Norwegian School of Economics
Bergen, Spring 2022

Estimating the Impact of Weather-Sensitive Cargo Risk on Transport Cost

Rikard Knive & Jon Mæle Liane

Supervisor: Roar Os Ådland

Master thesis, Economics and Business Administration

Major: Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis is written as a part of our master's degree at the Norwegian School of Economics (NHH) with a specialization in Business Analytics. We want to extend our sincere gratitude to the parties who assisted us during our thesis. First and foremost, we would like to thank our supervisor, Roar Os Ådland, for his pivotal guidance, feedback, and advice throughout the process. His extensive knowledge in the field of maritime economics has been immensely valuable to us. Additionally, he has introduced us to his broad network within the shipping industry and connected us with G2 Ocean. In this regard, we would like to express our profound appreciation to the people at G2 Ocean: Phil Curran, Trond Aga Haug, Arnt Davanger, and Henning Rebnord for supplying us with the necessary data, as well as providing unique insights into their operations.

Norwegian School of Economics

Bergen, June 2022

Jon Mæle Liane

Rikard Knive

Abstract

As a consequence of the global renewable energy transition, there is a rising demand for transportation of project cargo, such as wind turbine components. Transportation of this type of cargo requires special considerations as it is sensitive to adverse weather exposure. This thesis aims to determine what impact weather-sensitive cargo has on transportation cost, formulated as the expected incremental cost compared to vessels transporting ”regular“ cargo. The chosen methodology approach applies a ship weather routing model to identify the most cost-efficient route from Spain to Houston while accounting for the required weather considerations. The weather routing model comprises one of the literature's most prominent pathfinding algorithms combined with complex machine learning models to achieve realistic cost estimations. Our findings indicate that vessels carrying weather-sensitive deck cargo have a high tendency to deviate from the optimal route selected by vessels carrying regular cargo. This is particularly evident in the winter months, where our findings identify an incremental cost upwards of 13.80%. Conversely, the results reveal an upper limit on the incremental cost of 0.70% in the summer months, indicating a relatively modest disparity from the vessels transporting regular cargo. This asymmetry is found to be largely explained by the seasonal effect of adverse weather. Our findings suggest that vessels transporting weather-sensitive deck cargo are at a considerably higher risk during the winter months, where exposure to adverse weather effects is especially prominent.

Keywords – Weather-Sensitive Cargo Risk, Ship Weather Routing, Dijkstra's Algorithm, Optimization, Machine Learning, Shipping

Contents

List of Figures

List of Tables

1 Introduction

Maritime transport is the fundamental backbone of global trade. According to UNCTAD's annual review of maritime transport, goods carried by sea make up more than 80% of global trade (UNCTAD, 2021). UNCTAD also projects that international maritime trade will continue to grow in the coming years. An important contributor to this growth is the accelerating focus on renewable energy sources that increases new infrastructure investments (Project Cargo Journal, 2021). This generates increased demand for transportation of project cargo, such as wind turbine components, which is defined as weather-sensitive deck cargo (Liu, 2021).

The joint-venture company, G2 Ocean, is eager to exploit this growing trend and is therefore wagering on the project cargo segment (Project Cargo Journal, 2019). G2 Ocean is comprised of two of the world's largest open hatch shipowning companies: Gearbulk and Grieg Maritime Group (G2 Ocean, 2022a). The company currently holds a strong position in the onshore wind market and aspires to further strengthen its position in the growing offshore wind segment and oil and gas market. Both market segments require transportation of project cargo sensitive to adverse weather conditions. However, transportation of such cargo is likely to have an increased risk of additional costs as special precautions are required as a consequence of its sensitivity to external forces.

To avoid exposure to adverse weather, vessels with weather-sensitive deck cargo may have to deviate from the otherwise optimal route for vessels with regular cargo, thus increasing the costs of transportation. The cost of transporting shipments is largely influenced by the consumption of bunker fuel. According to Stopford (2009), two-thirds of a vessel's voyage costs are attributed to the fuel consumption. Weather-sensitive deck cargo has two additional cost drivers: increased sailing distance and weather exposure. In addition, an added day rate is charged if there occurs a delay in the estimated time of arrival (ETA). Furthermore, as pricing of forward cargo is usually determined weeks in advance, the chartering managers rely on uncertain weather forecasts during pricing and contract negotiations. This makes pricing of weather-sensitive cargo especially challenging. Our study will aid in this process by identifying the expected risk of increased costs, and thereby enable accurate pricing of voyages with such cargo.

For this purpose, an optimal ship weather routing model is applied to determine how the risk of adverse weather conditions affects the routing choices of vessels with and without weather-sensitive deck cargo. This methodology approach utilizes Dijkstra's algorithm, a pathfinding algorithm suitable for identifying the cost-minimizing route while accounting for the impact of environmental factors. In addition, machine learning (ML) regression algorithms are implemented to predict the vessel's fuel consumption, which serves as the main component of the ship weather routing model's cost functions. The model selection includes K-Nearest Neighbour (KNN), Random Forest (RF), Extra Trees (ET), as well as an Artificial Neural Network (ANN) to identify the empirical model, which provides a sufficient approximation of the effect of external weather conditions on fuel consumption.

There is a vast number of studies on ship weather routing. This field has seen escalating interest in recent years as increased computation power has enabled resolving increasingly complex optimization problems (Zis et al., 2020). In addition, regulatory pressure to reduce emissions, coupled with volatile bunkers costs, have provided significant economic incentives to identify the most fuel- and cost-efficient routes. Our study distinguishes itself from the relevant literature by providing a unique perspective on the weather-sensitive project cargo segment. The findings of this study supplement both the literature and industry by stipulating optimal routes and cost estimates for cargo restricted by weather constraints.

The contributions of our thesis are 4-fold: Firstly, we estimate the incremental cost associated with the risk of transporting weather-sensitive cargo, which prior to this has been determined based on previous experience. Secondly, we demonstrate how historic third-party meteorological data can be incorporated to identify the relation between external weather effects and fuel consumption on weather-sensitive deck cargo. Thirdly, we provide an operative tool that can aid in estimating future freight rates for weather-sensitive deck cargo based on weather forecasts. Lastly, our research contributes to establishing a benchmark for optimal modelling of vessel fuel consumption predictions, which helps facilitate further research within the maritime shipping sector.

The structure of this thesis is as follows: section 2 will elaborate on the background and scope of this thesis. Section 3 will cover existing literature on this topic and provide a theoretical framework for our optimization and machine learning models. Section 4 and

5 will explain our data and methodology approach, respectively. In section 6, we will provide our results and discussion of our findings. Section 7 will summarize the limitations of our study and propose how our work can be further developed. Lastly, in section 8, we will present our conclusion.

2 Background

Project cargo, such as wind turbine components, is sensitive to adverse weather effects, thus defined in this thesis as weather-sensitive deck cargo. This cargo is oversized, heavy and valuable components stored on the open deck of the vessel (G2 Ocean, 2022b). It requires tailor-made constructions and lashings to be properly secured on deck, which exposes the cargo to external forces such as adverse weather conditions. Due to this, weather-sensitive cargo requires special considerations when transported, such as applying weather constraints. This study restricts the vessels from sailing within an area with *any* mean hourly wave height exceeding 6 meters on the day of passing. Note that this is defined by G2 Ocean as a general directive when transporting weather-sensitive deck cargo, therefore it might not be applicable to other operators. In reality, the cargo limitation is often determined on a case-by-case basis dependent on the ship, cargo, and lashings.

