Big Data Analytics for Improving Energy Efficiency in Short-Sea Shipping

Mohamed Abuella, Halmstad University, email:mohamed.abuella@hh.se

Abstract

To meet the urgent requirements for climate change mitigation, several proactive measures of energy efficiency have been implemented in the maritime industry. In developing a ship routing approach to minimize the fuel consumption, the operational and enviromental conditions data collection and preprocessing are imperative for building a reliable framework. The weather data for short-sea shipping might have not been in temporal and spatial resolutions that accurately representing the actual environmental conditions. Therefore, the available onboard data as well as the weather data extracted from external resources should be investigated and deployed effectively to model the physics of the vessel, and hence the vessel performance.
This paper presents the major findings of data analytics for the impacts of weather and operating conditions on the vessel's fuel consumption.

Index terms— Short-sea shipping, Energy efficiency, Weather data, Fuel consumption, Standard performance, Ship routing.

1 Introduction

What are the main factors that impact the vessel's fuel consumption for short-sea shipping?
Let's consider an illustrative case of a passenger ship with actual profiles of fuel consumption as illustrated in figure 1. This ship has main routes between a port on the South side and a port on the North side, on April 1^{st}, 2020. By associating the voyage operating and weather factors with this illustrative case, we can observe the following:
As the ship started her route at $15:38$ and then at the last part of the route it changed her course direction toward west which is against the wind, current and wave directions, i.e., 270^o (degree). Nevertheless, it could be obvious that the weather has impacted the ship's fuel consumption, the vessel's speed was also increased simultaneously when the vessel headed westward. Thus, at this combination of weather conditions and speed, the vessel's resistance has increased and hence the fuel consumption became higher by about 38% more than just the previous route that started at $15:08$.

Therefore, different factors of the voyage status and their combinations with the weather conditions should be all considered to perform a rational reasoning for the vessel's fuel consumption.

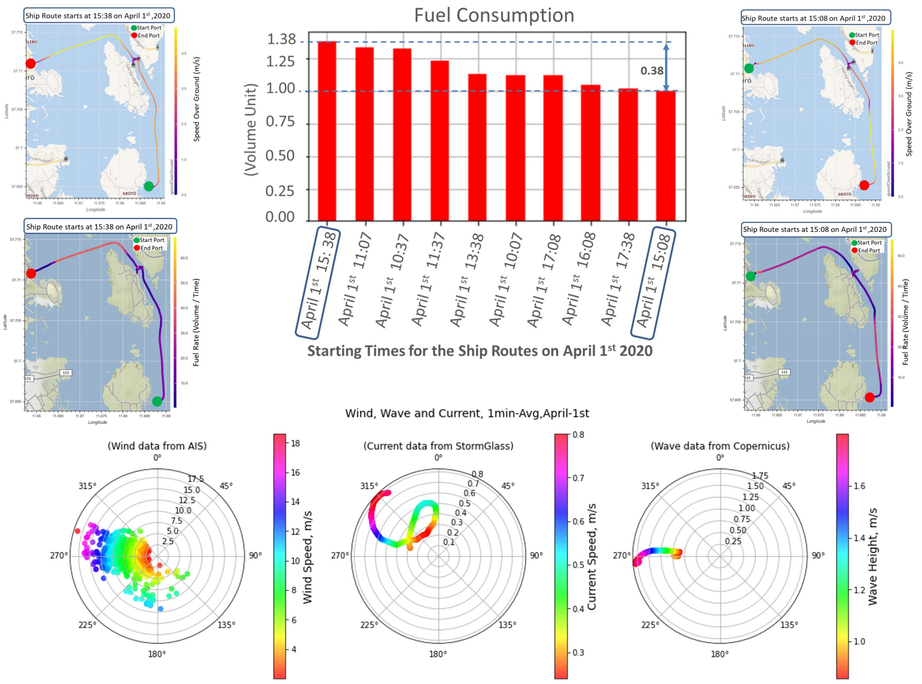


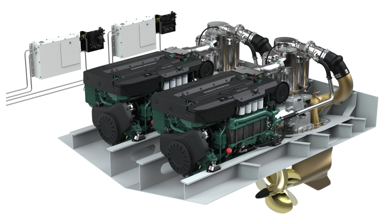
Figure 1: Case Study of Weather Impacts on Fuel Consumption, April 1st, 2020

