Journal of *Marine Science and Engineering*

Article

Navigating Energy Efficiency: A Multifaceted Interpretability of Fuel Oil Consumption Prediction in Cargo Container Vessel Considering the Operational and Environmental Factors

Melia Putri Handayani [1], Hyunju Kim [2], Sangbong Lee [3] and Jihwan Lee [1,*]

[1] Department of Industrial and Data Engineering, Major in Industrial Data Science and Engineering, Pukyong National University, Busan 48513, Republic of Korea; meliaph@pukyong.ac.kr
[2] Korea Marine Equipment Research Institute, Busan 49111, Republic of Korea; hjkim94@komeri.re.kr
[3] Lab021 Shipping Analytics, Busan 48508, Republic of Korea; sblee@lab021.co.kr
* Correspondence: jihwan@pknu.ac.kr; Tel.: +82-51-629-6492

Abstract: In the maritime industry, optimizing vessel fuel oil consumption is crucial for improving energy efficiency and reducing shipping emissions. However, effectively utilizing operational data to advance performance monitoring and optimization remains a challenge. An XGBoost Regressor model was developed using a comprehensive dataset, delivering strong predictive performance (R^2 = 0.95, MAE = 10.78 kg/h). This predictive model considers operational (controllable) and environmental (uncontrollable) variables, offering insights into complex FOC factors. To enhance interpretability, SHAP analysis is employed, revealing 'Average Draught (Aft and Fore)' as the key controllable factor and emphasizing 'Relative Wind Speed' as the dominant uncontrollable factor impacting vessel FOC. This research extends to further analysis of the extremely high FOC point, identifying patterns in the Strait of Malacca and the South China Sea. These findings provide region-specific insights, guiding energy efficiency improvement, operational strategy refinement, and sea resistance mitigation. In summary, our study introduces a groundbreaking framework leveraging machine learning and SHAP analysis to advance FOC understanding and enhance maritime decision making, contributing significantly to energy efficiency and operational strategies—a substantial contribution to a responsible shipping performance assessment under tightening regulations.

Keywords: maritime; ship energy efficiency; fuel oil consumption prediction; ship performance assessment; data analytics; machine learning; Explainable Artificial Intelligence (XAI)

check for updates

Citation: Handayani, M.P.; Kim, H.; Lee, S.; Lee, J. Navigating Energy Efficiency: A Multifaceted Interpretability of Fuel Oil Consumption Prediction in Cargo Container Vessel Considering the Operational and Environmental Factors. *J. Mar. Sci. Eng.* **2023**, *11*, 2165. https://doi.org/10.3390/jmse11112165

Academic Editor: Nikos Themelis

Received: 24 September 2023
Revised: 10 November 2023
Accepted: 11 November 2023
Published: 13 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Background of Study

A recent estimation reported that the ocean economy is worth between \$3 trillion and \$6 trillion annually, highlighting its significance to the world economy [1]. This is believed to be due to the fact that maritime transport has long been recognized as the most energy-efficient mode of transportation in terms of energy used per ton-kilometer transported [2]. Yet, the world seaborne trade that accounts for 80–90% of global volumes is reported to contribute 9% to greenhouse gas (GHG) emissions in the transport sector, mainly by international shipping [3]. It is suspected mainly due to fuel consumption emissions. The transport sector is reported to be responsible for 27% of global final-energy demand, and almost 3% of the world's final-energy demand, including 8% of the world's oil, is consumed by ships, mainly international cargo shipping [4]. Thus, fossil fuel dependency issues are plaguing the shipping industry [5]. These facts translated to the urgency of promoting energy efficiency in shipping, minimizing fuel oil consumption, and thus reducing GHG emissions.

As the response to the urgencies stated by the regulatory bodies, accurately predicting a vessel's fuel oil consumption is crucial for monitoring efficiency metrics mandated by regulators. The attempt can help ship owners and operators more precisely calculate metrics like EEOI, EEXI, and CII score; research that improves capabilities for forecasting a vessel's daily or voyage fuel use based on design factors, operational parameters, and weather conditions is necessary [6].

This research focuses on the pivotal role of accurate fuel oil consumption prediction in monitoring efficiency metrics mandated by regulatory bodies. By leveraging advanced data collection systems and employing cutting-edge statistical tools and machine learning techniques, we aim to develop a robust framework for forecasting a vessel's daily or voyage fuel use. This framework accounts for an extensive range of operational and environmental factors, providing essential insights to support the industry's transition to a sustainable, low-carbon future.

Advanced data collecting system is now enabling the enormous provision of data ranging from the operational factor to the environmental factors, supported by the more comprehensive modeling that can be developed due to the advancement of statistical tools and machine learning. However, limitations persist in fully leveraging information and transparently explaining predictions.

This is where the synergy between machine learning-based regression models and Explainable Artificial Intelligence (XAI) becomes paramount. These intricate interactions between controllable and uncontrollable variables demand sophisticated predictive models to navigate. Machine learning models, adept at unraveling such complexities, offer invaluable insights.

By complementing machine learning models with XAI, particularly the SHAP framework, we unveil the inner workings of these models. SHAP opens the door to an unprecedented level of transparency, enabling stakeholders to understand the specific contributions of each variable in the prediction process. This illumination translates to actionable insights for maritime professionals, empowering them to make informed decisions. In essence, the integration of machine learning and XAI shines a much-needed light on the intricate relationships within maritime fuel consumption, guiding the way toward improved efficiency, reduced environmental impact, and sustainable shipping practices.

1.2. Aim of the Study

Accurately predicting vessel fuel oil consumption is important for enhancing energy efficiency in shipping. A complex interplay of operational and environmental factors influences consumption levels. Developing machine learning models to forecast usage based on these predictors can provide insights to optimize routes and operations.

This study aims to advance such predictive modeling through three key objectives. First, a regression-based model will be developed using extensive operational and environmental data. Next, SHAP explanations will interpret the model to understand how the controllable (operational) and uncontrollable (environmental) predictors influence predictions globally. Finally, high consumption outliers will be identified for focused analysis of region-specific consumption dynamics.

Achieving highly accurate and transparent forecasts can support optimized planning and mitigation of fuel-intensive scenarios. Refining both predictive performance and interpretability through SHAP explanations aims to accelerate emissions reductions required under tightening regulations.

Understanding influential factors and contextualizing predictions regionally provides decision support for technical and operational changes to improve energy efficiency over time. This enhances compliance monitoring and continuous performance improvements mandated internationally. Ultimately, such analytics may support emerging policies through empirical evaluation of consumption determinants.

The overarching goal of this research is to develop a machine learning approach for predicting cargo vessel fuel oil consumption that provides interpretable insights to support energy efficiency efforts. Specifically, the procedures are:

1. Build an accurate predictive model using a regression-based model to forecast vessel fuel oil consumption (FOC) based on operational and environmental variables. The model performance will be evaluated using metrics like R-squared, RMSE, and MAE.
2. Interpret the key drivers of the model's predictions using SHAP values to understand how factors influence expected FOC under different conditions. This can reveal the most important controllable and uncontrollable factors overall.
3. Analyze regions or vessel routes where the model forecasts extremely high FOC and explore if differences exist in the factors' relative importance rankings—identifying region-specific consumption drivers to guide context-specific optimizations.
4. Provide recommendations for optimizing vessel operations, routing, and design based on the model interpretation. This includes suggestions for controlling influential operational parameters and mitigating uncontrollable factors where possible.
5. Present findings in a way that helps stakeholders in the maritime industry enhance decision-making regarding energy efficiency improvements through better comprehension of complex FOC determinants.

In essence, this research endeavors to harness the power of machine learning and interpretability techniques to provide an accurate and insightful understanding of cargo vessel fuel oil consumption. By constructing a predictive model with a regression model and interpreting its outcomes using SHAP values, we aim to shed light on the key drivers behind fuel consumption dynamics, encompassing both controllable and uncontrollable factors. Furthermore, our analysis extends to specific regions and vessel routes, offering a context-specific perspective on consumption determinants.

The following literature review will situate this study within the body of existing research on predictive modeling and interpretability techniques applied for fuel consumption analysis in maritime operations.

2. Literature Review

Navigating the complex seas of global trade necessitates an unwavering commitment to fuel efficiency within cargo vessels. This chapter embarks on an extensive expedition through existing scholarly works to unearth invaluable insights and identify knowledge gaps concerning the prediction of fuel oil consumption (FOC). The expedition explores critical variables, methodological paradigms, and the underlying motivations, setting the stage for a comprehensive investigation into FOC dynamics.

2.1. Existing Research

Maritime operations, including shipping, port operations, and offshore activities, are known to be energy-intensive. Thus, prioritizing energy efficiency is paramount, driven by the compelling need to curtail greenhouse gas (GHG) emissions and adhere to stringent environmental regulations. Energy-efficient practices not only reduce fuel consumption and carbon emissions but also result in cost savings for maritime companies [7].

As technology continues to advance and more data becomes available, machine learning, propelled by the advent of extensive datasets and advanced algorithms, has emerged as a transformative force, enabling the extraction of invaluable insights, process automation, and informed decision-making within the maritime sector to support energy efficiency efforts [8]. Maritime experts and scientists explore various approaches to improve transport efficiency on seas around the globe [9], which includes the implementation of machine learning. Existing research has underscored the instrumental role of machine learning in elevating energy efficiency within maritime operations, heralding a new era of innovation.

This technological prowess has given rise to a multitude of applications, including route planning [10], weather routing [11], speed optimization [12], trim optimization [13], ship routing and vessel scheduling [14], predictive maintenance [15], anomaly detection [16], shaft power prediction [17], safety management [18], etc. These diverse machine learning applications collectively hold the promise of bolstering operational efficiency, reducing fuel consumption, and curbing greenhouse gas emissions—a transformative shift within the shipping industry.

At the heart of this technological transformation lies a critical cornerstone: the precision of fuel oil consumption (FOC) prediction. Fuel oil consumption in the maritime industry has undergone significant changes over the years, driven by technological advancements, environmental regulations, and economic considerations. Accurate prediction of fuel consumption is crucial for efficient sailing, ship route network construction, and vessel management [19].

Researchers and industry experts have conducted comprehensive studies to analyze and understand fuel oil consumption in the maritime domain. They have developed models to predict fuel consumption based on various parameters such as vessel characteristics, weather conditions, and operational profiles.

Various machine learning algorithms that have been utilized in the attempt to develop the FOC prediction model include linear regression [20], multiple linear regression [21], ridge regression [22], support vector regressor [23], lasso regression [24], K-nearest neighbor regressor [25], extra tree regressor [26], random forest regressor [27], Gaussian process metamodel [28], artificial neural network (ANN) approach [29], and even deep learning [30]. Traditional methods, such as statistical analysis, have initially been used to examine historical consumption patterns, identify key factors influencing consumption, and develop predictive models, but these methods have been found to have low accuracy [31].

Machine learning models potentially demonstrate the potential to provide accurate real-time predictions for fuel management and optimization based on these datasets. This growing prominence of machine learning and deep learning techniques reflects the industry's increasing demand for transparency and interpretability.

Explainable Artificial Intelligence (XAI) has gained attention in various industries, including maritime, as a response to the need for transparent and interpretable AI systems [32]. XAI focuses on developing methods and models that are easily understood by humans [33]. In the maritime industry's evolving landscape, XAI plays a vital role in establishing credibility and accountability for AI systems [34]. One primary challenge in the maritime sector is the opacity of AI models, potentially hindering their acceptance [35].

XAI techniques, such as visualization and explanation methods, address this challenge by shedding light on AI systems' decision-making processes [36]. As machine learning techniques gain traction in maritime operations, transparency and accountability become increasingly important. Decision makers must understand the reasoning behind machine learning tasks, and regulatory bodies and stakeholders require explanations to ensure compliance with industry standards.

