-1.395	0.000
Wave Height (m)	1.860	1.067	0.020	10.190	1.764	0.991	0.020	7.330
Wave Factor (knots)	-0.352	0.693	-8.500	3.000	-0.272	0.507	-5.000	3.000
Current Speed (m/s)	0.497	0.442	0.019	3.922	0.458	0.411	0.019	3.922
Current Factor (knots)	-0.138	0.537	-3.690	3.260	-0.128	0.478	-3.140	2.120

During the feature selection phase, our focus was on selecting the most relevant features that contribute to fuel consumption. Based on the coefficients presented in the correlation matrix (Figure 7), we decided to prioritize the RPM feature, which plays a significant role in determining the vessel's speed and its fuel consumption. Additionally, we chose to include the averaged weather factors, namely the Averaged Wind Factor, Averaged Wave Factor, and Averaged Current Factor that provide a comprehensive representation of the combined influence of the direction and the speed/height of wind, waves, and currents as explained previously. Although the correlation coefficient for the propeller slip is important, it must be recognized that this value is often not available and cannot be reliably estimated before the sailing of the vessel. Consequently, despite its importance, we were unable to incorporate it into our predictive model at this stage. By selecting the RPM and the averaged weather factors, we ensure that our model, representing the pre-defined scenario 3, accurately captures the impact of these key elements on the vessel's performance.