2 Data Preparation and the Analysis Validation

2.1 Data Preprocessing

The ship's onboard data have been received from our industry partner CetaSol AB in Gothenburg [1]. It has been gathered over the entire year of 2020. It has a 3Hz frequency and records about the ship's position, course direction, and speed. It is also including some of operational and meteorological data, such as fuel rate, engine speed, torque, acceleration, wind speed and direction. Some information about the ship and its voyage can be found on Marine Traffic website [2]. For instance, the ship name: Buro, of type: passenger ship, with a ship size:(Length by Breadth) 19m by 6.41m, and the average speed is 8.2 knots (4.2 m/s), the picture of the vessel is provided as in figure 2.

(a) Buro ferry, photo by Owe Johansson [2].



(b) Engine Volvo Penta [3].

Figure 2: The vessel and its diesel engine.

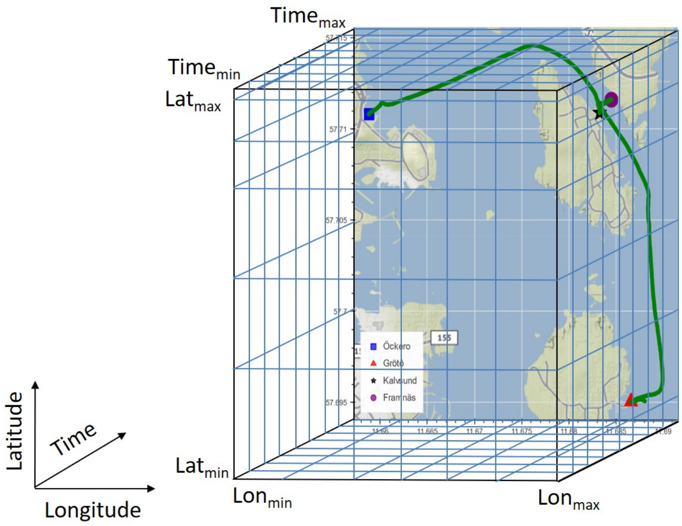


Figure 3: 3-Dimension Linear Interpolation

Other weather variables such as wave height and sea current speed and direction have been collected from external sources, Copernicus Marine Service [4] and Stormglass [5] APIs.

These external weather data are past forecasts (hindcasts), which have reanalysed to become hourly in temporal resolution and with 0.25 to 0.5 degree as a spatial resolution. Trilinear interpolation in time and space dimensions has been applied on external weather data to be more suitable for time and position frames of the given vessel routing. The schematic diagram of data interpolation in time and space is depicted by figure 3.
Therefore, the raw weather and onboard data are used in this analysis with two temporal resolution cases, 3-seconds and 1-minute in average.

2.2 The Analysis Validation

The analysis is divided to two parts, first the analysis with cruising-mode speed, and the second analysis with all vessel's speeds.
For the cruising speed, the weather analysis and validation of different weather data are carried out visually. This graphical analysis has been conducted based on the cruising mode to reduce the other vessel effects on the fuel consumption, and thus, producing graphs that can be then compared with the general ship's standard performance, as shown in figure 5.

These graphs are in form of the ship's standard performance, when the speed over ground is represented on x-axis, fuel on y-axis, while the weather variable as a third dimension, represented by the colorbar [6]. The standard performance graphs are also generated with data aggregated by fuel and speed over ground, in three ways of aggregations, mean, standard deviation, and count of data points.