As a part of our cooperation with G2 Ocean, we wanted to explore routes relevant to their operations. We have therefore opted to limit the scope of this thesis to cover trade from Spain to Houston (Project Cargo Journal, 2019). Furthermore, there are predominantly two vessel designs of G2 Ocean's vessels designated for project cargo: Class I and Class L. Data from these classes will serve as the foundation of our analysis. By looking at historic routes from these two classes, we have confined the relevant area to be within the northern part of the Atlantic[1]. We will examine the weather patterns found within this area to better grasp the importance of route selection for vessels carrying weather-sensitive cargo. All historical extreme weather observations[2] from 2006 to 2021 have been analyzed and mapped to determine which areas are more exposed to adverse weather conditions. This dispersion is seen in Figure 2.1.

[1]The Gulf of Mexico is excluded due to limited routing options and an observed low frequency of extreme weather conditions.

[2]Extreme weather observations is defined as wave heights equal or exceeding 6 meters

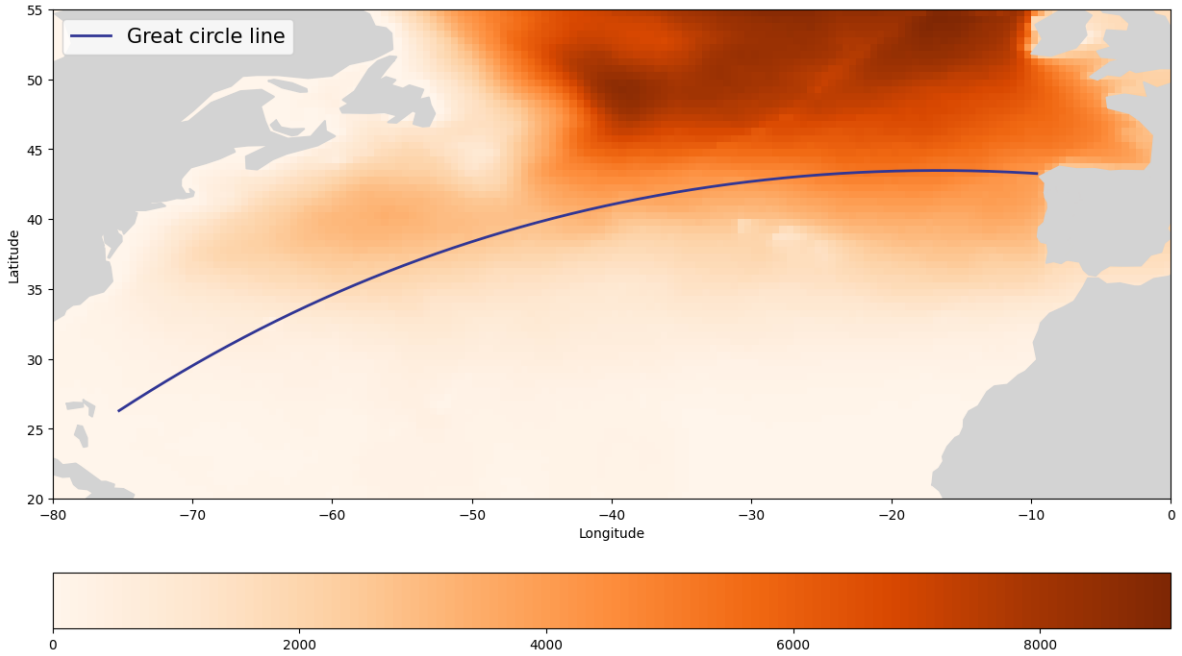

Figure 2.1: Mapping of the total number of hourly extreme weather observations (wave height >= 6m) found in the period from 2006 to 2021. The total number of observations ranges from 0 to 9000. The line represents the route following the great circle line from Spain to the Bahamas

It is apparent that the northern part of the area is more exposed to rougher weather conditions. This includes the area where the shortest path crosses and suggests that vessels transporting weather-sensitive deck cargo can potentially deviate from the shortest path due to its weather constraints. The shortest path is defined in terms of the shortest distance along the great circle line (Jie and Miao, 2021). Due to the concentration of adverse weather conditions along the shortest path, the shortest path might be undesirable for any vessel regardless of cargo. To create an accurate comparison, we will compare the cost-minimizing route of vessels carrying weather-sensitive cargo with the typically considered optimal route, which is the cost-minimizing route of vessels without weather-sensitive cargo.

Furthermore, we would like to examine the seasonal environmental conditions along the route. The Atlantic is known for seasonal weather incidents, such as the hurricane season, yielding periods of extreme weather conditions (Kossin, 2008). We will examine the seasonal weather patterns' effect on routing choices for weather-sensitive deck cargo and, thereby, the incremental costs of sailing suboptimal routes. The following plot, displayed in Figure 2.2, strongly indicates a substantial variation in extreme weather conditions

throughout the year; hence the optimal route selection for sensitive deck cargo might vary heavily depending on the time of year.

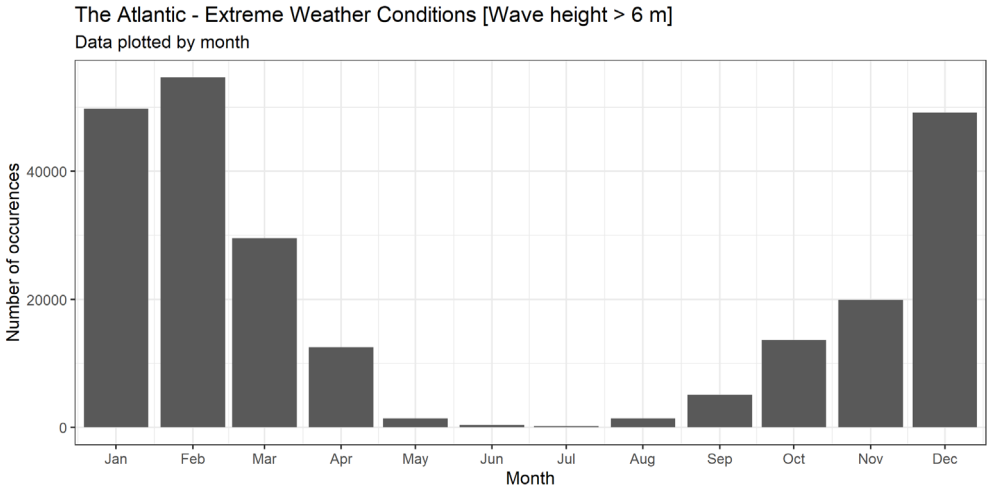

Figure 2.2: The total accumulated number of wave observations equal or exceeding 6 meters found per month within the relevant area from 2006 to 2021

Figure 2.3 displays the median wind speed and wave height for 2020, representing the median weather conditions observed from 2006 to 2021. This indicates an increased weather effect during the winter season, while the summer season experiences better weather conditions. We, therefore, expect a definite seasonal weather effect along the route.