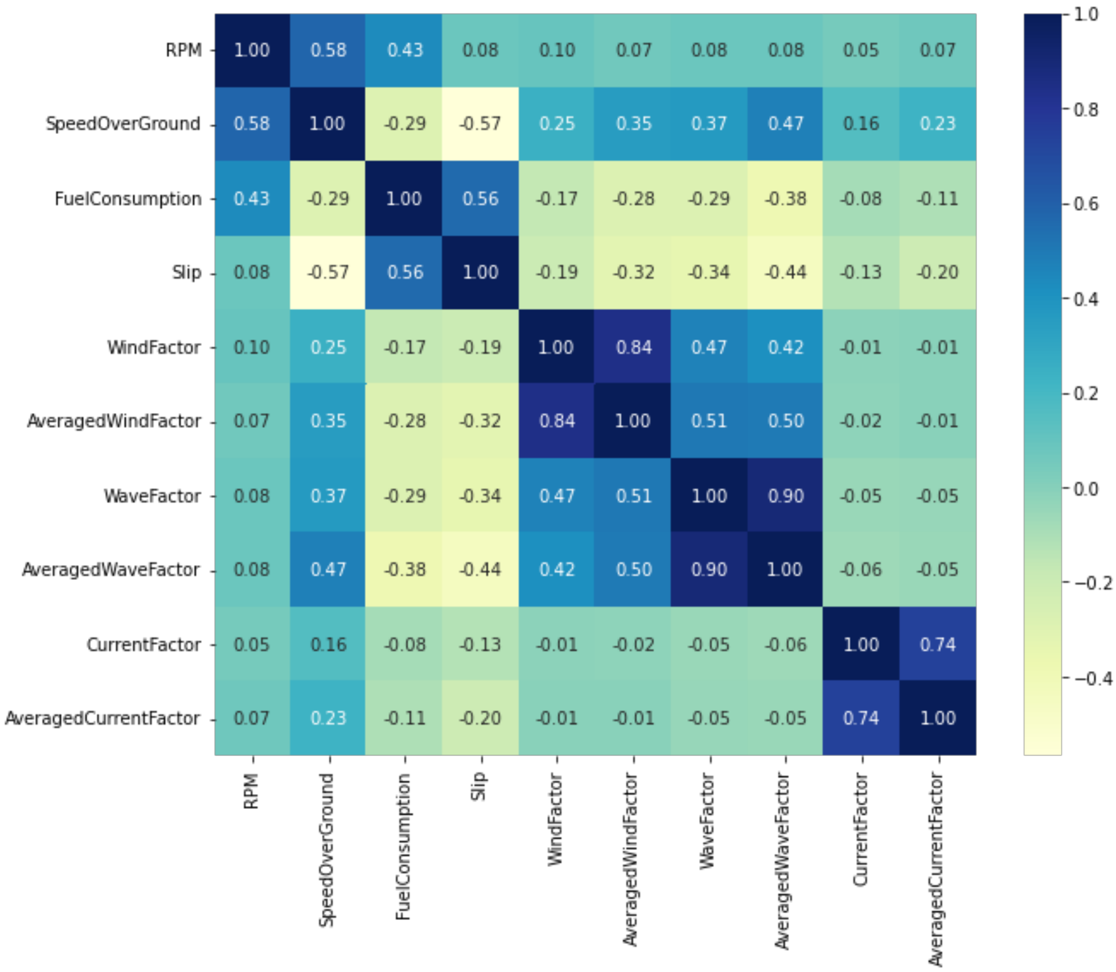

Figure 7: Correlation matrix

3.3. Modeling

Since we proceeded to the scaling of the data before the modeling step, it is necessary to do a reverse scaling of the weights obtained in order to have the final estimated regression function and to make an analysis of the results. A summary of the regression performed is presented in Table 5.

Table 3: Regression summary

Variable	**Coefficient**	**Standard error**	**t-statistic**	**P-value**
Intercept	3.89×10^{-2}	7.02×10^{-4}	55.53	0
RPM^3	4.28×10^{-8}	1.23×10^{-9}	34.85	0
Averaged Wind Factor	-9.79×10^{-3}	1.26×10^{-3}	-7.76	0
Averaged Wave Factor	-7.24×10^{-3}	3.25×10^{-4}	-22.32	0
Averaged Current Factor	-4.32×10^{-3}	3.36×10^{-4}	-11.92	0

3.4. Evaluation

The cross-validation was performed using 10 folds, and the proposed model, corresponding to the scenario 3, achieved a MAE of 5.49×10^{-3} metric tons per nautical mile, a RMSE of 7.36×10^{-3} metric tons per nautical mile, and a MAPE of 8.40 %. The errors were calculated on all noon reports, with the minimum value of fuel consumption of 22×10^{-3} metric tons per nautical mile and the maximum value of 113×10^{-3} metric tons per nautical mile. We also conducted the same evaluation for the remaining scenarios, and the summary of the results is provided in Table 6.

Table 4: A summary of the MAE, RMSE and MAPE errors across the entire dataset for the predefined scenarios.

	MAE	**RMSE**	**MAPE**
Scenario 1	6.66×10^{-3}	8.96×10^{-3}	10.26%
Scenario 2	6.10×10^{-3}	8.30×10^{-3}	9.42%
Scenario 3	5.49×10^{-3}	7.36×10^{-3}	8.40%
Scenario 4	4.59×10^{-3}	6.53×10^{-3}	7.10%
Scenario 5	4.37×10^{-3}	6.11×10^{-3}	6.67%

In addition to calculating the MAE. RMSE and MAPE on all the data, we also calculated residuals based on noon reports from the proposed model predictions using the following relationship:

$$\text{Residual}_i = y_i - \hat{y}_i \tag{12}$$

The distribution of the residuals (Figure 8) shows that 91.7% of the residuals are concentrated between -10^{-2} metric ton per nautical mile and 10^{-2} metric ton per nautical mile.