Whereas, for vessel's speeds other than the cruising speed, it is difficult to validate the weather impacts visually. Thus, for sake of more clarification and

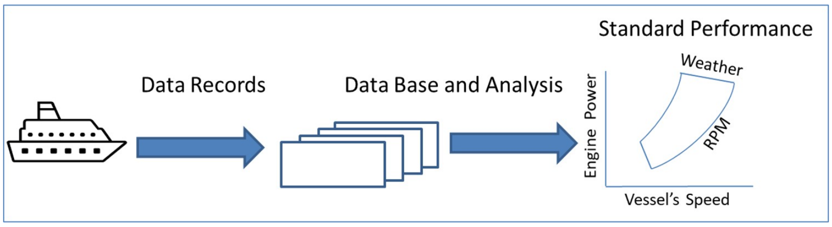
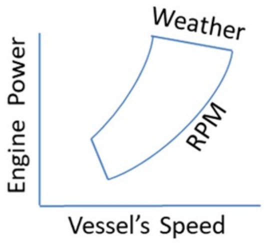
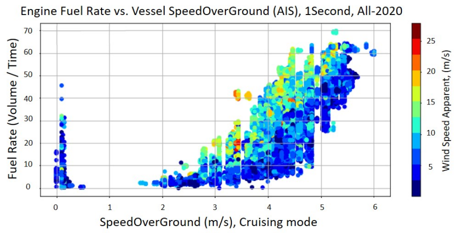


Figure 4: The vessel's data analytics.



(a) Ship standard performance [6].



(b) Vessel Buro's standard performance

Figure 5: The ship's standard performance graphs for the typical and the used cases

verification on how the weather variables interact to each other and influencing the vessel's fuel consumption, a regression model using artificial neural network (ANN) is utilized as an environment of several experiments with the main possible combinations of weather variables.

A comparison for the features importance with both cruising and all-speeds datasets is made as well. Instead of applying the ANOVA (i.e., analysis of variance) to the available features by using a linear regression model with feature polynomials to determine the regression intercept and coefficients, the Shapley values of the contribution of features to a predictive model are generated to determine the coalition or additive value of each feature to the overall prediction. The Shapley are illustrated by the beeswarm plots [7].

3 Summary of Major Findings

3.1 Remarks on the Analysis with Cruising-Mode Speeds

At the cruising mode to compare the vessel's performance with the standard performance, in this analysis numerous graphs have been generated for different weather variables from onboard data and external sources. The most important concluded remarks about the resulted graphs for the analysis of weather data as following:

- Firstly, the cruising mode only represents about 5% of the entire data. Therefore, it is not sufficient enough to rely on the cruising mode to analyze the weather impacts on the fuel consumption for short-sea shipping;

- The wind has a trend of 5 m/s with probability of frequency 12%, while the wind direction is most commonly from west degree of 200°, as shown in figure 6;

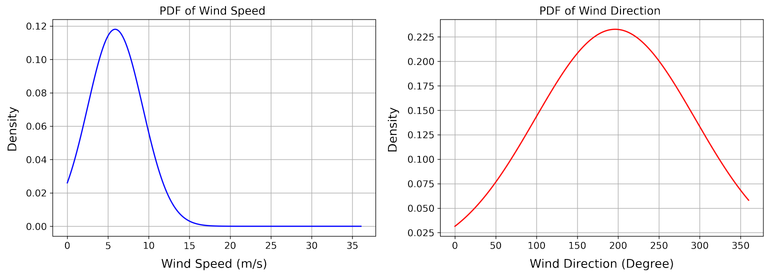


Figure 6: Probability Distribution of Wind Speed.