Only limited studies in the maritime domain have ever leveraged the machine learning implementation with XAI for some cases. Kim et al. (2023) developed a machine learning prediction model to estimate the vessel shaft power, then leveraged the prediction model with XAI using SHAP to obtain its feature attribution [37]. In other research, Kim et al. (2021) conducted vessel main engine anomaly detection and explained the prediction of whether an instance is considered an anomaly or not using SHAP [38].

Explainable Artificial Intelligence (XAI) holds immense potential for optimizing maritime operations, from individual vessels to industry-wide practices. Yet, with the considerably enormous amount of research utilizing machine learning, especially the black box model, the research that leverages the interpretability of machine learning using XAI is still very low. With the increasing use of AI in the maritime sector, the adoption of XAI techniques is necessary to help ensure the safe and reliable operation of maritime assets.

2.2. Research Gaps and Contributions

The literature review highlights some key gaps in existing research on fuel oil consumption modeling and interpretability techniques applied to shipping operations. While XAI methods are gaining attention for maritime applications, interpretability studies specific to fuel consumption predictions remain limited. Techniques like SHAP have yet to be leveraged for exploring location-specific consumption dynamics and evaluating changes in input importance under high consumption conditions.

To address these gaps, this study develops and interprets a machine learning model using a large and varied ship operational dataset. Table 1 outlines the machine learning models employed by the existing research in FOC prediction, the predictors used for fuel oil consumption prediction, and the incorporation of eXplainable Artificial Intelligence (XAI) techniques. Notably, this review underscores the distinctive nature of our research.

Table 1. Existing research on FOC prediction with machine learning.

Citation	Model	Number of Predictors				XAI
		Operational	Environmental	Engine	Vessel Characteristics	
[20]	Linear Regression	3	0	0	2	✗
[21]	Multiple Linear Regression	1	3	1	0	✗
[22]	Ridge Regression	5	4	0	0	✗
[23]	Support Vector Regressor	3	0	16	0	✗
[24]	Lasso Regression	8	7	0	3	✗
[25]	K-Nearest Neighbor Regressor	11	0	0	4	✗
[26]	Extra Tree Regressor	2	4	0	0	✗
[27]	Random Forest Regressor	2	9	0	0	✗
[28]	Gaussian Process Metamodel	3	4	0	0	✗
[29]	Artificial Neural Network (ANN)	1	3	1	0	✗
[30]	Deep Learning	2	2	0	1	✗
This Research	XGBoost Regressor	5	13	0	0	✓

In contrast to prior work, complex consumption determinants are considered by accounting for both controllable maneuverability factors and uncontrollable oceanographic variables. Moreover, our comparative analysis with existing models, as seen in Table 1, further supports the judicious selection of the XGBoost Regressor. This model exhibited superior predictive performance, with an impressive R-squared value of 0.95, showcasing its effectiveness in capturing the complexities of fuel consumption dynamics.

When compared to the respective metrics reported in the literature, such as the R-squared values of 0.84 for linear regression [20], 0.90 for ridge regression [22], 0.99 for support vector regressor [23], 0.74 for lasso regression [24], and 0.66 for K-nearest neighbor regressor [25], our chosen XGBoost Regressor consistently outperformed these models. The extra tree regressor [26] showed a greater R-squared score, which is 0.97; however, it is to be noted that they used considerably low numbers of predictors and only trained the model for a data sample with less than 4000 instances.

Additionally, models like multiple linear regression [22], random forest regressor [27], and Gaussian process model [28] were reported with RMSE values (1.61, 1.78, and 0.44, respectively). It is essential to highlight that Citation [30] presented the performance of deep learning with a MAPE of 0.58. This comparative overview aims to underscore the varied performance metrics across different models, providing a nuanced perspective on the quality of our proposed model while considering the inherent disparities in datasets and modeling approaches.

The predictive performance achieved, as mentioned in the abstract, demonstrates this model's potential. This research introduced a novel framework that not only leverages advanced machine learning, specifically XGBoost Regressor, but also emphasizes the interpretability aspect through SHAP (SHapley Additive exPlanations) analysis. The combination of a powerful predictive model and XAI techniques sets our work apart, as none of the reviewed studies adopted a similar approach.

Significantly, this research is the first known application of SHAP values to increase the comprehensibility of fuel consumption forecasts for the maritime industry. By identifying key driving factors and their behavior under different scenarios, this framework provides novel and actionable insights beyond what previous interpretability approaches have achieved. Overall, this study aims to advance both modeling accuracy and understanding of fuel use dynamics through an explainable machine learning methodology. Findings are expected to guide energy efficiency improvements in complex real-world shipping operations.

3. Data and Methodologies

This chapter provides the data overview and methodologies utilized in this research. The analysis commences with an in-depth examination of a real-world voyage dataset, meticulously scrutinizing the intricacies and nuances of the observed vessel dataset and the pre-processing procedure to better prepare the data.

3.1. Data Overview

3.1.1. Data Acquisition

This research acquired the data from a vessel noon report and sensor data collected from a general cargo ship with detailed vessel specifications, as seen in Table 2. This data was retrieved from the voyage noon report and onboard sensors system that was collected in a span of 15 months long voyage from 29 February 2020, until 8 June 2021.

Table 2. Vessel profile.

Register		Capacity		Size	
Type	General Cargo	Gross Tonnage	41,416	Length (m)	201
Year Built	2020	Summer DWT (t)	62,321	Breadth (m)	34

As one voyage number is registered as one route between at least two ports, this general cargo ship operated administratively with six voyage numbers and sailed across nine ports located in four different continents such as Asia, Africa, South America, and North America. The vessel voyage trajectory can be seen in Figure 1.

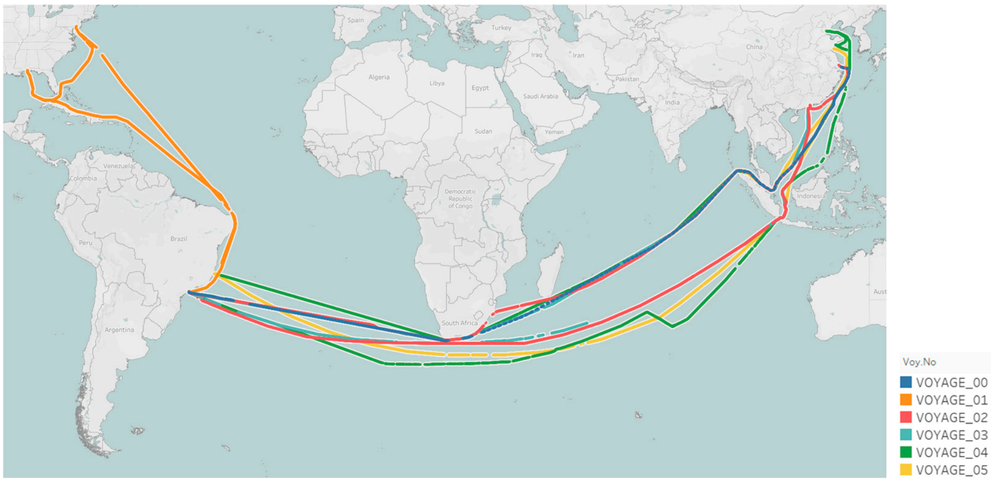


Figure 1. Vessel voyage trajectory.

3.1.2. Feature Selection

A key goal of this study was to develop a comprehensive fuel oil consumption prediction model incorporating both operational and environmental determinants. As such, features were carefully selected from the raw dataset to represent both controllable maneuverability factors and uncontrollable oceanographic conditions. On the operational side, features such as vessel speed, draught, cargo load percentage, and engine load provided insights into how management decisions could impact fuel usage. Environmental features like wind speed, wave height, and sea surface temperature were also included to account for external sea state influences out of the vessel's control. The list can be seen in Table 3.

Table 3. Original feature lists.

Feature Name	Unit	Description
Main Engine 1 FOC	kg/h	Fuel oil consumption of main engine 1
Speed Over Ground	knots	Vessel speed relative to the ground
Draught Fore	m	Distance between the waterline and the bottom of the vessel at the bow (front) end
Draught Aft	m	Distance between the waterline and the bottom of the vessel at the stern (back) end
Current Speed	m/s	Directional movement of seawater driven by gravity, wind, and water density
Wind Speed	m/s	Speed of the geographic or ground wind, assuming no tidal flow
Relative Wind Speed	m/s	Wind speed adjusted for the speed at which the vessel is traveling
Sea Surface Salinity	PSU	Salt concentration of the ocean water at the surface
Sea Surface Temperature	°C	Measure of heat in degrees at the top layer of the surrounding sea
Total Wave Height	m	Vertical distance between trough and crest including all wave components
Swell Wave Height	m	Height of the long-period waves not affected by local winds
Wind Wave Height	m	Height of gravity waves on the sea surface directly generated, sustained by wind
Wind Wave Period	s	Time interval between appearances of the same phase of a wind wave
Ship Heading	°	Direction the vessel's pointed end is facing, in degrees from north
Course Over Ground	°	Direction of progress over the ground actually covered, regardless of heading
Rudder Angle	°	Measured position of the vessel's side-to-side steering mechanism
Current Direction	°	Compass orientation of the flowing ocean water movement
Total Wave Direction	°	Direction from which the combined wind and swell waves are coming from
Swell Wave Direction	°	Direction the long regular ocean swells are originating from.
Wind Wave Direction	°	Orientation the wind-generated waves are coming from.
Wind Direction	°	Direction of the geographic or ground wind
Relative Wind Direction	°	Measured angle of the wind in relation to the heading of the moving ship.

Considering both types of factors simultaneously is expected to yield a more accurate and representative model of real-world fuel oil consumption dynamics. Past predictive studies have often focused on limited controllable or voyage-based attributes. However, complex consumption patterns are driven by interactions between operational profiles and varying oceanographic conditions that change over routes and locations. By accommodating both internal decision variables and external sea condition attributes, this model aims to capture these complex relationships to provide enhanced forecasting ability beyond existing approaches. Such an interpretable model can offer insights into both optimizing management strategies as well as mitigating external environmental impacts. The size of the original data set from the cargo vessel is 85,768 rows with a total of 257 columns, but only the 22 features in Table 3 were arbitrarily chosen for this analysis.

3.2. Data Preprocessing

3.2.1. Data Filtering

To ensure only realistic and meaningful data was used in modeling, certain filtering steps were implemented during preprocessing. Drawing from maritime domain expertise as well as general data cleaning practices, values considered erroneous or in error in features were identified and filtered out. This included removing values that fell far outside expected physical ranges based on common vessel specifications and operating conditions. For example, timestamps with incomplete or non-chronological time sequencing were discarded. Numerical features involving measurements like draught, speed, wave height, etc., that contained values severely deviating from standard measurement techniques and thereby deemed erroneous were also omitted. Additionally, any data records with high levels of missingness across fields were dropped. Documenting these filtering criteria allows reproducibility while focusing the modeling on valid operational scenarios for enhanced integrity of results.

3.2.2. Feature Transformation

Several new features were engineered by transforming existing variables using common nautical calculations. These transformations were introduced to better capture vessel motion and sea state impacts relative to the ship's heading. Important vessel geometry and stability metrics like average draught, trim, and heel were derived from fore, aft, and port/starboard draught readings. Relative angle features were also created by taking the differences between headings and environmental directionals to quantify interactions between the ship and waves/currents/wind. Specifically, leeway, tideway, wind wave angle, swell wave angle, and relative wind angle features provided additional insight beyond raw directional values. The transformed features produced through these calculations are summarized in Table 4. The final dataset has 58,548 rows and 18 columns.

Table 4. Feature transformation.