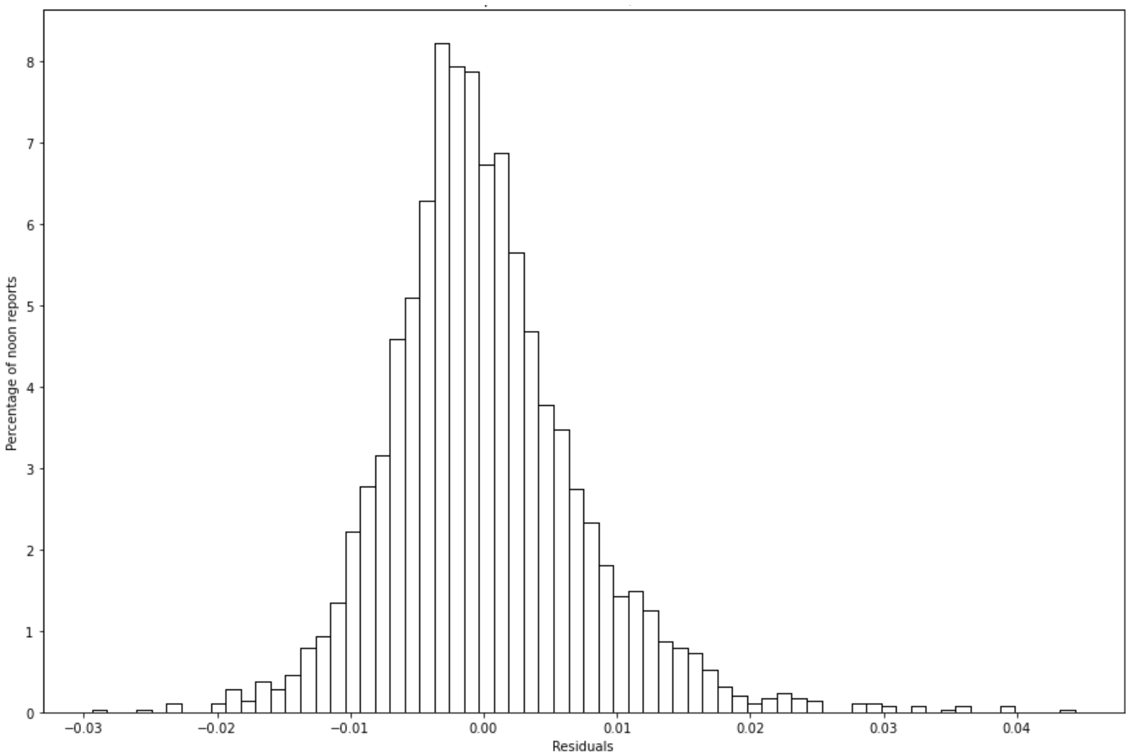

Figure 8: Distribution of residuals across noon reports for the proposed model.

Since our data concern several voyages of the sister bulk carriers, it would be relevant to analyze the distribution of different types of errors across all voyages rather than only focusing on the average errors presented in Table 6. According to Figure 9, for the proposed model, 85.71% of the voyages have a MAE of less than 10^{-2} metric ton per nautical mile, while 82.51% show a RMSE within the same 10^{-2} metric ton per nautical mile threshold. Furthermore, 64.18% of the voyages achieve a MAPE below 10%, indicating highly accurate forecasting according to Montaño et al. (2013). Also, 93.39% of the voyages maintain a MAPE lower than 20%, which aligns with the 'good forecasting' criteria outlined in Montaño et al. (2013).

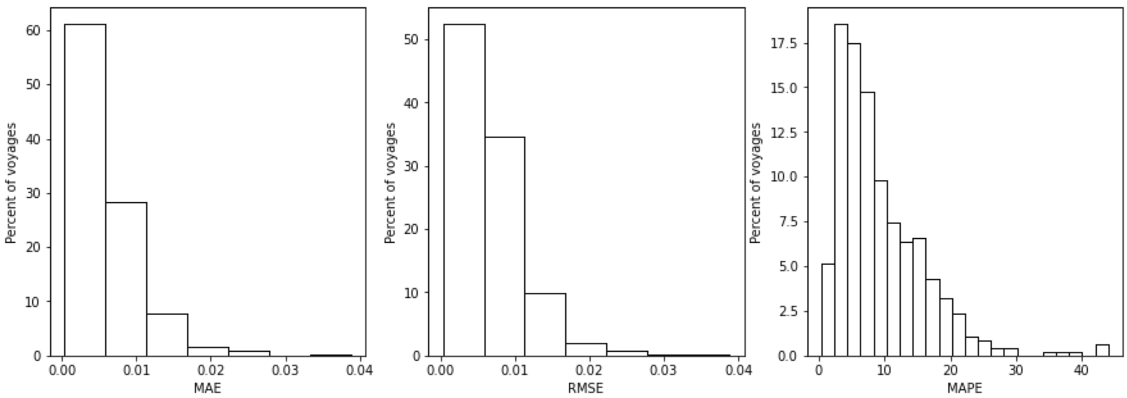

Figure 9: Distribution of errors across voyages for the proposed model.