- The graphs that grouped by mean are the most match to the standard performance;
- The graphs of speed apparent, whether they are generated by onboard or external data, they are the most matching graphs with the general ship's standard performance;
- Graphs of wave height, and current speed apparent are not matching with the standard performance;
- Comparing the wind speed graphs with each other (i.e., without wave or current graphs), it can be observed that:
  - The graphs of wind speed * cos(Φ) are better match to standard performance than true wind speed graphs;
  - The graphs of wind speed * cos(Φ) from the external resources are more matching with the standard performance than the graph of wind speed *cos(Φ) from onboard data.
- To express the weather impact as bounds in fuel and speed dimensions, the calculation of the weather factor (fw) may not be suitable for ships with length smaller than 150 m [8]. However, the generated cruising-mode graphs can be exploited to estimate approximately the weather factor. As it was depicted on figure 7, the estimated weather factor for this vessel is $fw = 0.66$ with cruising-mode speeds;
- As shown in table 1, the correlation of each individual weather variables with fuel in 3 time resolutions, second, minute and hour, it shows that after the vessel's speed over ground, the weather variable of measured wind, especially the wind true and apparent (relative) speed from onboard data has the highest correlation factor, while the wind data from Stormgalss API (sg) has the second highest correlation factor;

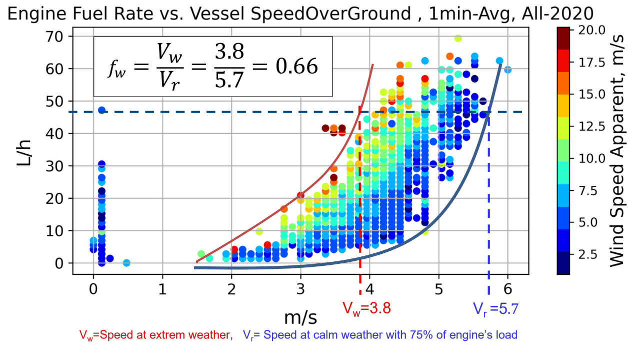


Figure 7: The weather factor obtained from the vessel's standard performance graph.

Table 1: Correlation of vessel's fuel with various weather variables.

Correlation with Engine Fuel Rate	Speed Over Ground (AIS)	Wind Speed (AIS)	Wind Speed (CDS)	Wind Speed (SG)	Wind Direction (AIS)	Wind Direction (CDS)	Wind Direction (SG)	Wind Apparent Speeed (AIS)	Wind Apparent Speeed (CDS)	Wind Apparent Speeed (SG)	Current Speed (SG)	Current Direction (SG)	Wave Height (CDS)	Wave Direction (CDS)	Wave Drag Coefficient (Cds)
Hour	0.6895	0.1680	0.1126	0.1365	0.0339	0.0506	0.0047	0.2200	0.1136	0.1280	0.0505	0.0602	0.1029	0.0260	0.1216
Minute	0.6631	0.1902	0.1172	0.1316	0.0321	0.0457	0.0059	0.2390	0.1714	0.1795	0.0522	0.0347	0.1125	0.0144	0.1256
Second	0.7059	0.1820	0.1241	0.1336	0.0318	0.0451	0.0044	0.2272	0.1773	0.1850	0.0506	0.0338	0.1168	0.0116	0.1314

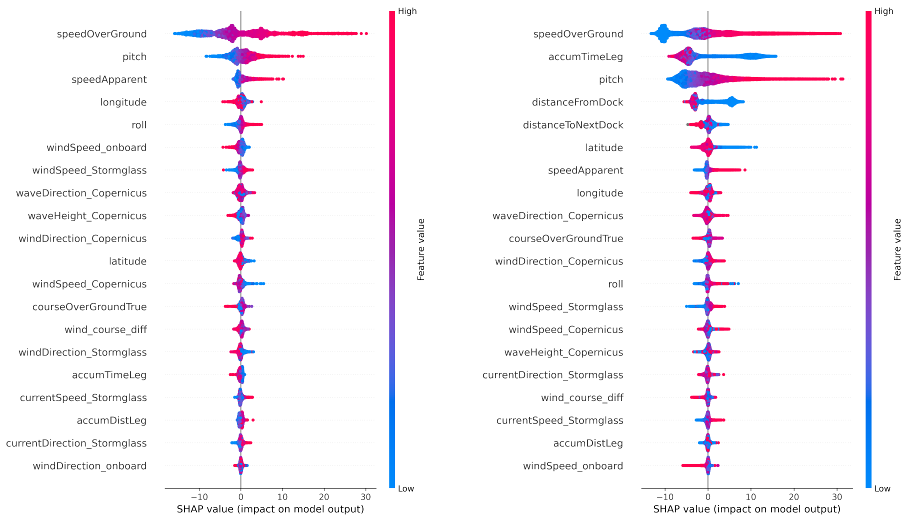
(AIS): Onboard Data, (CDS): Copernicus Weather Data, (SG): StormGlass Weather Data