Feature Name	Unit	Description
Average Draught Aft and Fore	M	Average of draught fore and draught aft
Vessel Trim	M	Difference between draught fore and draught aft
Vessel Leeway	°	Difference between ship heading and course over ground
Vessel Tideway	°	Difference between ship heading and current direction
Wind Wave Angle	°	Difference between ship heading and wind wave direction
Swell Wave Angle	°	Difference between ship heading and swell wave direction
Relative Wind Angle	°	Difference between ship heading and relative wind direction

3.3. Methodologies

The overall framework involved data preprocessing, predictive modeling using regression analysis, identification of extremely high consumption points, and use of interpretable machine learning techniques to explain model predictions.

Specifically, contributions were aimed at not just improving predictive performance through the inclusion of diverse operational and oceanographic factors but also enhancing comprehension of consumption patterns through a focus on explainability. This combined methodology sets out to advance both accuracy and transparency in fuel consumption modeling for the optimization of shipping operations. The methodology that is utilized in this research will be explained in the form of a research procedure, as seen in Figure 2.

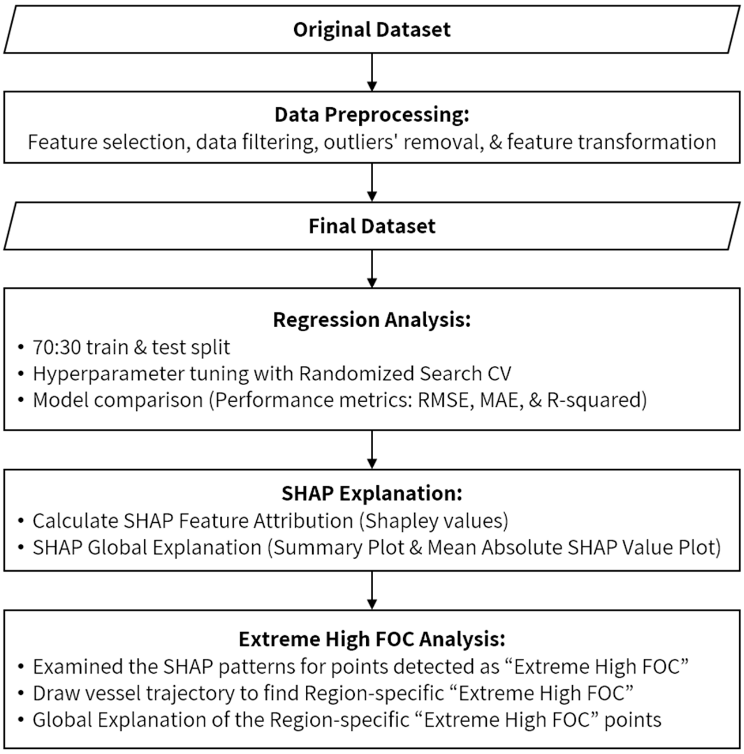


Figure 2. Research procedure.

- Data Preprocessing: The dataset underwent data cleaning including filtering outliers and implausible values. Key features were also engineered from raw measurements, as explained in Section 3.2. This prepared the data for modeling.
- Regression Analysis: The preprocessed data was split 70% for model training and 30% for testing. This commonly used split ratio provides sufficient data for fitting while reserving an independent portion for unbiased evaluation. Three regressor models were tuned and trained. Its performance was evaluated on the test set using metrics like R-squared, RMSE, MAE, and MAPE to quantify predictive accuracy. The best regressor was developed to predict FOC using the training set.
- IQR Analysis: The interquartile range was determined to identify outlier FOC values above the third quartile, indicating potentially extreme consumption. Focusing interpretation on these points aimed to uncover dynamics in atypical usage scenarios.
- SHAP Explanation: Model explanations are crucial to support complex decision-making. SHAP values were computed to highlight the relative impact of each input on FOC predictions, increasing the comprehensibility of consumption determinants.
- Extreme High FOC Analysis: Further analysis specifically examined the SHAP patterns for points detected as extremely high FOC. Region-specific analyses then explored whether certain areas exhibited divergent input importance trends compared to overall patterns.

Following the research procedure above, we conducted the multifaceted analysis that aims to predict the vessel FOC and explain the prediction result by investigating specifically the index of extremely high FOC, which resulted in a region-specific SHAP explanation. Each of the faceted procedures will be explained in the following subchapters.

3.3.1. Fuel Oil Consumption Prediction

To build a robust prediction model, it was important to consider a diverse set of influential factors. Previous studies often examined limited subsets of either operational or environmental attributes. However, fuel consumption is driven by complex interplays between controllable maneuvering profiles and varying uncontrollable ocean conditions. Therefore, this study aimed to develop a model using both operational (e.g., speed, draft) and environmental (e.g., waves, wind) predictors to more comprehensively represent real-world determinants.

The regressor utilized in this research was selected as the regression algorithm after a comprehensive comparative study involving three prominent algorithms: XGBoost, CatBoost Regressor, and Gradient Boosting Regressor. We undertook an extensive evaluation process, which included hyperparameter tuning for all three models to optimize their performance. In Section 4.1, XGBoost has been proven to have ultimately emerged as the superior choice. XGBoost, renowned for its robustness and efficiency in handling large, diverse datasets comprising both numerical and categorical features, aligns perfectly with the demands of our research. It employs an optimized distributed gradient boosting method, iteratively adding weak estimators [39]. This approach has consistently delivered highly accurate predictive models, surpassing other machine learning techniques such as random forest and neural networks in various domains. Key advantages of XGBoost include built-in measures against overfitting through regularization and the ability to efficiently train on massive datasets across computing clusters.

Randomized search cross-validation (CV) was employed for the tuning process to optimize the XGBoost regressor for this dataset and problem. These included parameters affecting the depth of trees, learning rate, number of estimators, and regularization factors to strike the right balance between underfitting and overfitting during training. Randomized search CV randomly samples hyperparameter configurations to evaluate the CV splits, allowing intensive searching of the hyperparameter space without defining a finite set of values to test exhaustively. The tuned model was selected based on CV scores to maximize predictive accuracy on unseen test data while controlling model complexity. Tuning helped ensure the final XGBoost regressor achieved high performance for fuel consumption prediction.

3.3.2. Extreme High Fuel Oil Consumption Detection

While predicting average fuel consumption is useful, understanding episodic extreme usage provides critical insights. Periods of extremely high FOC potentially signify operational inefficiencies or impacts of unusual conditions. Pinpointing such outliers allows targeted analysis that regular predictions obscure. IQR analysis was chosen to objectively identify extreme values as it is resistant to outlier effects compared to methods like standard deviation.

IQR works by defining an inner fence from the first (Q1) to the third (Q3) quartiles. Values above Q3 indicate potential extreme consumption of fuel oil. This interquartile range isolates observations within a dataset's main distribution by proportion unaffected by extreme values. It is a robust, statistically grounded approach to detecting anomalies across large, heterogeneous datasets like maritime operational data.

Focusing only on points above the IQR upper fence warranted deeper examination through SHAP analysis. SHAP values quantify feature impact on individual predictions, offering insights obscured in aggregate metrics. Comparing SHAP patterns for extreme versus normal predictions can uncover influential factors differentially affecting infrequent high FOC cases. These aid in understanding what distinguishes extreme events.

3.3.3. SHAP Model Explanation

To interpret the black box of machine learning and to reveal the logic of how our model makes the prediction, this research utilized the Explainable Artificial Intelligence (XAI) methodology. Various techniques have been developed to provide explanations for individual predictions from opaque black-box models, such as LIME [40], which fits an interpretable linear model locally, and Anchors [41], which determine threshold conditions. However, SHAP (SHapley Additive exPlanations) [42] values offer a theoretically grounded game-theoretic approach that provides globally consistent and locally accurate feature attribution values for any model type.

As the tree ensemble machine learning algorithm was employed for this research's fuel consumption predictive modeling task, SHAP's specially adapted TreeSHAP method [43] was ideally suited, as it leverages the tree structure to connect global importance weights to local attribute effects for individual predictions.

In this study, we implemented SHAP (Shapley Additive Explanations) explanation techniques to enhance the interpretability of our fuel oil consumption (FOC) prediction model. SHAP is a powerful tool that provides insights into the contribution of each feature or variable to the model's predictions. Using SHAP explanations, we aimed to shed light on the factors influencing FOC predictions and gain a deeper understanding of the model's decision-making process.

SHAP operates by demystifying black-box models and revealing the rationale behind their predictions. It builds on the Game Theory concept of Shapley values [44] to interpret machine learning models. In Game Theory, Shapley values are used to distribute rewards fairly among cooperative players. In the context of model interpretability, each feature is like a player, and the prediction model becomes the "game."

To calculate the SHAP values, we employed the Shapley value concept from cooperative game theory [45]. Suppose that there is a set of input $X = \{x_1, x_2, I, x_n\}$ and a machine learning model v for every subset of the inputs, and S is the subset of X with the size of $k(S)$, so that $v(S)$ is the value of the subset. Then, the Shapley value for a specific feature is estimated as the following:

$$\varphi_x(v) = \frac{1}{n}\sum_{s} \frac{[v(S \cup \{x\}) - v(S)]}{\binom{n-1}{k(S)}}, \tag{1}$$

where $[v(S \cup \{x\}) - v(S)]$ is the marginal contribution of x for a given subset S. This calculation is repeated for all observations in the data set, resulting in a set of feature importance values for each observation. Once all the feature importance values are calculated, they can be used to interpret the importance of each feature for each observation in the data set.

Adapted from [37], Figure 3 shows the individual prediction outcome of f(x) = 10 can be broken down by incorporating the combined contribution value (which is the sum of Shapley values) from all the features, resulting in a value of 1.6 + 0.7 − 2.9 − 0.9 = −1.5. This is then added to the model's fixed base value of 11.5. In the case of regression, the base value represents the average of the target variable across all data points. Consequently, following the revelation of these contribution values, the model's output becomes the prediction base value plus the summation of the Shapley values for the features. This allows us to quantify which feature has the most significant impact on the prediction for that specific individual forecast. Furthermore, the accumulation of attributions from all the features provides a comprehensive explanation of the model at a global level.

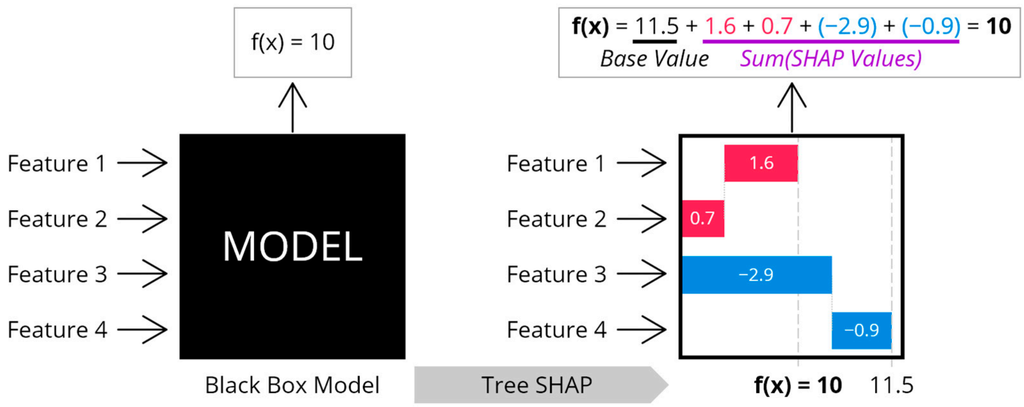


Figure 3. TreeSHAP concept [37] (red color to indicate the positive Shapley Value and blue color for the negative one).

SHAP values quantify the impact of each feature on the model's predictions by attributing a numerical value to each feature. Positive SHAP values indicate that a particular feature increases the predicted FOC, while negative values suggest a decrease in the predicted FOC. These values provide a clear understanding of how each feature contributes to the overall prediction.