Compared to a model based only on a cubic relationship with the RPM and lacking meteorological information (Scenario 1), our model demonstrates a lower error rate (Figure 10).

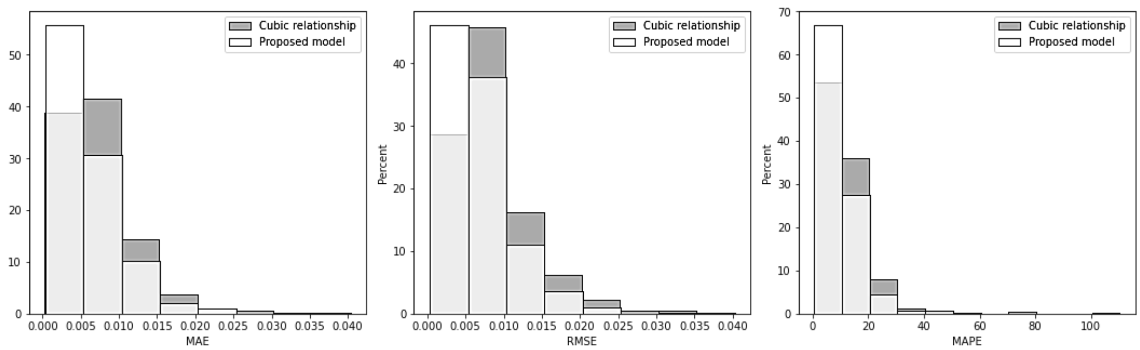

Figure 10: Comparison of the error distribution among voyages for the proposed model and the cubic relationship.

Specifically, 79.62% of the voyages have a MAE lower than 10^{-2} metric ton per nautical mile, 72.79% of the voyages have a RMSE lower than 10^{-2} metric ton per nautical mile, and 50.92% of the voyages have a MAPE lower than 10%. Furthermore, the distribution of residuals (Figure 11) from the noon reports indicates that the errors of the cubic relationship are spread over the same value ranges as our model, but with less concentration around zero.