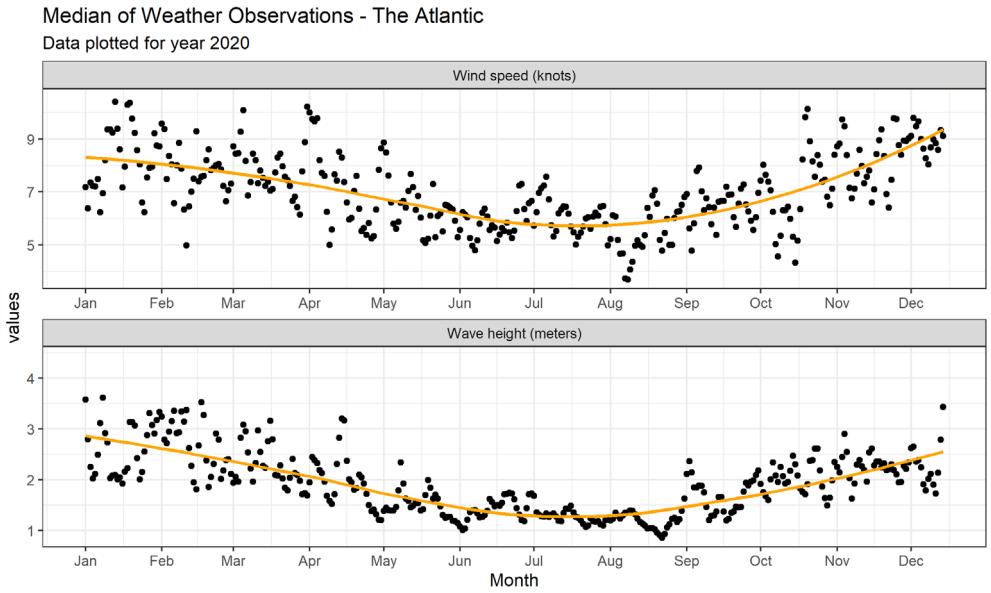

Figure 2.3: The median of monthly wind speed (knots) and wave height (meters) observations found in 2020

3 Theory and Literature review

This section provides an overview of scientific literature relevant to this thesis. First, a synopsis of various methodologies to optimize route selection in maritime transportation is examined. Then, the fundamental theory of vessel bunker consumption is presented. Lastly, we will introduce the concept of machine learning algorithms, one of the leading data-driven techniques to provide accurate bunker consumption estimations.

3.1 Ship weather routing

The thesis's methodology approach is heavily inspired by pre-existing literature on optimal ship weather routing. Ship weather routing has seen increasing attention in recent years in both academia and industry (Zis et al., 2020). The literature mentions a vast array of methodologies used to optimize route selection for a given voyage while considering the environmental effects of wind and waves. The most prominent methods related to weather routing optimization in maritime transportation include the Modified Isochrone method, dynamic programming, and pathfinding algorithms. These routing algorithms all share two fundamental components: waypoint grid generation and path selection and evaluation criteria (Wang et al., 2017).

The isochrones method was originally proposed by James (1957) and further extended and modified by Hagiwara and Spaans (1987) into the modified isochrone method. An isochrone represents a line in the waypoint grid system, which displays the different possible trajectories of a ship for a given route (Zis et al., 2020). Each line is connected by stages, where each stage is assumed to be sailed at an equivalent time. This assumes that the ship operates with a constant engine power during the voyage. The modification of Hagiwara and Spaans (1987) showed that when environmental factors are incorporated, the length of each isochrone will change, and thus different lengths can be traversed within the same time interval. The inclusion of stages reduces the complexity by splitting the problem into multiple sub-route optimization problems. To solve the optimization problem at hand, this method utilized a forward recursive algorithm for individual sub-routes (Wang et al., 2017). The forward recursion allows the ship to vary its heading for each interim stage. Furthermore, a three-dimensional isochrone method (3DMI) was proposed

by Lin et al. (2013), which enables varying ship speed and wave angle as well. This makes this method suitable for identifying the optimal route in terms of the expected time of arrival and fuel consumption (Zis et al., 2020).

Dynamic programming is based on Bellman's principle of optimality (Bellman, 1954): "An optimal policy has the property that whatever the initial state and initial decision is, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". This method constructs a grid system based on the great circle reference path (Wang et al., 2017). Along this path, the route is divided into multiple stages, and for each stage, multiple grids are generated perpendicular to the reference path. A two-dimensional dynamic programming (2DDP) variation has been utilized by Chen (1978) and Wang and Chrétienne (1993) to minimize expected voyage cost and fuel consumption, respectively. The dimensions refer to the number of optimization criteria the algorithm is capable of considering. Shao et al. (2012) proposed a new forward three-dimensional dynamic programming (3DDP) method. This enabled fuel consumption minimization while also considering safety constraints for safe ship operations.

There are mainly two pathfinding algorithms that have been applied in ship weather routing, namely A* and Dijkstra. These algorithms divide all defined paths from point A to point B into a series of sub-routes and sub-points (Wang et al., 2017). Each sub-point is connected to its neighboring sub-points by a path, which has an associated cost of travel. The main principle of these algorithms is that the optimal sub-route follows the path with the lowest associated cost. The path with the lowest cost will be used as a reference when comparing the corresponding paths. This means that when a new point is reached, the new path will be compared to the cost of the reference path. If the associated cost of the new path is smaller than the reference, it will be assigned as the new reference. This procedure will continue until the destination point is reached and only the optimum route remains.

Multiple variations of the pathfinding algorithms have been implemented for ship weather routing. Padhy et al. (2008) studied the North Indian Ocean using Dijkstra's algorithm, while Szlapczynska (2015) implemented A* in her multi-objective weather routing algorithm. These methods have proved suitable for fuel cost minimization with a focus on studying the impact of environmental uncertainties on a ship's optimization planning

(Wang et al., 2017). Mannarini et al. (2013) for instance, used a modified version of Dijkstra's algorithm in combination with meteorological and oceanographic data for environmental factors to identify the optimal route.

3.2 Vessel fuel theory

According to Zis et al. (2014), the fuel consumption of any voyage can be isolated into four separate phases: sailing, anchorage, maneuvering, and at-berth. The fuel consumption during sailing is predominantly determined by the effective horsepower, P_E, produced by the main engine of the ship (Meng et al., 2016). P_E expresses the power required by the engine to propel the ship forward through the water given the total resistance R_T and sailing speed V, in accordance with the formula:

$$P_E = R_T \times V \tag{3.1}$$

The total resistance is further comprised of three components:

$$R_T = R_F + R_R + R_A \tag{3.2}$$

R_F is the frictional resistance representing the induced resistance of the hull and the propeller due to friction and wave-making (Meng et al., 2016). This is affected by the loading condition of the vessel and the total deterioration of the hull. The R_R is the residual resistance of waves, which causes pitch and heave motions of the vessel, while R_A is the air resistance caused by the wind. Arribas (2007) finds that the residual resistance can be approximated by a function of the significant wave height and the modal period of the sea. The air resistance can be explained by features such as wind speed and relative wind direction (Magnussen, 2017).

As evident from equation 3.1, the vessel speed is a highly influential predictor of bunkers consumption, supported by previous literature by Adland et al. (2020); Gkerekos et al. (2019); Meng et al. (2016). The consensus is that the total resistance is roughly proportional to the square of vessel speed, V^2, the bunkers consumption can therefore be approximated by a cubic function of speed (Meng et al., 2016). However, this speed-consumption

relationship is challenged by Adland et al. (2020). They argue that while the "cubic law" seems to hold near the design speed of vessels, it is not necessarily applicable for all speed ranges. Due to this, Adland et al. (2020) emphasizes the importance of accurate data-driven estimates of the speed-consumption relationship for optimization of ship operations.