Once we obtained the SHAP values of every single instance, we visualized them using plots such as SHAP summary plots, SHAP dependence plots, and SHAP interaction plots. These visualizations allowed us to interpret the effects of individual features on the FOC predictions and understand any nonlinear relationships or interactions between features. By implementing SHAP explanations, we aimed to address the black-box nature of our FOC prediction model and provide transparent and interpretable insights into the factors driving the predictions. These explanations are crucial for gaining stakeholders' trust in the model's predictions, complying with regulatory requirements, and making informed decisions based on the model's insights.

4. Results and Discussion

On the shores of discovery, this chapter unveiled the research findings and leveraged the discussion about the results. It commences with an assessment of the predictive model's performance with regression-based machine learning models, highlighting FOC predictability and the detailed hyperparameter tuning settings with randomized search cross-validation. Then, it delves into the model interpretability using SHAP to better investigate the FOC dynamics in general and specific manners. This chapter reflects the empirical findings of the analysis and then expands on the implications and actionable insights for efficient and sustainable maritime operations.

4.1. Fuel Oil Consumption Prediction

The first step of our analysis is to develop a prediction model that estimates the Fuel Oil Consumption (FOC) while considering both operational and environmental factors. To ensure the most suitable model for our objectives, we meticulously selected three machine learning algorithms: XGBoost Regressor, Catboost Regressor, and Gradient Boosting Regressor. These models were chosen for their ability to provide accurate predictions while accommodating our need for interpretability, aligning perfectly with our research goals. With these model candidates in place, we conducted a comprehensive comparative study to evaluate their performance and identify the most effective predictor for FOC estimation. Seen in Table 5 are the hyperparameter tuning settings for all models.

Table 5. Hyperparameter settings with randomized search CV.

Model	Parameters		
	Grid Parameters	**Grid Values**	**Best Values**
XGBoost Regressor	n_estimators	: [100, 200, 300, 400]	{'n_estimators': 400,
	reg_lambda	: [0, 0.1, 0.5, 1.0]	'reg_lambda': 0.5,
	reg_alpha	: [0, 0.1, 0.5, 1.0]	'reg_alpha': 0,
	min_child_weight	: [1, 2, 3, 4]	'min_child_weight': 2,
	max_depth	: [3, 4, 5, 6]	'max_depth': 6,
	learning_rate	: [0.01, 0.1, 0.2, 0.3]	'learning_rate': 0.2,
	gamma	: [0, 0.1, 0.2, 0.3]	'gamma': 0,
	colsample_bytree	: [0.6, 0.7, 0.8, 0.9, 1.0]	'colsample_bytree': 0.6}
CatBoost Regressor	n_estimators	: [100, 200, 300, 400]	{'n_estimators': 400,
	max_depth	: [3, 4, 5, 6]	'max_depth': 6,
	learning_rate	: [0.01, 0.1, 0.2, 0.3]	'learning_rate': 0.3,
	l2_lear_reg	: [1, 3, 5, 7]	'l2_leaf_reg': 5,
	bagging_temperature	: [0.0, 1.0, 2.0, 3.0]	'bagging_temperature': 2.0}
Gradient Boost Regressor	n_estimators	: [100, 200, 300, 400]	{'n_estimators': 400,
	min_samples split	: [2, 3, 4, 5]	'min_samples_split': 5,
	min_samples_leaf	: [1, 2, 3, 4]	'min_samples_leaf': 1,
	max_features	: ['auto', 'sqrt', 'log2']	'max_features': 'log2',
	max_depth	: [3, 4, 5, 6]	'max_depth': 6,
	loss	: ['ls', 'lad', 'huber']	'loss': 'huber',
	learning_rate	: [0.01, 0.1, 0.2, 0.3]	'learning_rate': 0.3,
	alpha	: [0.0, 0.1, 0.5, 1.0]	'alpha': 0.5}

Using all the data that we aimed to train, extensive hyperparameter tuning is performed on the three selected models to optimize their performance for fuel oil consumption prediction. This process involved specifying wide parameter ranges and adjusting settings tailored to each model's distinct characteristics.

Such meticulous tuning is critical to ensure that the models are fine-tuned to provide the most accurate predictions while leveraging their unique capabilities, ultimately enhancing the accuracy of our research results. Turned out, the XGBoost regressor was proven to be the best-performing predictor compared to the CatBoost Regressor and Gradient Boosting Regressor, as seen in Table 6.

Table 6. Performance Metrics of the Compared Models.

Model	R^2 Score		RMSE		MAE	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**
XGBoost Regressor	0.99	0.95	9.47	19.12	7.09	11.62
CatBoost Regressor	0.96	0.94	16.04	21.03	11.5	13.69
Gradient Boost Regressor	0.96	0.93	16.77	22.89	8.79	13.68

To further improve its performance, we trained the XGBoost Regressor using the optimized hyperparameters determined through a randomized search CV with a wider range of values to be tuned, as seen in Table 7.

Table 7. Hyperparameter settings for XGBoost Regressor with randomized search CV.

Grid Parameter	Grid Values	Best Value
n_estimator	: [100, 200, 300, 400, 500]	300
max_depth	: [2, 3, 5, 7, 10]	10
learning_rate	: [0.005, 0.01, 0.1, 0.3]	0.1
reg_alpha	: [0.1, 1, 10, 50, 100, 200]	1
min_child_weight	: [1, 5, 10, 25, 50]	10
subsample	: [0.5, 0.75, 1.0]	1.0
colsample_bytree	: [0.8, 0.9, 1.0]	0.8
reg_lambda	: [0.1, 1, 10, 50, 100, 200]	1

With the best parameter, we evaluate the XGBoost Regressor model performance on the held-out test data, and the regression model achieved strong prediction metrics, as seen in Table 8. It recorded an R-squared value of 0.95, indicating the model explained 95% of the variance in actual fuel consumption values. The RMSE of 18.54 kg/h further confirms the predictive ability, with errors comparable to typical consumption deviations. Additional MAE of 10.78 kg/h verifies accurate estimation of central tendencies without disproportionate impacts of outliers.

Table 8. XGBoost performance metrics.

R^2 Score		RMSE		MAE	
Train	Test	Train	Test	Train	Test
0.99	0.95	8.73	18.54	6.14	10.78

These quantitative performance metrics verify the model's high quality for predicting fuel oil consumption. However, to qualitatively demonstrate predictive power, a line plot seen in Figure 4 was generated comparing predicted versus actual consumption values for test vessels and voyages.

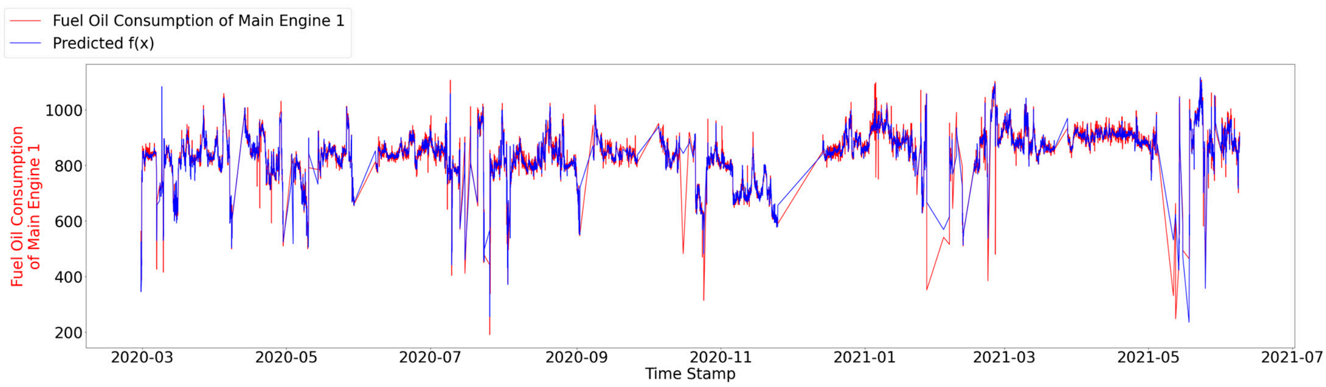


Figure 4. Actual FOC and the predicted FOC.

The close alignment between the two lines, with the predicted values nearly overlaying the actual reported fuel usage, provides visual verification of the model's ability to accurately forecast consumption levels across new unseen operational patterns and conditions. This validates the efficacy of the optimized XGBoost Regressor approach for the focal prediction task.

4.2. SHAP Global Explanation

Beyond predictive accuracy, the interpretability of model decisions is critical to guiding efficient operations. SHAP values were therefore computed to provide global feature importance explanations of the consumption estimation model. The SHAP approach quantifies the impact of each attribute on predictions, shedding light on primary drivers according to the learned patterns. This section delves into SHAP results under different contexts to comprehensively explore influential factors. Specifically, global explanations will be presented for overall fuel usage as analyzed in Section 4.2.1, extremely high consumption cases in Section 4.2.2, and region-specific extremely high FOC in Section 4.2.3. Together, these explain the prediction model at different levels and contexts to facilitate sustainability efforts through transparency into key consumption determinants.

4.2.1. Global Explanation of Overall Data

To begin, a global overview of feature importance for fuel oil consumption predictions across all data was established using SHAP explanations. A SHAP beeswarm summary plot as seen in Figure 5 and a mean absolute SHAP value bar plot as seen in Figure 6 were generated based on the prediction model. The beeswarm plot shows the distribution of SHAP values for each feature, visualizing their individual effects. The average draft of the vessel (aft and fore) had the strongest impact, followed by total wave height, relative wind speed, and speed over the ground.

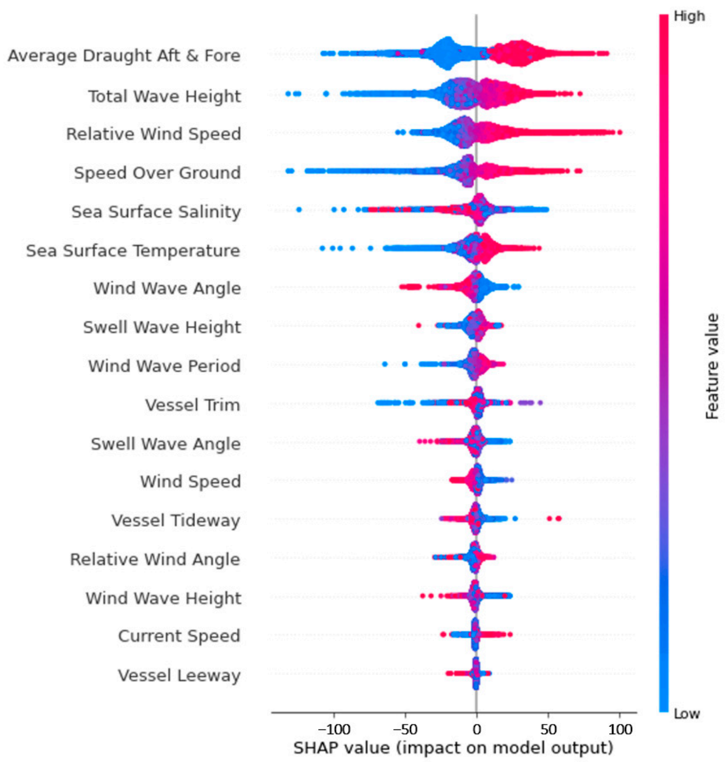


Figure 5. SHAP Beeswarm Summary Plot.

In the beeswarm plot that is seen in Figure 5, features towards the right span a wider range of SHAP values, indicating their positive impact on the prediction value. The color represents the actual value of the feature. So, for example, the average draft influenced predictions both positively and negatively to a great degree, and the higher its actual value, the more it tends to impact the prediction value of FOC positively. Features with points clustered around zero, like current speed or vessel leeway, have little overall effect.