3.2 The Findings of the Analysis with All Speeds

It can be concluded from the modeling of combined weather variables and investigation their interactions for all vessel's speeds (not only cruising speed) that:

- The comparison of features importance by implementing Shapley values are provided in figure 8, which illustrating the feature importance for the cruising-mode speeds and all-speeds datasets. The Gradient Boosting (XGB) Regression model is implemented as a predictive model for studying the features importance by using Shap package, publicly available in Python [9], to interpret the predictions variance with the value changes in every feature. The features are ranked based on their Shapley values (i.e., importance). It is noticeable that the rank of wind speed from onboard data changes from being the most important feature of weather variables in the cruising speeds to become the less important in all-speeds dataset. The ranks of space features, such as the latitude and longitude, indicating the fuel consumption has some pattern with the vessel routing and its location. In general, the feature importance of weather variables, especially with the case of all speeds, come after features of the vessel's motion in space and time.
  These features are also deployed as inputs of the ANN model with all speeds of the vessel, as shown in table 3;

- At higher apparent speeds (head winds), the fuel rate is also high, as shown in figure 9, which depicts the apparent speed against the difference directions of vessel and true wind. But, there are also high fuel-rate

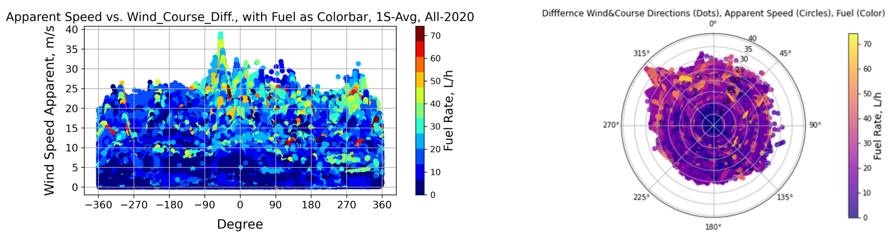


(a) Shapley values for the available features with cruising-mode speeds

(b) Shapley values for the available features with all speeds

Figure 8: Beeswarm plots of Shapley values for the available features

samples at lower apparent speeds (tail winds);



(a) Colorbar Plot of the Speed Apparent vs. Differences of Wind and Vessel's Directions with Fuel Rate

(b) Polar Plot of Differences of Wind and Vessel's Directions with Fuel Rate

Figure 9: Differences of Wind and Vessel's Course with Fuel Rate (Color)

- To further verify the graphs in figure 9, a statical analysis has been conducted as shown in table 2. The difference of wind and course directions are distributed to 8 intervals to signify and investigate the impacts of head winds and tail winds on the fuel consumption. When all vessel's speeds are considered, it is not obvious how the speed apparent individually impacts the fuel consumption. In other words, the wind is not always directly related to the fuel consumption;
- The wave and current variables from external sources, before the interpolation, they came with lower spatio-temporal resolution than the onboard wind data. Therefore, from the resolution perspective, the external weather data are not perfectly capturing the weather impacts, especially, with this analysis on fuel data from vessels for short-sea shipping;