Aside from the vessel speed, the vessel's bunker consumption is known to be influenced by various features, such as its draft, hull fouling condition, wave height and direction, wind force and direction (Adland et al., 2018). Water salinity, currents, and temperature are found to have a slight effect on consumption as well, however, the influence of wave and wind is considerably more significant (Carlton, 2018).

3.3 Artificial intelligence and machine learning

The use of artificial intelligence and machine learning methodologies have garnered much attention in recent years (Zis et al., 2020). Most papers in the literature attempt to utilize this methodology to predict fuel consumption while accounting for various environmental factors to minimize either fuel consumption or voyage arrival time. Various data-driven machine learning approaches have been implemented in an effort to model the complex and often non-linear relationships associated with vessel fuel consumption. This approach has been proven effective, as many authors are reporting substantial fuel, cost, or emission savings (Wang et al., 2016; Zheng et al., 2019; Du et al., 2019).

Machine learning can be defined as the use of formal structures (machines) to do inference (learning)(Clarke et al., 2009, p. 2). This means that the machine learning algorithms improve their predictive performance by learning from a set of training data. The machine learning algorithms can generally be classified into two groups: supervised and unsupervised problems (Gkerekos et al., 2019). The latter refers to machine learning problems where the algorithm is left to its own to discover hidden patterns and structures in its input data without any labels are provided to the learning algorithm (Bishop and Nasrabadi, 2006, p. 3). Regarding supervised learning, the algorithm's goal is to learn a general rule that maps the input data to a target value (output) (Gkerekos et al., 2019). For our thesis, we will utilize the supervised machine learning approach to estimate a vessel's fuel consumption (target value), which is dependent on various parameters (inputs)

affecting the total resistance in a regression problem.

3.4 Performance metrics

Performance metrics are necessary to evaluate the accuracy of prediction models. There is a wide variety of metrics for this purpose; however, none can be universally used for all instances. Each metric has its inherent weaknesses and strengths depending on the nature of the problem. Thus, an educated selection of metrics can be beneficial to measure the predictive power of each prediction model objectively. Some of the most common choices are the Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

A key performance indicator of model accuracy for regression analysis is the R^2 (Abebe et al., 2020). It measures the proportion of the variance of the response variable that can be attributed to the model's features. This is achieved by dividing the *sum of squares of residuals* and *the sum of squares total*. The former quantifies the total deviance of predicted values ($\hat{y_i}$) to observed data (y_i), while the latter identifies the total difference between observed data and its mean ($\overline{y_i}$). The resulting quotient is subtracted from 1 to produce the coefficient of determination. R^2 scales from 0 to 1; 0 indicates an inability to explain the variance, and 1 indicates a perfect prediction. The definition is presented in Equation 3.3:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y_i})^2} \tag{3.3}$$

MAE is a relatively simple yet effective measure of error, which identifies the expected magnitude of error of the model (Willmott and Matsuura, 2005). It involves calculating the *sum of absolute difference* between the predicted value ($\hat{y_i}$) and the true observed value (y_i), then dividing by the number of samples in y (n). A low value of MAE is desired as it corresponds to a low expected prediction error. The formula is displayed in Equation 3.4.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}| \tag{3.4}$$

RMSE is the standard deviation of the residuals, which can be calculated through a

sequence of three steps. First, the *sum of squared error* is obtained as the sum of the individual squared difference between predicted value ($\hat{y}_i$) and observed value (y_i) (Willmott and Matsuura, 2005). Then, the squared error sum is divided by n, and lastly, the square root of the resulting equation is computed, yielding the RMSE. Similar to MAE, a lower value of RMSE denotes a better prediction accuracy (Abebe et al., 2020). The definition of RMSE can be seen in Equation 3.5.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.5}$$

Though widely used in climatic and environmental literature, Willmott and Matsuura (2005) argues that RMSE is an inappropriate measure of error. This is due to its function of three error characteristics; making is misleading and not as easily interpretable. One unfortunate trait of RMSE is that each error influences the total error in proportion to its square. This results in larger error margins having a relatively greater influence on the total error than the more minor ones. Contrary to RMSE, MAE provides an unambiguous measure of average error and is less biased for increasingly higher error magnitudes. On the other hand, MAE becomes a less intuitive measure of a model's performance when dealing with large error values.

The Mean Absolute Percentage Error (MAPE) is a variant of MAE expressed in percent form. MAPE has the advantages of scale-independence as well as providing excellent interpretability (Kim and Kim, 2016). However, a substantial drawback of MAPE is that it produces infinite or undefined values for zero or values approaching zero (Kim and Kim, 2016). The error measure also wrongly favors underestimated predictions, as there exists a ceiling of 100% error for underestimated predictions, whereas there exists no ceiling for overestimation (Gkerekos et al., 2019). To resolve this, the measure can be transformed into a symmetric measure, seen in equation 3.6.

$$sMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| - |\hat{y}_i|)/2} \tag{3.6}$$

The symmetric Mean Absolute Percentage Error (sMAPE) provides an improved alternative to MAPE as it corrects the drawbacks by defining a lower and upper bound.

4 Data

The data section will provide details of data acquisition, pre-processing methodology and the resulting feature selection for our dataset. Our approach for data processing is largely comparable to the data process workflow used by Abebe et al. (2020), depicted in Figure 4.1.

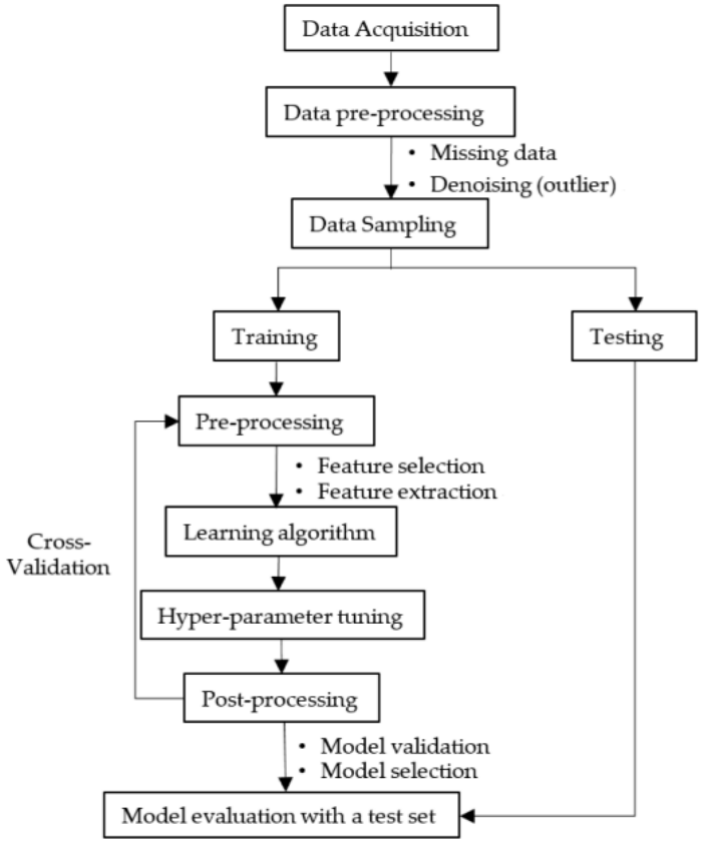

Figure 4.1: Data and methodology workflow (Abebe et al., 2020)

The data acquired is processed and handled in R, Python, and Go on an Apple M1 chip with 16 GB RAM and 8-core CPU. The datasets acquired are of significant size, which poses an immense challenge when handling and processing due to our computer equipment's limited processing and storage capabilities. These issues are managed by separating our operations into multiple steps, which will be further clarified in the methodology section.