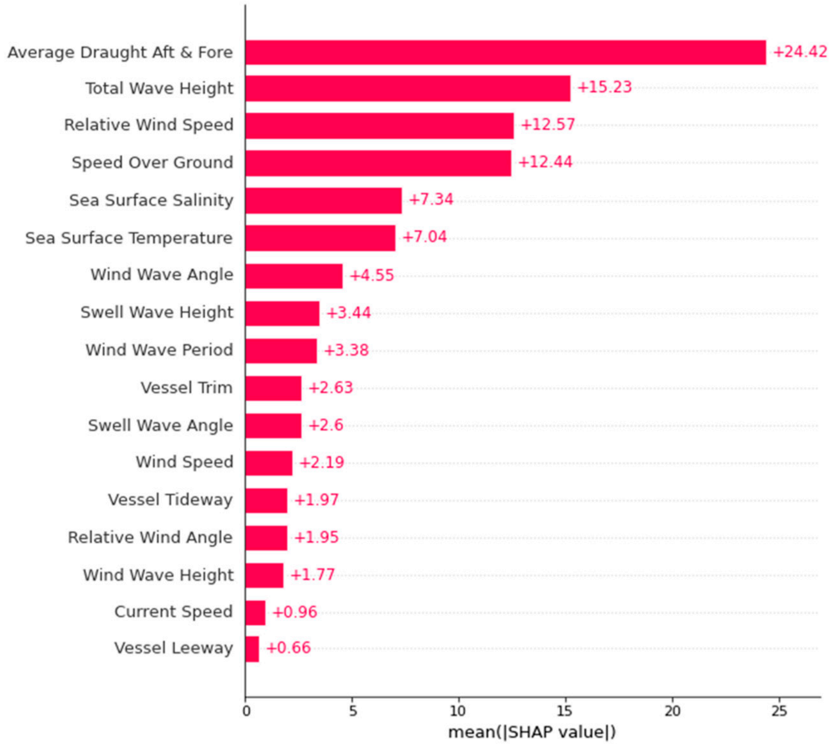


Figure 6. SHAP mean absolute SHAP value bar plot.

The mean absolute SHAP bar plot that is seen in Figure 6 meanwhile summarizes the average influence of each attribute, with longer bars corresponding to features that more consistently drive consumption up or down across the dataset. Together, these plots provide an informative high-level overview of drivers as perceived by the prediction model with automated feature importance selection.

4.2.2. Global Explanation of Extremely High FOC

Periods of extremely high FOC potentially signify operational inefficiencies or impacts of unusual environmental conditions that warrant focused inspection beyond typical usage patterns. We hypothesized that the influence of predictive factors may differ during such conditions that caused the extreme consumption of fuel oil.

To test this, SHAP explanations were generated focusing only on consumption values above the third quartile threshold identified through IQR analysis. Isolating these outliers allowed testing of whether predictive factor importance differed during atypical episodes versus general usage, helping explain the causes of abnormal high consumption scenarios.

The resultant feature importance ranking for extremely high FOC cases, as seen in Figures 7 and 8, saw some notable changes compared to general consumption. Overall changes are concluded in Table 9. The rise of relative wind speed from third to top-ranked feature provides key insight, as this attribute represents both environmental winds and vessel operations. Its prominence under outliers suggests periods of abnormally high fuel consumption were driven most strongly by needing to overcome significantly higher mechanical power requirements under very windy relative conditions. This change is further contextualized by the shifts seen in speed over ground and average draft. Speed over ground climbed from fourth to second, indicating vessels may have pushed operational limits to maintain schedules in strong opposing winds.

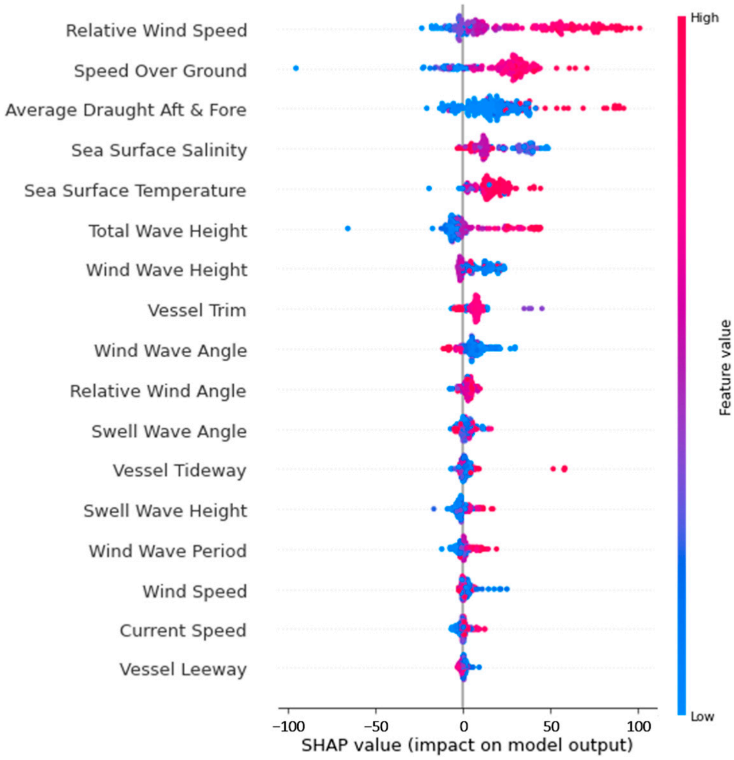


Figure 7. SHAP beeswarm summary plot during extremely high FOC episodes.

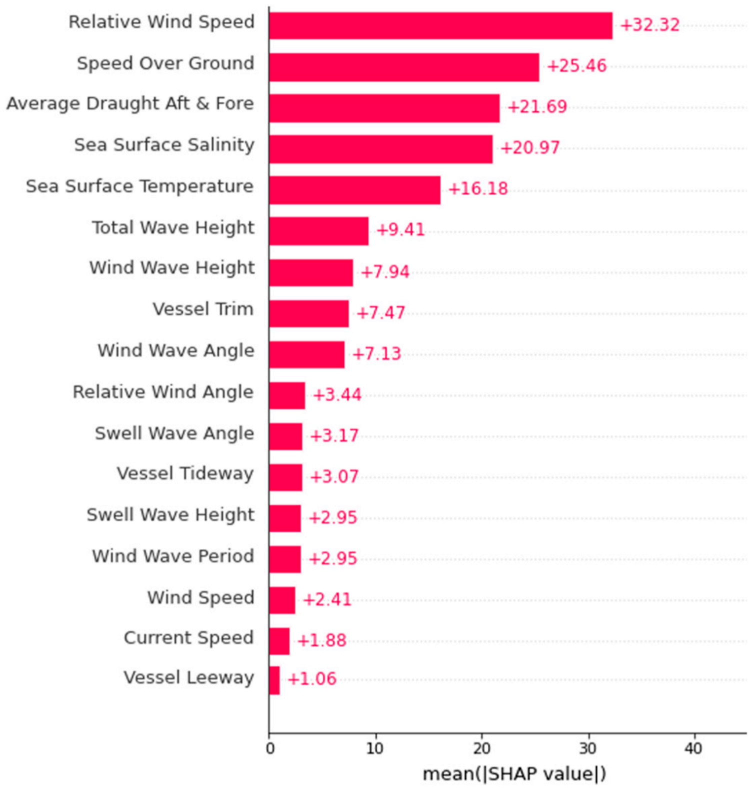


Figure 8. SHAP mean absolute SHAP value bar plot during extremely high FOC episodes.

Table 9. Comparison table of feature importance on overall data and extremely high FOC.

Feature Importance	Ranking		Mean (∣SHAP Value∣)		Mean (SHAP Value)	
	All Data	High	All Data	High	All Data	High
Relative Wind Speed	#03	#01	12.57	32.32	−0.30	29.92
Speed Over Ground	#04	#02	12.44	25.46	−1.84	22.04
Average Draft (Aft and Fore)	#01	#03	24.42	21.69	3.64	19.88
Sea Surface Salinity	#05	#04	7.34	20.97	−0.60	20.90
Sea Surface Temperature	#06	#05	7.04	16.18	−0.03	15.97
Total Wave Height	#02	#06	15.23	9.41	0.52	2.36
Wind Wave Height	#15	#07	1.77	7.94	−0.33	7.13
Vessel Trim	#10	#08	2.63	7.47	0.51	6.94
Wind Wave Angle	#07	#09	4.55	7.13	−0.28	5.96
Relative Wind Angle	#14	#10	1.95	3.44	−0.65	2.99
Swell Wave Angle	#11	#11	2.60	3.17	−0.60	2.17
Vessel Tideway	#13	#12	1.97	3.07	0.00	2.38
Swell Wave Height	#08	#13	3.44	2.95	−0.54	−0.19
Wind Wave Period	#09	#14	3.38	2.95	0.05	0.90
Wind Speed	#12	#15	2.19	2.41	−0.03	2.14
Current Speed	#16	#16	0.96	1.88	−0.04	0.08
Vessel Leeway	#17	#17	0.66	1.06	−0.01	0.46

Meanwhile, the average draft fell from first to third-ranked. This suggests draft-dependent static resistance was less influential than dynamic factors like wind and operational speeds under outlier scenarios. Together, these ranking changes reinforce the interpretation that extreme fuel usage was dictated most prominently by a vessel's need to counter very high total resistance from combined wind and self-propelled water flows rather than ambient static resistances alone. This helps explain disproportionate fuel burn in some outliers versus typical operations.

Sea surface salinity and temperature both maintained rankings within the most influential features, moving only slightly from 5th to 4th place and 6th to 5th, respectively, between the general and outlier SHAP explanations. This stability indicates ocean condition attributes like salinity and temperature levels remain important determinants of fuel usage, even under extremely high consumption scenarios.

Their consistent significance provides an indication that variations in sea properties could underlie some outlier events. The minor ranking adjustments also suggest salinity and temperature may still affect fuel consumption but are perhaps less directly sensitive to abnormal operations compared to factors like wind speeds that dominate under outlier conditions.

Total wave height, which was assessed second overall, dropped to seventh in the outlier analysis. This suggests that while wave impacts are consistently important, they may be lesser determinants of fuel usage during unusually high consumption periods compared to direct mechanical power needs accommodated by wind speeds. The shifted feature importance identifies potential optimization focus areas to mitigate consumption outliers driven prominently by high relative wind conditions over wave impacts alone.

Additionally, wind wave height rose substantially in the rankings, from 15th when considering all data to 7th for extreme outliers. This shift indicates that wind-generated sea state, beyond total wave impacts, may play a more pronounced role in driving exceptionally high fuel usage. Periods with higher wind wave heights could reflect rougher sea conditions, necessitating more engine output from vessels. The increase in the influence of this predictor helps shed light on why specific outlier events saw an elevation in fuel consumption levels compared to usual operations.

With this targeted approach, the SHAP explanations yielded valuable insight, confirming that identifying atypical points for isolated investigation can reveal the shifted influence of predictors in non-routine operational scenarios. The overall comparison of the SHAP Feature Importance Rank of the overall data and the extremely high FOC conditions can be seen in Table 9.

4.2.3. Region-Specific Extreme FOC Explanation

In addition to analyzing outliers across all data, it was also valuable to inspect their spatial distribution patterns to identify any regional attributes contributing to abnormal fuel consumption. Mapping extremely high FOC points could reveal geography-tied trends beyond vessel-specific or global environmental factors. If outliers clustered within certain localized transit areas rather than dispersing randomly throughout the extensive sailings, it indicated regions meriting focused study. By visualizing the voyages on a map, we aimed to detect any outlier concentrations that may point to locale-dependent influences particular to those locations. This could guide targeted regional analyses to better understand causative features and optimize operations, especially within problematic transit corridors.

Among the high FOC points, it is found that most of them are closely clustered in two regions. As seen in Figure 9, there are two regions marked with red trajectories that indicate the extremely high FOC with a value above 1017.48 kg/h. These two clusters, respectively, belong to Voyage Number 4 and Voyage Number 5. It was suspected that during those voyages, when crossing the Strait of Malacca and the South China Sea on that specific trajectory, certain factors among the operational and environmental variables fluctuated out of the normal range.