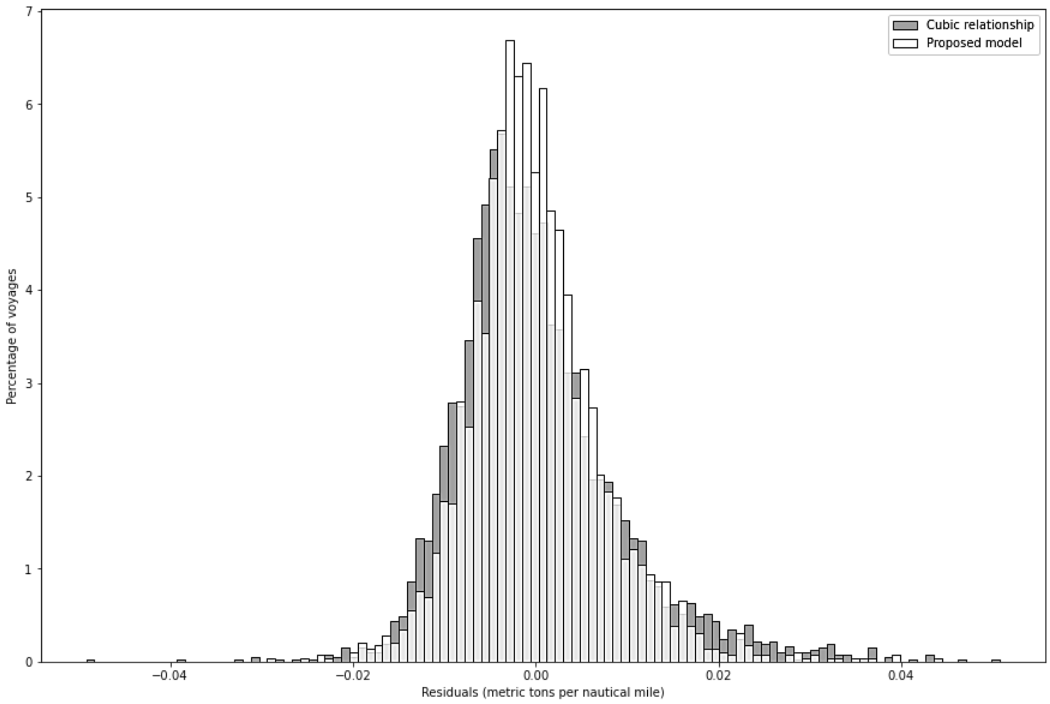

Figure 11: Comparison of the residuals' distribution among noon reports for the proposed model and the cubic relationship.

Moreover, the correlation matrix in Figure 7 highlights the significance of propeller slip in influencing fuel consumption. Therefore, despite being a post-voyage calculation in our dataset, its inclusion among the features offers valuable insights into fuel consumption. Consequently, a comparative analysis is conducted between our model (Scenario 3) without propeller slip and the same model augmented with propeller slip as an explanatory variable (Scenario 5).

After incorporating propeller slip into our model, the distribution of voyage-based errors shows that 91.68% of voyages have an MAE below 10^{-2} metric ton per nautical mile. 89.34% indicate an RMSE below this threshold, and 78.25% manifest an MAPE below 10%. The inclusion of propeller slip has resulted in a reduction in error rates, as illustrated in Figure 12.

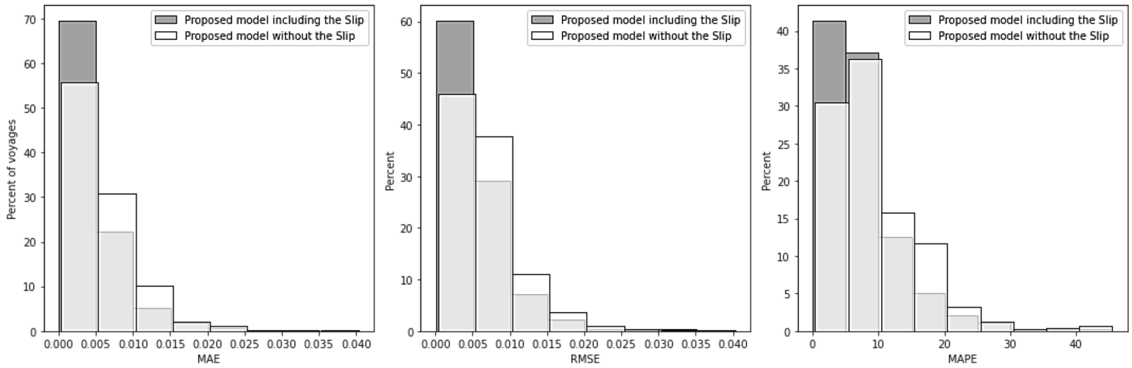

Figure 12: Comparison of the error distribution among voyages for the proposed model with and without the slip information.

The evaluation of the distribution of different types of errors (MAE. RMSE. and MAPE) across voyages was extended to the other scenarios. and a summary of the results is presented in Table 7.

Table 5: A summary of the percentage of voyages with MAE and RMSE lower than 10^{-2} Mt/nm and MAPE lower than 10% then lower than 15% for the predefined scenarios.

	MAE<10^{-2} Mt/nm	**RMSE <10^{-2} Mt/nm**	**MAPE<10%**	**MAPE<15%**
Scenario 1	79.63%	72.79%	50.92%	74.29%
Scenario 2	81.64%	76.79%	58.26%	78.30%
Scenario 3	85.71%	82.52%	64.18%	81.45%
Scenario 4	90.48%	87.15%	76.96%	89.48%
Scenario 5	91.68%	89.34%	78.25%	90.83%

4. Discussion

As previously mentioned in the results section, the proposed model successfully predicted the fuel consumption of nine ships with a MAE of 5.49×10^{-3} metric ton per nautical mile and a RMSE of 7.36×10^{-3} metric ton per nautical mile, respectively. However, since the calculated performance measures do not account for the direction of errors, these average error values (MAE and RMSE) across the entire dataset alone may not fully reflect the accuracy of our model. By averaging the errors over voyages, we reduced the impact of direction cancellation (overestimation and underestimation cancellation) and observed that approximately 85% of voyages have average errors lower than 10^{-2} metric ton per nautical mile (Figure 9), which is still considered good performance considering the average fuel consumption.