4.1 Noon reports

This thesis will mainly rely on noon reports as the primary data source. A *noon report* is a low-resolution dataset prepared by the ship's chief engineer containing relevant standardized information (Anish, 2021). The fields reported include, but are not limited, to the ship's position, fuel consumption, average speed, and average experienced environmental conditions. The reports are prepared approximately once every 24 hours and are then sent to the respective company around noon, hence the name *noon report*. The noon report datasets are coarse but still widely used across the global fleet due to their convenience and cheap compilation (Aldous et al., 2015). An alternative data source is using continuous monitoring (CM) systems that utilize automated onboard sensors. These sensors can provide more nuanced and accurate data with a significantly increased sampling frequency. However, due to its lack of historical use, the CM methods cannot provide an adequate sample size for the relevant ships of this thesis.

However, noon report datasets should be used with caution, as the reports are prone to be erroneous due to their reliance on human input, which makes the datasets inherently uncertain (Aldous et al., 2015). As stated earlier, the chief engineer prepares the noon reports by hand. This poses multiple problems, as it exposes the reports to potential human errors such as misinterpreted readings and input errors. In addition, the chief engineers might not maintain a consistency with units, rounding, report intervals or the number of fields reported. Lastly, there is also a risk for equipment malfunction, which can yield inaccurate measurements.

Aldous et al. (2013) have analyzed and quantified the intrinsic uncertainty of noon reports as a data source. Their approach was to utilize a regression model in combination with pre-existing knowledge of the physical influence of fuel consumption, which enabled them to capture the physical interaction between parameters. They fitted a multi-linear regression model consisting of variables closely linked to a ship's fuel consumption in a physical manner while making sure that the remaining residuals were normally distributed. This produced a model which was a close approximation of the true underlying model. They found that the relative standard error is in the range of 1-8% for various types of oil tankers and 15.8% for LNG carriers, which is thought to be due to aleatory and measurement uncertainty in the noon reports. This proved the viability of the underlying

data, making it credible for use in performance analysis. However, before the noon report datasets can be used in our models, it is crucial to perform extensive pre-processing to address the inherent uncertainty.

Through our collaboration with G2 Ocean, we have received access to 54 882 noon reports, covering the period from 2009 to 2021. Since G2 Ocean was launched in 2017, any noon report prior to this is from Grieg Star ships. These reports consist of information from their Class I and Class L vessels exclusively, as these are commonly used to transport weather-sensitive deck cargo. A complete list of all the vessels within these classes can be seen in Table 4.1.

Table 4.1: G2 Ocean's vessels

Vessel	IMO	Vesselcode	DWT	Year of build
Class I				
Star Istind	9182954	I001	46,428	1999
Star Ismene	9182966	I002	46,428	2000
Star Isfjord	9182978	I003	45,740	2000
Class L				
Star Laguna	9593854	LAG | L001	50,761	2012
Star Lima	9593866	LIM | L002	50,761	2012
Star Lindesnes	9593878	LIN | L003	50,761	2012
Star Louisiana	9593880	LOU | L004	50,748	2013
Star Lofoten	9593892	LOF | L005	50,761	2013
Star Livorno	9593907	LIV | L006	50,761	2013
Star Loen	9603790	LOE | L007	50,761	2013
Star Luster	9603805	LUS | L008	50,761	2013
Star Lygra	9616838	LYG | L009	50,761	2013
Star Lysefjord	9616840	LYS | L010	50,728	2014

The reports include all the historic voyages of the I-class and L-class vessels, displayed in Figure 4.2, not only the voyages within the Atlantic area that we focus on in this thesis. We opt to use the dataset in its entirety to secure a sufficient sample size. This assumes that the ship's speed and fuel consumption are not affected by the ship's position but rather the conditions at the given location. Therefore, we emphasize that our models will not include time or location, just the vessel-performance data.

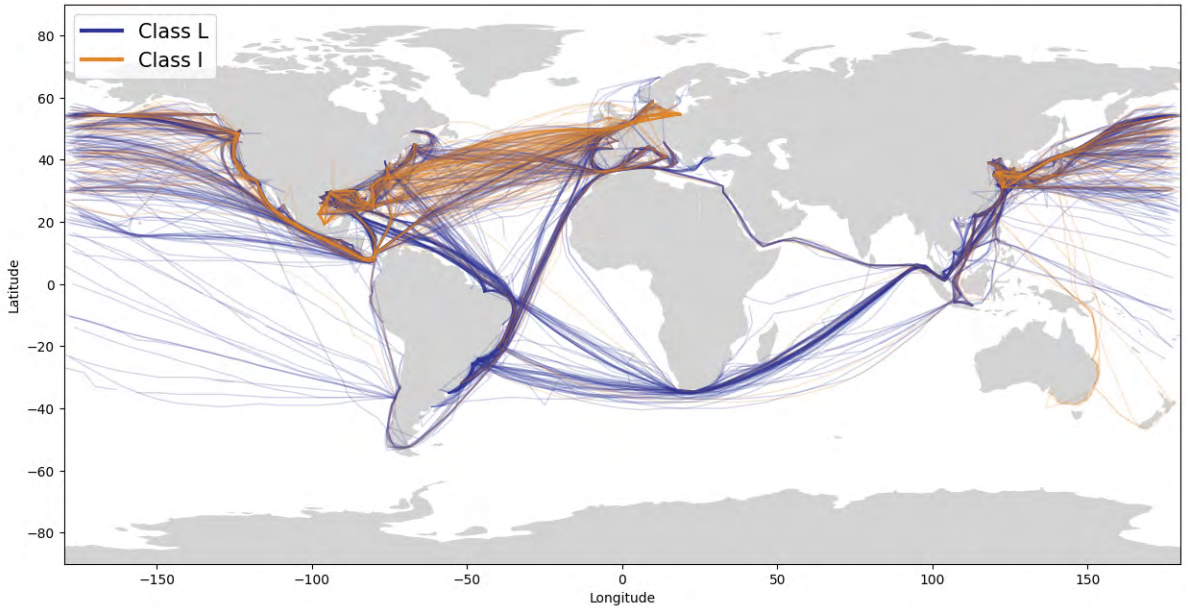

Figure 4.2: All historic voyages found in noon reports from G2 Ocean's I-class and L-class vessels between 2009-2021

4.2 Third-party weather data

The extent to which our methods and models can be applied and provide accurate results depends on the data quality. Chu et al. (2015) found that meteorology and oceanography (METOC) systems can contribute to upwards of 20% in fuel savings. More accurate weather data enables increased model performance and consequently a significant increase in cost savings. The noon reports contain information regarding environmental conditions such as relative wind direction, average wind force, relative sea direction, and average sea state. However, as previously mentioned, the noon reports are fundamentally uncertain either due to human error or aleatory and measurement uncertainty.

To ensure high-quality weather data, we opt to use third-party meteorological data rather than the weather recordings found in the noon reports. We have retrieved historical weather data from the European service Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Hersbach et al., 2018), which provides weather data using satellite altimetry data, numerical models, and in situ data. It contains information regarding waves, wind, currents, precipitation, salinity, and temperature. This thesis will mainly focus on wave and wind features because we assume that these contain the highest explanatory impact on fuel consumption in a physical manner. We acknowledge that

the remaining features might have some predictive power; however, the influence of wave and wind is considerably more significant (Carlton, 2018). This is further supported by Abebe et al. (2020) and Adland et al. (2020) that find that sea current speed and current direction have a relatively low impact on fuel consumption.