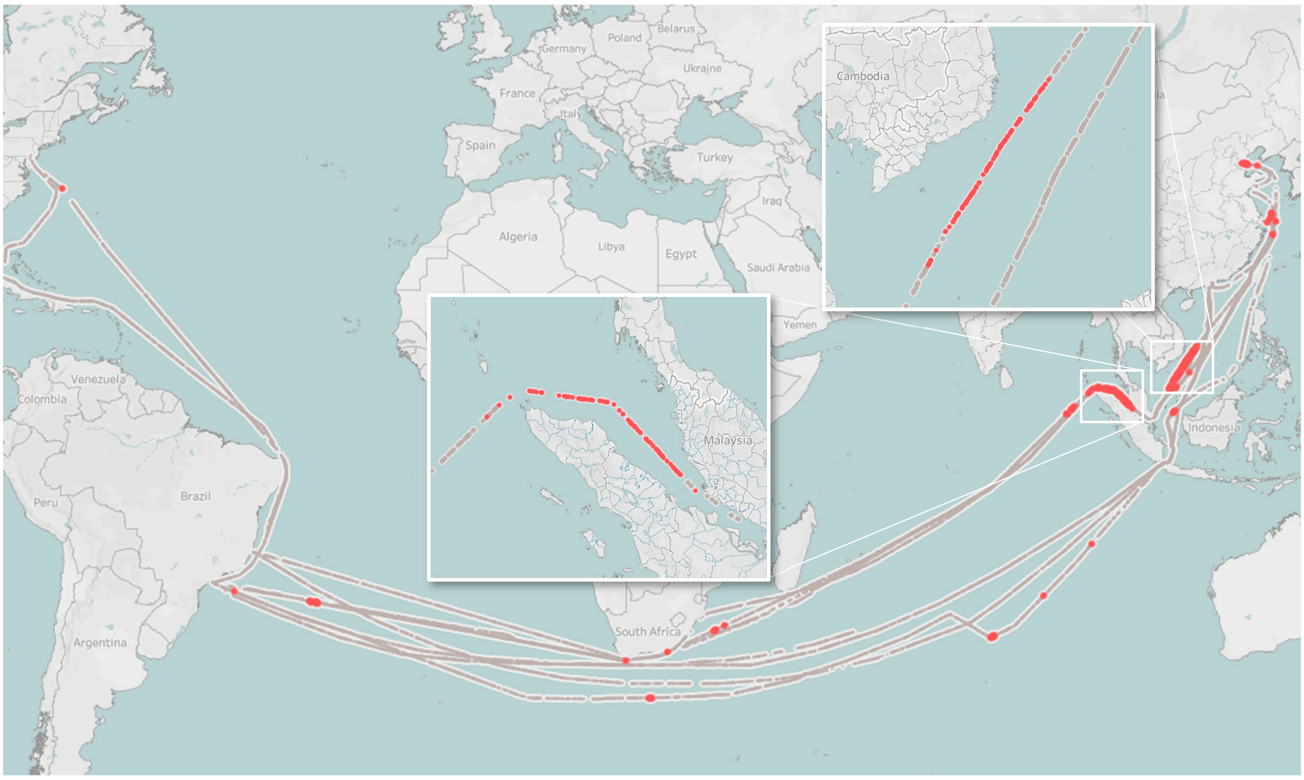
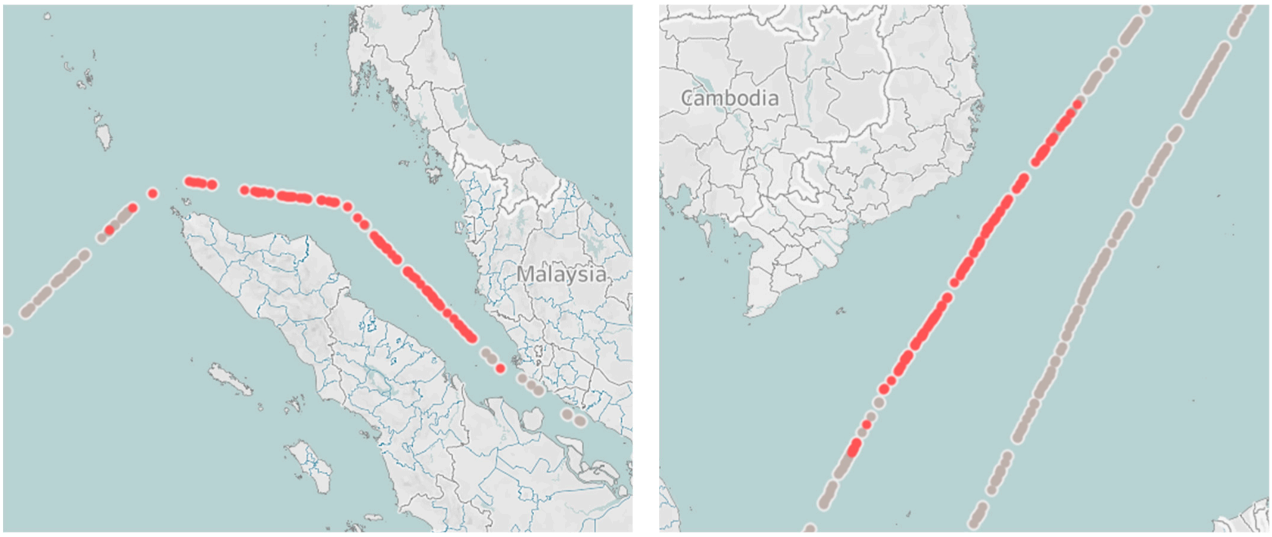


Figure 9. Extremely high FOC (Red Points) on vessel voyage trajectory.

By mapping out vessel trajectory using recorded latitude and longitude data over the 15-month voyage period, an interesting pattern emerged among points identified as extremely high FOC through IQR analysis. Visualizing the full voyage in gray, with outliers colored red, revealed two distinct regions containing unusually long connected strings of abnormal usage episodes. Upon inspection, these corresponded to routes passing through the Strait of Malacca between Indonesia and Malaysia during Voyage 4 and another route traversing the South China Sea near Vietnam in Voyage 5. The concentrated nature of outliers in just two localized areas suggested region-specific factors could be influencing fuel consumption beyond global patterns. This discovery prompted isolating SHAP explanations to investigate whether predictive importance differed for these locations compared to overall trends as well as other transit areas.

The findings highlighted the Strait of Malacca and the South China Sea as priority regions warranting closer examination through specialized SHAP analyses, intending to elucidate any location-tied drivers of fuel usage that set these apart from other sections of the extensive sailings.

Further investigation was carried out to better understand the cause of those extreme FOC values in the discovered regions. Having identified the Strait of Malacca (Figure 10a) and the South China Sea (Figure 10b) as concentrated areas of extreme FOC outliers, it was important to directly compare the predictive influence patterns within these regions to the overall trends. Rather than separate evaluations, overplotting the regional mean absolute SHAP values on the global SHAP importance bar plot allowed for a visual effectiveness comparison in a single figure. This facilitated the inspection of similarities and differences from typical usage across all locations. By depicting the Malacca Strait and South China Sea datasets in distinctive colors, any variances in average predictor impacts between even just these two pinpointed areas could also be immediately assessed.



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

Figure 10. Region-specific extremely high FOC Clusters (Red Points): (**a**) Strait of Malacca, between Indonesia and Malaysia (Registered with Voyage Number: 4); (**b**) South China Sea, near Vietnam (Registered with Voyage Number: 5).

Figure 11 provided an efficient preliminary screening to rapidly highlight regional variances prior to dedicated separated analyses, helping focus subsequent explorations on features displaying divergence across the fleets' diverse transit environments.

Overlapping region-specific SHAP lines aided the identification of where average prediction driver importance aligned or diverged when extreme outliers arose. Areas of mismatch suggested locale-dependent influential attributes that could help explain specific outlier incidents through concentrated regional studies. This set the stage for targeted SHAP evaluations of each region to uncover location-tied contextual predictors alongside globally consistent factors.

From Figure 11, it can be investigated how the regional SHAP plots revealed some notable variances from general trends. In the Strait of Malacca, sea surface salinity emerged as the predominant factor, with a mean absolute SHAP value of 31.84—markedly higher than the typical 5th-ranked salinity attribute. This suggests salinity levels in this locale disproportionately influenced fuel usage. Relative wind speed also showed a decreased importance there versus its usual third-place standing globally.

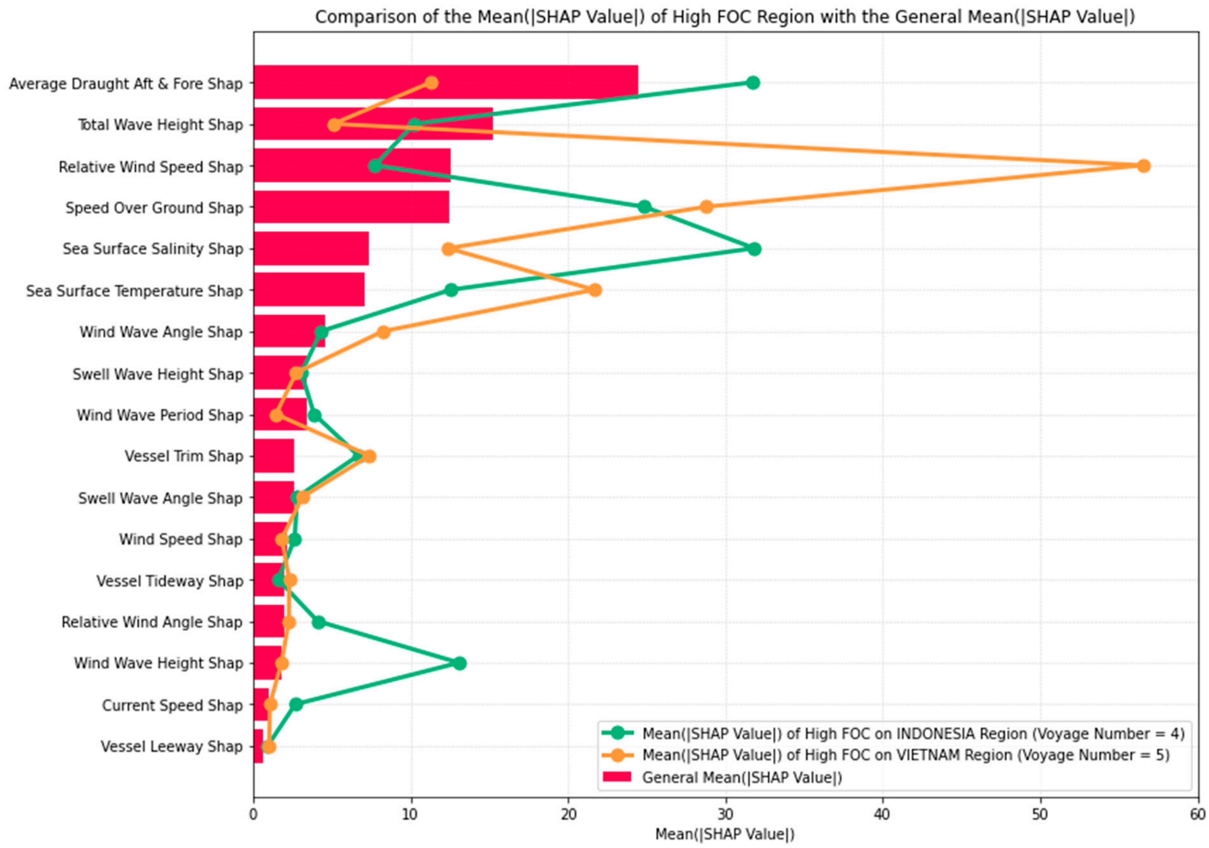


Figure 11. Comparison of overall mean (|SHAP Value|) and region-specific mean (|SHAP Value|).