Table 2: Statistics of the graphs in figure 9

Diff_Wind_Course Intervals	Samples Count	Fuel Avg.	Vessel Speed Avg.	Speed Apparent Avg.	Diff_Wind_ Course Avg.	Wind Speed Avg.	Wind Direction Avg.	Vessel. Course Avg.
(0, 90]	8147300	15.42	2.21	7.40	45.06	5.76	219.91	174.86
(90, 180]	7478087	15.23	2.27	4.99	131.87	5.66	253.69	121.83
(180, 270]	2573125	16.70	1.75	4.86	218.78	5.57	286.16	67.38
(270, 360]	940072	18.55	1.86	7.20	298.07	6.05	322.35	24.28
[0, -90)	8452311	15.79	2.22	8.05	-41.39	6.27	188.71	230.10
[-90, -180)	4080962	13.65	1.78	5.03	-135.41	5.45	130.83	266.23
[-180, -270)	3430024	13.69	2.36	5.07	-219.81	5.97	79.45	299.26
[-270, -360)	1479646	15.22	2.58	8.35	-300.98	6.74	35.80	336.78

- The preliminary results of ANN regression models with all-speeds data show that using the wind data from Copernicus leads to the highest accuracy (i.e., $R^2 = 0.8687$), as shown in table 3;

Table 3: Results of ANN regression, where fuel-rate is the ANN output, and ANN input with different cases, for both cruising-mode speeds and all speeds

Index	No.# ANN Inputs	All Speeds			Cruising-Mode Speeds		
		RMSE	R2	MAE	RMSE	R2	MAE
Using onboard wind for weather conditions	11	0.0643	**0.8298**	0.0476	0.0663	**0.8304**	0.0515
Using wind-Copernicus for weather conditions	11	0.0669	**0.8156**	0.0496	0.0679	**0.8218**	0.0539
Using wind-Stormglass for weather condtions	11	0.0664	**0.8188**	0.0490	0.0715	**0.8024**	0.0563
Excluding wind-related variables	7	0.0728	**0.7853**	0.0534	0.0781	**0.7788**	0.0587
All available variables	22	0.0557	**0.8722**	0.0411	0.0557	**0.8799**	0.0417
Excluding weather-related from all available variables	10	0.0732	**0.7829**	0.0532	0.0743	**0.8002**	0.0566
ANN info.: Model=Sequential, Layers=10, #Neurons=100, Activation=relu, Optimizer=adam, Loss=rmse, Epochs=100							

Index	No.# of Inputs	Inputs of ANN
Inputs including wind-related variables	11	speedOverGround, courseOverGroundTrue, windSpeedOverGround, directionTrue, distanceFromDock, accumDistLeg, accumTimeLeg, distanceToNextDock, speedApparent, pitch, wind_course_diff
Inputs excluding wind-related variables	7	speedOverGround, courseOverGroundTrue, distanceFromDock, accumDistLeg, accumTimeLeg, distanceToNextDock, pitch
Inputs of all available variables	22	speedOverGround, courseOverGroundTrue, distanceFromDock, accumDistLeg, accumTimeLeg, distanceToNextDock, speedApparent, pitch, wind_course_diff, roll, latitude, longitude, windSpeed_onboard, windDirection_onboard, currentSpeed_Stormglass, currentDirection_Stormglass, windSpeed_Stormglass, windDirection_Stormglass, windSpeed_Copernicus, windDirection_Copernicus, waveHeight_Copernicus, waveDirection_Copernicus
Excluding weather-related variables	10	speedOverGround, courseOverGroundTrue, distanceFromDock, distanceToNextDock, accumDistLeg, accumTimeLeg, pitch, roll, latitude, longitude

- The preliminary results of ANN regression model with all-speeds data show that the direct and extracted features of onboard wind data can improve the accuracy (i.e., R^2) by range of 5%, as shown in table 3;

- The preliminary results of ANN regression model in table 3 indicates that using inputs that excluding weather-related variables reduce the accuracy by about 9% for R^2 compared to the case of the ANN inputs that including all available variables;

- The accuracy of ANN regression with inputs that excluding weather variables in table 3, indicating that the vessel motion parameters, such as the vessel's speed and course, can influence the fuel consumption significantly, when various vessel's speeds are considered;
- As an alternative of weather forecasts, the historical values of onboard wind data can be used with time-series forecasting models in short-term route planning for short-sea sh.