We will be using the ERA5 hourly data on single levels dataset containing data from 1979 to the present. This dataset contains the wind speed components *u10* and *v10*, which report the wind speed measurements from the eastward and northward directions, respectively. The dataset also includes the wave measurements for the mean wave period (*mwp*), mean wave direction (*mwd*), as well as the significant height of combined wind waves and swell (*swh*). Note that the wave data have a spatial resolution of 0.5 degrees compared to the wind data's 0.25-degree resolution. Further details of the third-party weather data can be found in table 4.2.

Table 4.2: Third-party weather data (Hersbach et al., 2018)

Source	Storage size	Temporal resolution	Spatial resolution	Features
ERA5 hourly data on single levels from 1979 to present (CDS)	222 GB	1 hour	0.5°	Mean wave period (mwp)
				Mean wave direction (mwd)
				Combined height of wind waves and swell (swh)
	582 GB	1 hour	0.25°	Eastward component of wind (u10)
				Northward component of wind (v10)

4.3 Pre-processing

4.3.1 Merging datasets

Before performing an extensive pre-processing procedure, the noon reports and third-party weather data are merged into a single dataset. To enable matching with data from the noon reports, we interpolate the coordinates in the noon reports to the closest latitude and longitude value to a grid of 0.5 degrees resolution. The interpolation will yield a maximum deviation of 27 kilometers from the actual observation, which we deem acceptable. The wind and wave features are then matched with the noon report dataset along three dimensions: *longitude*, *latitude*, and *time (UTC)*. This ensures that the correct weather condition is matched with the corresponding location of the vessel at the appropriate time.

However, due to missing values for the required dimensions, the sample size is reduced from 54 882 to 35 182. Finally, the dataset is split by class ID, creating a separate dataset for Class I and Class L vessels. Table 4.3 displays descriptive statistics of the numeric features comprised in the combined dataset.

Table 4.3: Descriptive statistics of combined noon reports and external weather

n = 35 182								
Feature	n_missing	mean	sd	p0	p25	p50	p75	p100
longitude	0	-13.76	95.58	-180.00	-85.77	-39.02	79.47	180.00
latitude	0	24.59	23.16	-55.10	13.82	30.95	40.00	66.30
fuel_consumption	2611	16.75	13.00	0.00	0.00	19.30	27.80	150.10
speed_made_good	4386	13.18	31.24	0.00	11.96	13.33	14.44	4923.00
draft_aft	4517	9.72	1.77	0.00	8.20	9.65	11.25	40.60
draft_forward	4517	9.29	13.36	0.00	7.50	9.19	10.85	1096.00
u10	189	-0.32	4.97	-19.84	-3.89	-0.46	2.86	21.50
v10	189	-0.02	4.73	-22.69	-3.06	0.02	3.01	20.24
wave_deg	5445	176.82	98.96	0.03	90.64	180.66	258.93	359.99
wave_period	5445	7.27	2.32	1.67	5.52	7.32	8.90	17.69
wave_height	5445	1.74	1.09	0.03	0.95	1.57	2.28	10.59
h10	189	6.03	3.30	0.03	3.49	5.71	8.12	30.73

4.3.2 Data pre-processing

The scope of this thesis is limited to estimating the fuel consumption during operating periods of the ship, therefore the feature *fo_at_sea* indicating fuel consumption from the main engine, is defined as the dependent variable. This means that mooring and anchoring periods, as well as the use of auxiliary engines to generate electricity for the vessel are discarded. The inclusion of the auxiliary engines could affect the accuracy of our models as their consumption is dependent on different features compared to the main engine.

The speed of the vessel can decrease during a voyage due to increased weather exposure, alternatively, it might be because of maneuvering near ports and shores or navigation in constrained waters (Adland et al., 2020). The low temporal resolution of noon reports cannot accurately differentiate between these scenarios, but it would still impact the fuel consumption. To address this, we adopt a similar approach as Adland et al. (2020), by excluding data with less than 5 knots of *speed_made_good*. This limits our scope to fuel consumption estimation during open sea sailing.

An examination of the residuals and a descriptive analysis following an initial visual inspection of the raw dataset find that the noon reports contain outliers for various features. These can result from poorly calibrated or malfunctioning sensors or human input errors. To reduce the overall noise in our data, a default outlier value is set for fo_at_sea, speed_made_good, and draft based on the maximum values declared in the vessel's technical data plus an additional safety margin (Clarksons Research, 2022). This gives an upper limit of fo_at_sea, speed_made_good and draft of 60, 20 and 14, respectively

Feature selection methods are important to filter out unnecessary and redundant features (Bommert et al., 2020). First, we remove all endogenous determined features, meaning features heavily correlated to the dependent variable, such as RPM and distance. Furthermore, to identify and discard of features with lower predictive power, a high-correlation and near-zero variance filters are applied. The high correlation filter calculates the correlation coefficient between the input features, discarding input features with a coefficient higher than 0.8. Input features with an exceedingly low variance are removed as they represent noise and provide no support in differentiating groups of data (Bommert et al., 2020). Finally, every row containing a missing value is dismissed as they would disrupt the model performance. Because of the aforementioned selection and filter methods, the dataset was further reduced to a total of 22 459 observations.

4.3.3 Feature engineering

The wind variables *u10* and *v10* are converted from their respective vector forms to the magnitude of wind speed and direction angle. To calculate the magnitude and corresponding angle, we use the following equations:

$$|V| = \sqrt{u10^2 + v10^2} \tag{4.1}$$

$$\theta = 180 + \frac{180}{\pi} \tan^{-1}(v10, u10) \tag{4.2}$$

The wind and wave features contain separate features for their respective magnitude and direction. To interpret the effect of the forces on the vessel, an interaction term

that describe both the size and direction relative to the vessel should be included. This requires knowledge of the vessel's bearing, which the noon reports do not include. To resolve this, we approximate the vessel's bearing by calculating the true course along the rhumb line between every connecting points (Karney, 2013). With the inclusion of the vessel's bearing, the wind and wave directional features can be transformed based on the direction of the vessel. This transformation is inspired by Nilsson and Nilsson (2021), who assign directional features moving in the same direction to 90, while directional features moving in the opposite direction are assigned a value of -90. This is visualized in 4.3, where the outer circle represents the difference in angle between the directional weather features and the vessel's bearing, and the inner circle represents the transformed relative direction. Subsequently, the magnitude of wind speed and wave height are multiplied by their corresponding relative directional feature. This approach accomplishes to distinguish wind and waves coming from astern or from the bow of the vessel by affecting the fuel consumption in opposite directions. Based on the findings of Adland et al. (2018), we expect head wind to have an increasing effect on fuel consumption and wind from astern to have a decreasing effect. However, this does not capture the sway and roll effect of waves and wind moving perpendicular to the vessel's direction. Due to the vessel's symmetrical design, we can account for this by retaining the features that describe the magnitude of wave height and wind speed.