Meanwhile, in the South China Sea, relative wind speed stood out the most, attaining the top spot with a value of 56.53, far exceeding its common third-place rank. This points to wind conditions having an unusually strong impact on fuel consumption within this region. Sea surface temperature, too, demonstrated elevated significance, moving up several positions compared to global rankings.

Lesser but still substantial divergences included the 6th place ranking attained by vessel trim in both outliers' areas versus a worldwide 10th, as well as the 13th rank of wind wave height in the Malacca Strait, much higher than usual. These discrepancies hint at locale-specific operational aspects and sea state attributes holding more sway over fuel demand when transiting these distinct corridors.

The region-wise plots provided initial signals highlighting where focused SHAP analyses could best identify contextual predictors to help explain abnormal fuel usage incidents particular to each emphasized location.

4.3. Discussion

In this discussion section, we synthesize the key findings from the preceding sections, connecting the dots between our research endeavors. With the foundational stages of our prediction model development, the insights gleaned from SHAP global explanations, and the examination of region-specific extreme FOC patterns, we aim to provide a holistic understanding of fuel oil consumption prediction within the maritime industry.

4.3.1. Prediction Model Development

The initial development and evaluation of an XGBoost Regressor model to accurately forecast FOC based on various inputs established a way to quantitatively examine relationships between predictors and fuel consumption. This established a starting point for subsequent interpretations. We evaluated three established machine learning algorithms:

Gradient Boosting Regressor, CatBoost Regressor, and XGBoost Regressor. These algorithms have gained recognition for their predictive power in various domains.

Hyperparameters play a pivotal role in shaping the model's behavior, and fine-tuning them can significantly enhance its performance. Thus, we involved an exhaustive exploration of the hyperparameter space, setting a wide range of values for each parameter of the three prediction models. The performance metrics utilized to evaluate the prediction performances are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2) values.

Upon acquiring XGBoost Regressor as the best model, we engaged in hyperparameter tuning; with the hyperparameters optimized, we took it a step further by leveraging these values for even finer-tuned settings. This meticulous process aimed to maximize the predictive capabilities of our XGBoost Regressor.

This attempt enabled us to accurately predict FOC while considering both operational (controllable) and environmental (uncontrollable) factors, proven by its performance metrics considered to show accurate prediction performance (R^2 = 0.95, MAE = 10.78 kg/h). The subsequent sections reveal the outcomes of this model development journey and the valuable insights gained into the factors influencing FOC.

4.3.2. SHAP Global Explanation

SHAP (SHapley Additive exPlanations) emerges as a powerful tool in the realm of Explainable Artificial Intelligence (XAI) for the maritime industry, offering essential insights into predictive models. One of the key features of SHAP is its ability to generate a Summary Plot, a valuable resource for understanding the impact of different features on the predicted output. This plot presents feature importance based on Shapley Values, shedding light on which factors exert the most influence on Fuel Oil Consumption (FOC). Such feature attribution emphasizes the fundamental utility of SHAP in our predictive model development. By clearly outlining the feature importance, SHAP enables us to pinpoint and interpret the key drivers of FOC prediction.

This transparency, derived from XAI, aids stakeholders in the maritime sector in comprehending the factors influencing FOC and, consequently, making informed decisions for energy efficiency. Using SHAP values to interpret the model's predictions helped uncover how factors differentially influenced expected FOC under changing conditions.

Overall, the controllable operational variable that affected the FOC prediction the most is the average draft (Aft and Fore), and the uncontrollable environmental variable is the Total Wave Height. The identification of key controllable and uncontrollable drivers provided a baseline understanding to ultimately help guide the optimization of the operational feature and mitigation of the environmental one.

Moreover, the significance of features like average draft, total wave height, relative wind speed, and speed over ground ranked at the top for both the SHAP explanation of overall data and the extremely high FOC-specific points, as these variables are intuitively associated with FOC. The relevance of these features in maritime operations has been well-established, with their influence on fuel consumption widely recognized by experts. While it is no surprise that these features are considered important, SHAP provides the means to quantify their individual contributions precisely. It goes beyond the apparent influence of these variables to reveal the extent to which they affect FOC.

However, our analysis has unearthed a noteworthy revelation: sea surface salinity and sea surface temperature also hold prominent positions in the feature importance ranking. Although domain experts may find this unexpected at first glance, SHAP illuminates how these factors indirectly impact FOC through their influence on ship resistance.

This intriguing finding may indeed be linked to their indirect impact on vessel resistance through factors such as hull and propeller fouling.

- Fouling Influence: Higher sea surface temperatures and specific salinity levels can create more favourable conditions for the growth and distribution of fouling organisms. The accumulation of fouling on a ship's hull and propeller increases hydrodynamic drag and leads to higher resistance, necessitating increased fuel consumption for the ship to maintain its operational efficiency. These conditions can also affect the kinematic viscosity of seawater, further elevating resistance.

- Kinematic Viscosity: The alteration of seawater properties by temperature and salinity influences kinematic viscosity, impacting the flow of water around the ship's hull. Higher kinematic viscosity can result in elevated resistance, requiring more energy to propel the vessel. The pronounced effect of sea surface salinity and temperature on this viscosity may explain their high-ranking impact on fuel consumption.

However, it is important to note that our primary focus in this study was leveraging Explainable Artificial Intelligence (XAI) to enhance fuel oil consumption prediction and promote transparency in operational optimization within the maritime industry.

4.3.3. Region-Specific Extreme FOC Explanation

Focusing on regions exhibiting concentrated outlier episodes allowed location-specific differences in influential predictors to surface. Leveraging the spatial context helped address sources of variability not detectable from uniformly considering all voyages; insights into locale-dependent consumption behaviors aimed to enable targeted mitigations.

As seen in Table 10 is the comparison of the entire SHAP analysis when considering overall data, then separately examined the extremely high FOC and investigated region-specific SHAP of regions that clustered in the Strait of Malacca, between Indonesia and Malaysia (Registered with Voyage Number: 4) and the South China Sea, near Vietnam (Registered with Voyage Number: 5).

Table 10. Comparison of SHAP feature importance of all data, extremely high FOC points, and region-specific points.

Feature Importance	Feature Importance Ranking				Mean (∣SHAP Value∣)				Mean (SHAP Value)			
	Overall	High	ID	VN	Overall	High	ID	VN	Overall	High	ID	VN
Average Draught (Aft and Fore)	#01	#03	#02	#05	24.42	21.69	31.73	11.30	3.64	19.88	30.68	9.80
Total Wave Height	#02	#06	#06	#08	15.23	9.41	10.24	5.10	0.52	2.36	0.49	−0.23
Relative Wind Speed	#03	#01	#07	#01	12.57	32.32	7.74	56.53	−0.30	29.92	3.42	56.46
Speed Over Ground	#04	#02	#03	#02	12.44	25.46	24.84	28.76	−1.84	22.04	21.04	26.25
Sea Surface Salinity	#05	#04	#01	#04	7.34	20.97	31.84	12.39	−0.60	20.90	31.84	12.39
Sea Surface Temperature	#06	#05	#05	#03	7.04	16.18	12.57	21.67	−0.03	15.97	12.55	21.20
Wind Wave Angle	#07	#09	#09	#06	4.55	7.13	4.32	8.25	−0.28	5.96	1.93	8.25
Swell Wave Height	#08	#13	#12	#10	3.44	2.95	3.06	2.74	−0.54	−0.19	0.56	−1.54
Wind Wave Period	#09	#14	#11	#15	3.38	2.95	3.88	1.45	0.05	0.90	0.04	0.86
Vessel Trim	#10	#08	#08	#07	2.63	7.47	6.68	7.38	0.51	6.94	6.44	7.15
Swell Wave Angle	#11	#11	#13	#09	2.60	3.17	2.79	3.12	−0.60	2.17	1.42	2.30
Wind Speed	#12	#15	#15	#13	2.19	2.41	2.64	1.83	−0.03	2.14	2.41	1.45
Vessel Tideway	#13	#12	#16	#11	1.97	3.07	1.59	2.29	0.00	2.38	0.84	1.51
Relative Wind Angle	#14	#10	#10	#12	1.95	3.44	4.10	2.27	−0.65	2.99	3.54	2.05
Wind Wave Height	#15	#07	#04	#14	1.77	7.94	13.08	1.80	−0.33	7.13	13.07	−0.02
Current Speed	#16	#16	#14	#16	0.96	1.88	2.68	1.06	−0.04	0.08	−0.89	0.78
Vessel Leeway	#17	#17	#17	#17	0.66	1.06	0.92	1.02	−0.01	0.46	0.56	0.10

As distinctive drivers were exposed, recommendations could be developed. Suggestions focused on optimally controlling predictive operational parameters and proposing mitigation strategies for prominent unalterable factors, especially within challenging outlier-prone areas.

Table 11 summarizes the key takeaways that mainly suggested the optimization of operational measures and mitigation of environmental conditions that aim for energy efficiency through fuel oil consumption management.

Presenting successive findings in an accessible manner supported the overarching goal of aiding industry decision-making. The multi-scale investigations and interpretations strived to offer a more nuanced perspective of the complex socio-environmental determinants shaping maritime energy usage to ultimately enhance efficiency through informed action.

Table 11. Key takeaways.

	General Condition	Extreme High FOC	Strait of Malacca	South China Sea
Operational Optimization	Draft and speed control and monitoring Load adjustments	Speed adjustment given the wind conditions	Speed adjustment on the region that is suspected to have busy marine traffic	Speed reduction in response to high wind speed
Environmental Mitigation	Minimize voyage through high waves	Rerouting in high wind conditions	Monitor the distribution of salinity around the strait and a narrow shallow water area	Investigate the regional wind pattern and Munson effect

5. Conclusions

In the maritime industry, achieving optimal vessel fuel oil consumption (FOC) is pivotal, impacting both energy efficiency and operational effectiveness. Our study has undertaken this challenge by harnessing the predictive power of the XGBoost Regressor, yielding a model that boasts impressive accuracy with a mean absolute error (MAE) of 10.78 kg/h and an R-squared of 0.95.

Crucially, our research extends beyond predictive modeling, delving into the complex interplay of oceanographic (uncontrollable) and maneuverability (controllable) factors affecting FOC. Through SHAP analysis, we've enhanced the interpretability of our model, elucidating the prominent roles of "Average Draught (Aft and Fore)" and "Relative Wind Speed" as key influencers.

One of the highlights of our investigation is the in-depth analysis of conditions characterized by extremely high FOC. Notably, we uncovered distinct shifts in SHAP feature importance rankings within the closely clustered high FOC trajectories in the Strait of Malacca, located between Indonesia and Malaysia, and the South China Sea near Vietnam. These findings illuminate the critical importance of region-specific factors in FOC optimization and decision-making.

Our research introduces a pioneering framework that combines a robust prediction model with SHAP analytics, offering a deeper understanding of the intricate dynamics underpinning FOC. By focusing on the influence of region-specific factors, this framework provides actionable insights for enhancing energy efficiency, refining operational strategies, and mitigating sea resistance—an invaluable contribution to advancing decision-making processes within the maritime industry.

In conclusion, our study not only presents a powerful predictive model but also elucidates the complex landscape of FOC, emphasizing the need to consider both controllable and uncontrollable factors, particularly in high FOC scenarios. This research empowers vessel operators with the knowledge needed to optimize energy efficiency and operational performance while navigating the challenges of sea resistance in diverse maritime regions, a pioneering framework combining robust prediction and transparency through SHAP explanations. A deeper exploration of region-specific contextualization advances knowledge beyond single-factor or vessel-centric views, instead incorporating environmental and operational spatial variability.

Insights support targeted efficiency strategies tailored to challenges across diverse maritime territories. Overall, this work informs planning and decision-making aimed at responsible shipping growth alongside tightening emissions regulations.