4 Future Work

The main objective of this analytics is to optimize the ship's speed and course to minimize the fuel consumption. The optimization of fuel consumption problem can be modeled by several techniques [10–12]. Some of those optimization methods have been used with short-sea shipping [13, 14].

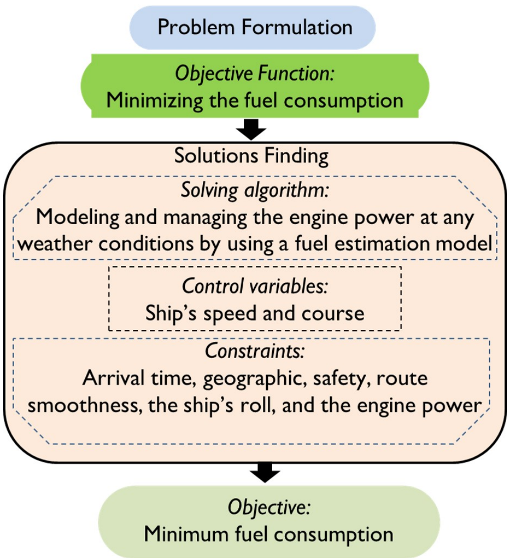


Figure 10: Flow chart for minimization of ship's fuel consumption.

References

[1] "CetaSol AB ." [Online]. Available: https://cetasol.com .

[2] "Marine Traffic." [Online]. Available: https://www.marinetraffic.com/en/ais/details-/ships/shipid:1088282/mmsi:265513810/imo:8602713/vessel:BURO, July 2022.

[3] "Engine Volvo Penta tier13 A2022 8398." [Online]. Available: https://www.volvopenta.com/about-us/news-page/2022/jun/imo-tier-iii-range-expands-with-new-d13-solutions/.

[4] "Copernicus Marine Service ." [Online]. Available: https://marine.copernicus.eu .

[5] "StormGlass API ." [Online]. Available: https://stormglass.io .

[6] J. Carlton, *Marine propellers and propulsion.* Butterworth-Heinemann, 2018.

[7] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[8] Y. Kim, D. Hayden, D. Fathi, G. Hermanski, D. Hudson, P. d. Jong, K. Tanizawa, G. Thomas, and W. Chengshen, "Seakeeping committee final report and recommendations to the 27th ittc," 2014.

[9] "Shap package." [Online]. Available: https://github.com/slundberg/shap.

[10] L. Walther, A. Rizvanolli, M. Wendebourg, and C. Jahn, "Modeling and optimization algorithms in ship weather routing," *International Journal of e-Navigation and Maritime Economy*, vol. 4, pp. 31–45, 2016.

[11] A. Fan, J. Yang, L. Yang, D. Wu, and N. Vladimir, "A review of ship fuel consumption models," *Ocean Engineering*, vol. 264, p. 112405, 2022.

[12] M. H. Moradi, M. Brutsche, M. Wenig, U. Wagner, and T. Koch, "Marine route optimization using reinforcement learning approach to reduce fuel consumption and consequently minimize co2 emissions," *Ocean Engineering*, vol. 259, p. 111882, 2022.

[13] A. Zakaria, A. Md Arof, and A. Khabir, "Instruments utilized in short sea shipping research: A review," *Design in Maritime Engineering*, pp. 83–108, 2022.

[14] M. Grifoll, F. X. Martínez de Osés, and M. Castells, "Potential economic benefits of using a weather ship routing system at short sea shipping," *WMU Journal of Maritime Affairs*, vol. 17, no. 2, pp. 195–211, 2018.