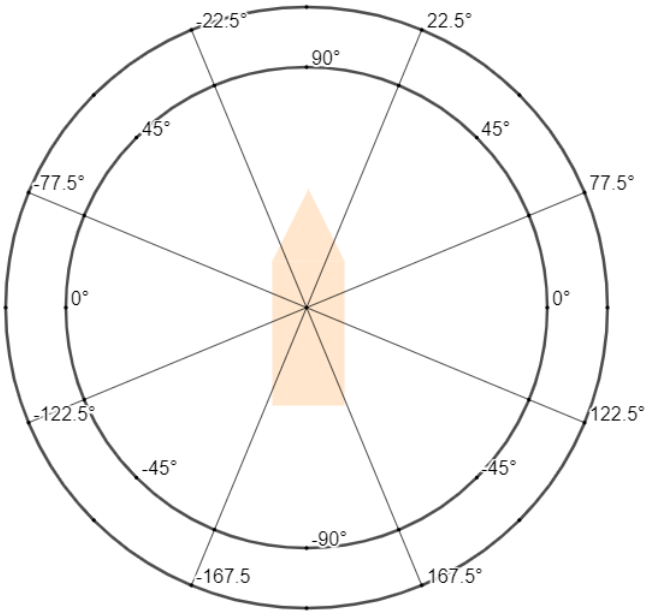

Figure 4.3: Transformation of relative directional features

Furthermore, the draft is comprised of the two measures draft_forward and draft_aft. The draft of most ships varies longitudinally from forward to the aft of the vessel; therefore, the position of the measure is often included. We opt to calculate *draft* as an average of draft_forward and draft_aft, as the two measures are highly correlated. This is displayed in Equation 4.3.

$$draft = \frac{draft\_forward + draft\_aft}{2} \tag{4.3}$$

The draft defines the vertical distance between the waterline and the ship's keel, which is useful to determine the loading conditions of the vessel (Menon, 2020). As the load of the vessel increases, so does its draft in order to support the additional weight. A vessel that sits lower in the water will have increased resistance (ceteris paribus) and consequently a higher fuel consumption (Diesel and Turbo, 2004).

Lastly, the vessel's age is identified by calculating the difference between the time of reporting (y_r) and each vessel's year of build (y_b), resulting in the equation: $age = y_r - y_b$. We opt to include the vessel's age to represent the state of a given vessel. The vessels experience a gradual hull deterioration as well as decreasing propulsion system efficiency over time. Because of this, an older ship is generally expected to consume more fuel given otherwise equal conditions.

4.3.4 Feature scaling

The selection of features in our dataset varies in magnitude, range, and units. Some machine learning algorithms are highly sensitive to a feature's magnitude, assigning features of higher value a higher weight (Aniruddha, 2020). This increases the bias towards certain features in the dataset, which can reduce the performance of the machine learning algorithm. To prevent this, we implement two feature scaling methods: standardization and normalization.

Feature standardization scales all numerical attributes to be centered around the mean with unit standard deviation, which is done by subtracting the mean and subsequently dividing by the standard deviation, as seen in Equation 4.4 (Gkerekos et al., 2019). This removes the unit differences among the features; however, the features are not restricted

to a particular range (Aniruddha, 2020).

$$X' = \frac{X - \mu}{\sigma} \tag{4.4}$$

The alternative method is normalization, also referred to as Min-Max scaling, which maps values to a range between 0 and 1. The formula for normalization can be seen in Equation 4.5.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.5}$$

The two feature scaling methods will be favored for different instances depending on specific attributes of the various machine learning algorithms (Aniruddha, 2020). For instance, algorithms that utilize gradient descent as an optimization technique, such as artificial neural network and linear regression model, require data to be scaled. Distance-based algorithms, like k-nearest neighbours, on the other hand are influenced by the range of the features. Finally, tree-based algorithms, such as Random Forest and Extra Trees, are relatively insensitive to feature scaling since these algorithms process each feature separately. By applying the most suitable scaling method to each algorithm we make sure that all features contribute equally to the final result.

The final selection of features can be found in table 4.4 below. A detailed descriptive analysis of the resulting I-Class and L-Class datasets can be found in Appendix A1.2 and A1.1, respectively.

Table 4.4: Feature descriptions

	Feature	Description	Unit
Noon Report			
1	fo_at_sea	Fuel consumption since last report	mt
2	speed_made_good	Mean speed since last report	knt
3	draft	Mean draft	m
4	age	Age of vessel	years
Copernicus			
5	mwp	Mean wave period	s
6	swh	Significant height of combined wind, waves and swell	m
7	h10	Mean wind speed	knt
8	wind_speed/dir	Wind speed relative to vessel direction	m . deg. forwards
9	wave_height/dir	Waves relative to vessel direction	m . deg. forwards

5 Methodology

The first part of this section will provide an explanation of our chosen machine learning algorithms: K-Nearest Neighbours, Random Forest, Extra Trees, and Artificial Neural Network. The main focus will be on the ensemble decision trees: Random Forest and Extra Trees. The motivation for this is based on Nilsson and Nilsson (2021) and Abebe et al. (2020) findings that ensemble decision tree models provide superior performance when predicting a vessel's fuel consumption. With this in mind, we will examine a few models of varying sophistication to substantiate their findings. A linear regression model is included to serve as a benchmark for our models' performance.

The second part will provide a detailed explanation of our optimization algorithm of choice, namely Dijkstra's algorithm. This will cover the fundamental architecture and objective function of the algorithm and its underlying assumptions and limitations. Lastly, we will demonstrate how everything comes together in our extensive simulation approach.

5.1 Machine Learning

For the optimization algorithm to function as intended, a cost function must be defined to provide context to the cost of moving along the path. This cost will largely be explained by the fuel consumed during the voyage. There are multiple alternatives to estimating the bunker fuel consumption; one alternative is to choose a selection of predetermined attributes and assign each attribute an associated weight based on its presumed relative importance until the output resembles observed values. A similar approach is achieved by fitting a linear regression on the acquired data. However, both provide imperfect estimates as they fail to explain the nonlinear dynamics of fuel consumption. Machine learning algorithms can train and learn from the collected data, which subsequently enables the algorithms to better identify the patterns, trends, and nonlinear relationships. This will conceivably provide improved bunkers fuel predictions, which is crucial to achieving realistic cost estimates.

All machine learning models presented below are related to regression analysis. The models span from simple linear regression to complex neural networks. Note that the associated complexity of each model does not necessarily correlate directly to improved

predictive performance, as this is affected by the quantity and quality of available data, as well as the complexity of the problem at hand (Gkerekos et al., 2019).

5.1.1 Cross-validation

In order to measure the models' performance, we split the data into train, test, and validation sets. The test set is used as an evaluation measure to determine how well the trained model will generalize to an independent dataset (Prashant, 2017). This requires the test set to remain untouched during the training process, which means an additional set, referred to as the validation set, is required for measuring the accuracy during this process. To avoid the risk of overfitting our models, as well as ensuring an unbiased validation set, a repeated k-fold cross-validation method is applied (Gkerekos et al., 2019). This method splits the training data into k subsets (or folds). The model is trained on the *k-1* subsets, while the remaining subset is treated as the validation and used to evaluate the model's performance. This is an iterative process that runs k number of times, and for every iteration, a model's accuracy score is obtained and averaged. This process will be repeated a specified number of times before obtaining the overall accuracy score by calculating the average of the prediction accuracy measures from each repetition. This process is visualized in figure 5.1. This approach adds significant complexity and running time to the training process; however, this is preferred to ensure an optimal training procedure for our models.