However, limitations include representing a single vessel, precluding fleet-wide comparisons. This could help identify optimization strategies with broader applicability. Other than that, supplementing the dataset with additional data sources like AIS, cargo load details, and high-frequency engine metrics provides opportunities to incorporate further dynamic attributes and develop even more accurate and spatially refined models. Finally, leveraging model outputs to simulate "what-if" scenarios applying various mitigation tactics under different conditions would offer tangible support for decision-making. For example, rerouting analyses or estimations of efficiency gains from retrofits/operational changes applied specifically within outlier-prone regions. Expanding the scope and real-world applicability testing represents exciting paths for continuing progress in fuel-efficient maritime operations.

In conclusion, this study presents a powerful yet transparent predictive modeling approach with the scope to advance fuel-efficient operations through understanding regionally distinct influential factors and abnormal behaviors.

Author Contributions: Conceptualization, M.P.H. and H.K.; methodology, J.L. and M.P.H.; software, M.P.H.; formal analysis, M.P.H.; resources, S.L.; data curation, S.L. and M.P.H.; writing—original draft preparation, J.L. and M.P.H.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the Ministry of Trade, Industry and Energy, Korea, under the "Regional Innovation Cluster Development Program (R&D) (P0025328)" supervised by the Korea Institute for Advancement of Technology (KIAT).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We could not share our data as is not available for public due to privacy and confidentiality.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. UNCTAD. *Trade and Environment Review 2023: Building A Sustainable and Resilient Ocean Economy Beyond 2030*; United Nations Conference on Trade and Development (UNCTAD): Geneva, Switzerland, 2023. Available online: https://unctad.org/system/files/official-document/ditcted2023d1_en.pdf (accessed on 15 May 2023).
2. DNV AS. *Energy Transition Outlook 2022: A Global and Regional Forecast to 2050*; DNV: Oslo, Norway, 2022.
3. IRENA. *A Pathway to Decarbonise the Shipping Sector by 2050*; International Renewable Energy Agency (IRENA): Abu Dhabi, United Arab Emirates, 2021. Available online: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2021/Oct/IRENA_Decarbonising_Shipping_2021.pdf (accessed on 15 May 2023).
4. DNV GL AS. *Energy Transition Outlook 2020: A Global and Regional Forecast to 2050*; DNV: Oslo, Norway, 2020.
5. Seddiek, I.S.; Ammar, N.R. Harnessing Wind Energy on Merchant Ships: Case Study Flettner Rotors Onboard Bulk Carriers. *Environ. Sci. Pollut. Res.* **2021**, *28*, 32695–32707. [CrossRef] [PubMed]
6. IBarreiro, J.; Zaragoza, S.; Diaz-Casas, V. Review of Ship Energy Efficiency. *Ocean Eng.* **2022**, *257*, 111594. [CrossRef]
7. Bayraktar, M.; Cerit, G.A. An Assessment on the Efficient Use of Hybrid Propulsion System in Marine Vessels. *World J. Environ. Res.* **2020**, *10*, 61–74. [CrossRef]
8. Mansoursamaei, M.; Moradi, M.; González-Ramírez, R.G.; Lalla-Ruiz, E. Machine Learning for Promoting Environmental Sustainability in Ports. *J. Adv. Transp.* **2023**, *2023*, 2144733. [CrossRef]
9. Akyuz, E.; Cicek, K.; Celik, M. A Comparative Research of Machine Learning Impact to Future of Maritime Transportation. *Procedia Comput. Sci.* **2019**, *158*, 275–280. [CrossRef]
10. Li, H.; Yang, Z. Incorporation of AIS Data-Based Machine Learning into Unsupervised Route Planning for Maritime Autonomous Surfave Ships. *Transp. Res. Part E Logist. Transp. Rev.* **2023**, *176*, 103171. [CrossRef]
11. Zis, T.P.V.; Psaraftis, H.N.; Ding, L. Ship Weather Routing: A Taxonomy and Survey. *Ocean Eng.* **2020**, *213*, 107697. [CrossRef]
12. Li, X.; Sun, B.; Jin, J.; Ding, J. Speed Optimization of Container Ship Considering Route Segmentation and Weather Data Loading: Turning Point-Time Segmentation Method. *J. Mar. Sci. Eng* **2022**, *10*, 1835. [CrossRef]
13. Du, Y.; Meng, Q.; Wang, S.; Kuang, H. Two-Phase Optimal Solutions for Ship Speed and Trim Optimization over a Voyage Using Voyage Report Data. *Transp. Res. Part B Methodol.* **2019**, *122*, 88–114. [CrossRef]
14. Ksciuk, J.; Kuhlemann, S.; Tierney, K.; Koberstein, A. Uncertainty in Maritime Ship Routing and Scheduling: A Literature Review. *Eur. J. Oper. Res.* **2023**, *308*, 499–524. [CrossRef]
15. Makridis, G.; Kyriazis, D.; Plitsos, S. Predictive Maintenance Leveraging Machine Learning for Time-Series Forecasting in the Maritime Industry. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020. [CrossRef]
16. Petry, L.M.; Soares, A.; Bogorny, V.; Brandoli, B.; Matwin, S. Challenges in Vessel Behavior and Anomaly Detection: From Classical Machine Learning to Deep Learning. In *Advances in Artificial Intelligence. Canadian AI 2020*; Lecture Notes in Computer Science; Goutte, C., Zhu, X., Eds.; Springer: Cham, Switzerland, 2020; Volume 12109. [CrossRef]
17. Laurie, A.; Anderlini, E.; Dietz, J.; Thomas, G. Machine Learning for Shaft Power Prediction and Analysis of Fouling Related Performance Deterioration. *Ocean Eng.* **2021**, *234*, 108886. [CrossRef]
18. Kretschmann, L. Leading Indicators and Maritime Safety: Predicting Future Risk with a Machine Learning Approach. *J. Shipp. Trade* **2020**, *5*, 19. [CrossRef]

19. Park, H.J.; Lee, M.S.; Park, D.I.; Han, S.W. Time-Aware and Feature Similarity Self-Attention in Vessel Fuel Consumption Prediction. *Appl. Sci.* **2021**, *11*, 11514. [CrossRef]
20. Soleymani, A.; Sharifi, S.M.H.; Edalat, P.; Sharifi, S.M.M.; Zadeh, S.K. Linear Modelling of Marine Vessels Fuel Consumption for Ration of Subsidized Fuel. *Int. J. Marit. Technol.* **2018**, *10*, 7–13. [CrossRef]
21. Uyanik, T.; Arslanoglu, Y.; Kalenderli, O. Ship Fuel Consumption Prediction with Machine Learning. In Proceedings of the 4th International Mediterranean Science and Engineering Congress, Antalya, Turkey, 25–27 April 2019. Available online: https://www.researchgate.net/profile/Tayfun-Uyanik/publication/332717845_Ship_Fuel_Consumption_Prediction_with_Machine_Learning/links/5ecd0608299bf12a632d479f/Ship-Fuel-Consumption-Prediction-with-Machine-Learning.pdf (accessed on 15 May 2023).
22. Ren, F.; Wang, S.; Liu, Y.; Han, Y. Container Ship Carbon and Fuel Estimation in Voyages Utilizing Meteorological Data with Data Fusion and Machine Learning Techniques. *Math. Program Eng.* **2022**, *2022*, 4773395. [CrossRef]
23. Uyanık, T.; Karatuğ, Ç.; Arslanoğlu, Y. Machine Learning Approach to Ship Fuel Consumption: A Case of Container Vessel. *Transp. Res. Part D Transp. Environ.* **2020**, *84*, 102389. [CrossRef]
24. Wang, S.; Ji, B.; Zhao, J.; Liu, W.; Xu, T. Predicting Ship Fuel Consumption Based on LASSO Regression. *Transp. Res. Part D Transp. Environ.* **2018**, *65*, 817–824. [CrossRef]
25. Peng, Y.; Liu, H.; Li, X.; Huang, J.; Wang, W. Machine Learning Method for Energy Consumption Prediction Of Ships in Port Considering Green Ports. *J. Clean. Prod.* **2020**, *264*, 121564. [CrossRef]
26. Zhao, S.; Yin, Q.; Chen, X.; Zhao, F.; Zhao, K.; Zheng, J. Influence of Different Machine Learning Algorithms on Prediction Model of Fuel Consumption of Inland Ships. In Proceedings of the 2021 6th International Conference on Transportation Information and Safety (ICTIS), Wuhan, China, 22–24 October 2021. [CrossRef]
27. Yan, R.; Wang, S.; Yuquan, D. Development of a Two-Stage Ship Fuel Consumption Prediction and Reduction Model for a Dry Bulk Ship. *Transp. Res. Part E Logist. Transp. Rev.* **2020**, *138*, 101930. [CrossRef]
28. Yuan, J.; Nian, V. Ship Energy Consumption Prediction with Gaussian Process Metamodel. *Energy Procedia* **2018**, *153*, 655–660. [CrossRef]
29. Moreira, L.; Vettor, R.; Soares, C.G. Neural Network Approach for Predicting Ship Speed and Fuel Consumption. *J. Mar. Sci. Eng.* **2021**, *9*, 119. [CrossRef]
30. Bui-Duy, L.; Vu-Thi-Minh, N. Utilization of a Deep Learning-Based Fuel Consumption Model in Choosing a Liner Shipping Route for Container Ships in Asia. *Asian J. Shipp. Logist.* **2021**, *37*, 1–11. [CrossRef]
31. Hu, Z.; Jin, Y.; Hu, Q.; Sen, S.; Zhou, T.; Osman, M.T. Prediction of Fuel Consumption for Enroute Ship Based on Machine Learning. *IEEE Access* **2019**, *7*, 119497–119505. [CrossRef]
32. Arrieta, A.B.; Díaz-Rodríguez, N.; Ser, J.D.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
33. Tiwari, R. Explainable AI (XAI) and Its Applications in Building Trust and Understanding in AI Decision-Making. *Int. J. Sci. Res. Eng. Manag.* **2023**, *7*, 1. [CrossRef]
34. Angelov, P.P.; Soares, E.A.; Jiang, R.; Arnold, N.I.; Atkinson, P.M. Explainable artificial intelligence: An analytical review. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2021**, *11*, e1424. [CrossRef]
35. Samek, W.; Wiegand, T.; Müller, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *arXiv* **2017**, arXiv:1708.08296.
36. Sarp, S.; Kuzlu, M.; Wilson, E.; Cali, U.; Guler, O. A Highly Transparent and Explainable Artificial Intelligence Tool for Chronic Wound Classification: XAI-CWC. *Preprints* **2021**, *1*. [CrossRef]
37. Kim, D.; Handayani, M.P.; Lee, S.; Lee, J. Feature Attribution Analysis to Quantify the Impact of Oceanographic and Maneuverability Factors on Vessel Shaft Power Using Explainable Tree-Based Model. *Sensors* **2023**, *23*, 1072. [CrossRef]
38. Kim, D.; Antariksa, G.; Handayani, M.P.; Lee, S.; Lee, J. Explainable Anomaly Detection Framework for Maritime Main Engine Sensor Data. *Sensors* **2021**, *21*, 5200. [CrossRef]
39. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Fransisco, CA, USA, 13–17 August 2016; pp. 787–794. [CrossRef]
40. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Fransisco, CA, USA, 13–17 August 2016. [CrossRef]
41. Ribeiro, M.T.; Singh, S.; Guestrin, C. Anchors: High-Precision Model-Agnostic Explanations. In Proceedings of the AAAI Conference of Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018. [CrossRef]
42. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017. Available online: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf (accessed on 15 May 2023).
43. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]

44. Shapley, L.S. A Value for N-Person Games. In *Contribution to the Theory of Games II*; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.
45. Hart, S. Shapley Value. In *Game Theory*; Palgrave Macmillan: London, UK, 1989; pp. 210–216.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.