The distribution of residuals in noon reports (Figure 8) indicates that the majority of them are centred around zero, with roughly equal proportions of negative and positive values. This implies that our model's predictions are generally accurate despite the presence of certain outliers, which may be attributed to human factors affecting data quality. Since fuel consumption data in noon reports is manually recorded by captains, input errors may occur. Additionally, captains may intentionally report lower fuel consumption values due to contractual obligations or to avoid penalties, resulting in inaccurate or manipulated data. Furthermore, external factors such as measurement errors. equipment inaccuracies or variations in operating conditions could also contribute to the observed bias in the distribution of errors.

The comparison between our model and the scenario 1 model shows that our model has fewer errors as it takes into consideration the meteorological conditions while retaining the cubic relationship between fuel consumption and RPM speed. This suggests that while the cubic relationship may be reliable in calm water conditions, it does not reflect the realities of sailing, which involves navigating through meteorological conditions such as wind, waves, and currents.

Moreover, the results show that incorporating propeller slip into a predictive model (Scenario 5), where this information is available, reduces error rates and improves accuracy. This provides a deeper insight and a better understanding of the complex dynamics that influence fuel consumption as it accounts for the combined effects of engine resistance and weather conditions.

The various predefined scenarios are based on the availability of meteorological data between noon reports or their absence, and then on the availability of information on the propeller slip. The summary Tables 6 and 7 allow us to conclude the importance of integrating meteorological conditions, given that the highest error rate is observed in scenario 1 when compared to the other scenarios. Additionally, the availability of meteorological reports between noon reports and their consideration (through the averaging of meteorological factors over the past 24 hours) has shown lower error rates than other scenarios that rely on a single meteorological value available at each noon report (scenarios 2 and 4). The availability of the propeller slip further reduced the error rate, but since this information is not available before the ships' voyage in our case, the best scenario is the scenario 3.

5. Conclusion

Fuel consumption is a crucial factor in optimizing the performance of ships during sailing and reducing greenhouse gas emissions, as targeted by the IMO in the coming years. In this study, we have presented a systematic methodology to develop a fuel consumption prediction model based on a commonly observed cubic relationship in the literature, while incorporating meteorological factors that reflect the speed loss due to wind, waves, and currents. The Multiple Linear Regression model was utilized to establish this cubic relationship between fuel consumption and the propeller rotational speed along with the meteorological factors.

Our model was trained and tested on a dataset of nine sister bulk carriers with similar physical characteristics, resulting in accurate fuel consumption predictions with over 80% of voyages having average errors of less than 10^{-2} metric ton per nautical mile. Furthermore, a comparison with the cubic relationship highlighted the impact of considering meteorological conditions in predicting fuel consumption, resulting in improved accuracy and reduced errors. The incorporation of the slip factor improved the model's performance too. However, a crucial limitation arises from the impracticality of utilizing slip as a predictive feature in our case. This is due to the unavailability of slip data before the vessel's voyage. In addition, our model enhances its representation of meteorological conditions by overcoming the traditional methodologies. Instead of using only the speed/magnitude of the meteorological components and projecting their direction, we use an approach in which we merge speed/magnitude and direction into a single measure: the speed loss induced by each of the winds, waves and currents.

Further improvements to our model could be achieved by acquiring more high-quality data to minimize the biases observed in our results. This could involve collecting data from a wider range of vessel types, routes, and operating conditions to increase the accuracy of the model. Additionally, it is important to note that the model currently relies on averaged weather factors, which may limit its ability to capture the full range of weather variations. Moreover, incorporating real-time data from onboard sensors, such as fuel flow meters, weather monitoring equipment, and navigation data, could improve the accuracy of the fuel consumption prediction model.

In addition, the model could be extended to take into account other factors that affect fuel consumption such as vessel load, maintenance practices and operational procedures. By incorporating these variables into the model, it could provide more comprehensive and accurate fuel consumption predictions for different vessels in various operational scenarios.

The model presented in this study demonstrates significant potential as a valuable addition to weather routing systems when applied to various types of vessels. Integrating this model into existing systems could provide optimized routes that consider not only distance but also emissions in different weather scenarios. By incorporating emissions into the routing process, the routing system ensures that the proposed routes prioritize the reduction of environmental impact while maintaining efficiency. This integration aligns with the increasing emphasis on sustainability within the maritime industry and contributes to more environmentally conscious navigation practices.

6. 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.

7. References

Bayraktar M, Sokukcu M. 2023. Marine vessel energy efficiency performance prediction based on daily reported noon reports. Ships and Offshore Structures 1-10.

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 1:157–166.

Coraddu A, Oneto L, Baldi F, Anguita D. 2017. Vessels fuel consumption forecast and trim optimisation: a data analytics perspective. Ocean Engineering 130:351–370.

Fan A, Yan X, Bucknall R, Yin Q, Ji S, Liu Y, Song R, Chen X. 2020. A novel ship energy efficiency model considering random environmental parameters. Journal of Marine Engineering & Technology 19:215–228.

Fan A, Yang J, Yang L, Wu D, Vladimir N. 2022. A review of ship fuel consumption models. Ocean Engineering 264:112405.

Farag YBA, Ölçer AI. 2020. The development of a ship performance model in varying operating conditions based on ANN and regression techniques. Ocean Engineering, 198, 106972.

Gkerekos C, Lazakis I, Theotokatos G. 2019. Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study. Ocean Engineering, 188, 106282.

Hu Z, Jin Y, Hu Q, Sen S, Zhou T, Osman MT. 2019. Prediction of fuel consumption for enroute ship based on machine learning. IEEE Access 7:119497–119505.

Işıklı E, Aydın N, Bilgili L, Toprak A. 2020. Estimating fuel consumption in maritime transport. Journal of Cleaner Production 275:124142.

Joung TH, Kang SG, Lee JK, Ahn J. 2020. The IMO initial strategy for reducing Greenhouse Gas (GHG) emissions, and its follow-up actions towards 2050. Journal of International Maritime Safety, Environmental Affairs, and Shipping 4(1):1-7.

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.

Kim YR, Jung M, Park JB. 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.

Krstajic D, Buturovic LJ, Leahy DE, Thomas S. 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. Journal of Cheminformatics 6(1):1-15.

Moreno JJM, Pol AP, Abad AS, Blasco BC. 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema 25:500-506. DOI: 10.7334/psicothema2013.23.

Psaraftis HN, 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 100111.

Simonsen M, Walnum HJ, Gössling S. 2018. Model for estimation of fuel consumption of cruise ships. Energies, 11 (5), 1059.

Solutions ME. 2018. Basic principles of ship propulsion. MAN Energy Solutions: Copenhagen, Denmark.

Statistics E. 2020. The 2020 world merchant fleet statistics from equasis. Statistics from Equasis.

Stopford M. 2008. Maritime economics 3e. Routledge.

Tarelko W, Rudzki K. 2020. Applying artificial neural networks for modelling ship speed and fuel consumption. Neural Computing and Applications 32(23):17379–17395.

Tran TA. 2021. Comparative analysis on the fuel consumption prediction model for bulk carriers from ship launching to current states based on sea trial data and machine learning technique. Journal of Ocean Engineering and Science 6:317–339.

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.

Wang S, Meng Q. 2012. Liner ship route schedule design with sea contingency time and port time uncertainty. Transportation Research Part B: Methodological 46(5):615-633.

Weiqiang Y, Honggui L. 2013. Application of grey theory to the prediction of diesel consumption of diesel generator set. Proceedings of the 2013 IEEE International Conference on Grey systems and Intelligent Services (GSIS). IEEE.

Yuan J, Nian, V. 2018. Ship energy consumption prediction with Gaussian process metamodel. Journal of Ocean Engineering and Science 152:655-660.