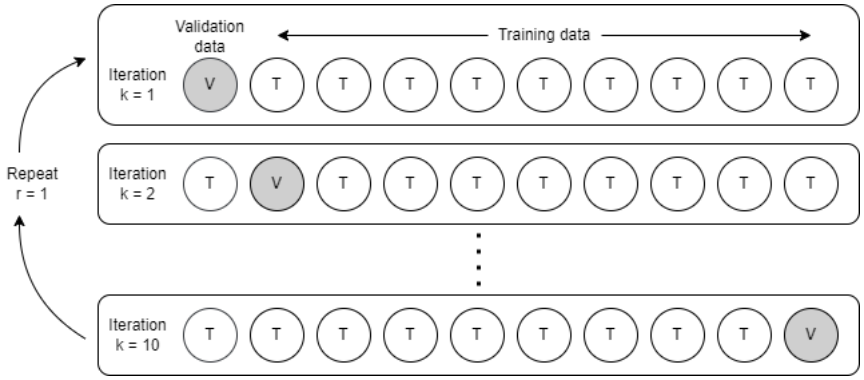

Figure 5.1: Visualization of repeated 10-fold cross validation method (Gkerekos et al., 2019)

5.1.2 Model parameters

There are two main modeling approaches, namely parametric and non-parametric models. The former assumes a finite set of parameters that are attained from the training set during the learning phase (Bishop and Nasrabadi, 2006). This set of parameters is expected to capture all variance within the dataset and thereby binding the complexity of the model (Russel and Norvig, 2010, p. 737). Linear regression and Artificial Neural Networks are examples of parametric models.

Models such as K-Nearest Neighbours, Random Forest, and Extra Trees are non-parametric models. Non-parametric models assume that the dataset distribution cannot be defined by a finite number of parameters (Gkerekos et al., 2019). For this reason, the non-parametric models can learn any functional form from the training data. Consequently, the predictive performance of these models will grow with the sample size of the training data. This approach can potentially provide higher-performance models as parametric assumptions do not restrict them; however, a sufficient dataset is required, and it will impose a higher computational cost.

5.1.3 Hyperparameter tuning

Parameters that define a model's structure are referred to as hyperparameters. A change in hyperparameter values can significantly affect the performance of the constructed model. Therefore, a process for searching for the ideal architecture should be carried out to select the proper values for each model (Abebe et al., 2020). The two most common methods of optimization are grid search and random search. The grid search tests all the potential combinations of a predefined list of hyperparameters, while the random search tests a randomly selected hyperparameter configuration within a given range. The grid search requires extensive trial and error of each hyperparameter to identify the optimal combination. Due to some of the models possessing multiple tunable hyperparameters, this process can get demanding. Furthermore, Bergstra and Bengio (2012) finds that random search can provide equal or better results than grid search; thus, we opt to apply random search as our preferred tuning approach.

5.1.4 Machine learning algorithms

We choose to implement a diverse selection of machine learning regression algorithms of varying complexity. The most basic algorithm is K-Nearest Neighbours (KNN), a non-parametric supervised method. Given a new data point, the algorithm identifies the k closest surrounding points in terms of the Euclidean distance (Russel and Norvig, 2010, p. 738-741). The predicted value of the new data point is found by calculating the mean of the identified k closest points (Christopher, 2021). The predictive performance of the model is heavily influenced by the value of k, a wide array of values should therefore be examined to identify the appropriate value of k for our model. A low value imposes a higher risk for error as the algorithm is more exposed to potential outliers. Contrarily, a higher value risks the inclusion of too many data points, which might reduce the model's accuracy, as well as increase the computational cost of the algorithm. Note that the distance of all data points is weighted equally, which causes the KNN-model to be inherently sensitive to outliers and the scale of the input data (Gkerekos et al., 2019). A tune length of 15 is chosen for our KNN-model. As a result, the model will be trained on a selection of 15 randomly sampled values of k. The value of k providing the lowest mean average error (MAE) during the validation process is chosen as the optimal value.

For our next model, we opt to implement a relatively simple configuration of a neural network, that is, a single hidden layer feed-forward neural network (NNET) (Ripley and Venables, 2022). (James et al., 2013, Chapter 10) refer to this as the most widely used "vanilla" neural network model. A typical feed-forward neural network consists of interconnected neurons (nodes) structured in three layers: input layer, hidden layer, and output layer. Each neuron in a particular layer is connected with all neurons in the following layer, where each neuron is assigned a weight based on the degree of importance of the given connection in the neural network (Svozil et al., 1997, p. 45).

In our case the input features are passed to a set of neurons consisting of a function of its input value and associated weight. These weighted values are further passed to the output layer and produce a prediction based on the sum product of all neuron values above a threshold. To recalculate the weights, gradient descent is commonly used following each learning sequence (Ruder, 2016). Our models require two tuning hyperparameters: (1) the number of nodes in the hidden layer and (2) the value of weight decay (Ripley and Venables,

2022). The number of potential configurations is vast, and an ample running time training the model prevents an exhaustive search for the optimal configuration of hyperparameters. Therefore, we implement a grid search of a small subset of hyperparameters showing promising results.

The remaining algorithms in our selection are Random Forest (RF) and Extra Trees (ET), which are algorithms that combine ensemble learning methods with the decision tree framework (Beheshti, 2022). Decision trees are structured in the form of trees with nodes and branches. It functions by forking decisions based on specific criteria until a leaf node (prediction) is reached. Ensemble methods utilize multiple weak learner models, which are trained on the same dataset, then aggregate the results of each model to produce a more robust prediction. The RF-algorithm combines these two methods by drawing multiple randomly sampled decision trees and averaging the results to improve prediction performance. The ET-algorithm is a slight modification to RF, by performing randomized splitting of the nodes rather than informed splits (Tunnicliffe, 2021).

5.2 Ship weather routing

To quantify the associated incremental cost of weather-sensitive cargo risk, we must first identify how the optimal routes of vessels with weather-sensitive deck cargo differentiate from vessels with no such cargo. We opt for a routing approach heavily inspired by Dijkstra's algorithm to achieve this. Dijkstra is one of the most prominent pathfinding algorithms, achieving accurate route cost minimization while allowing for multi-dimensional considerations, such as environmental effects, as displayed by Wang et al. (2017); Mannarini et al. (2013).

5.2.1 Dijkstra's algorithm

The fundamental structure of Dijkstra's algorithm is based on a generated waypoint grid system. To create this, we first select our starting point and ending point that marks the ship's starting position and ending position. Then, we split the space between the starting and ending positions into *stages*. We select 15 stages for our route and space them equally. After the number of stages are selected, we fill the stages with nine *nodes* each. To make sure that the nodes are within a reasonable travel path for the selected

route, the placements of the nodes are dictated by the historical positioning data from the noon reports. The grid we have created represents all possible positions a ship can visit in our model. We name the connecting line between two nodes in two different stages an *edge*. The optimal route will be a series of connecting edges from the starting point to the ending point. Further, we limit the combinations of our model by having each node only be able to connect to three nodes on the adjacent stage. This rule does not apply to the starting point and the ending point, as each node in the first and last stage must be able to access or be accessed by the starting and ending point. It is then impossible for the route to go back to a previous stage or laterally within the same stage. By limiting neighboring nodes, we drastically reduce the possible routes we need to calculate. With this approach, we are left with:

$$nodes\_per\_stage * connections\_per\_node^{(stages-1)} = 43\ 046\ 721\ possible\ paths \quad (5.1)$$

A visualization of the generated waypoint grid network is displayed in Figure 5.2.

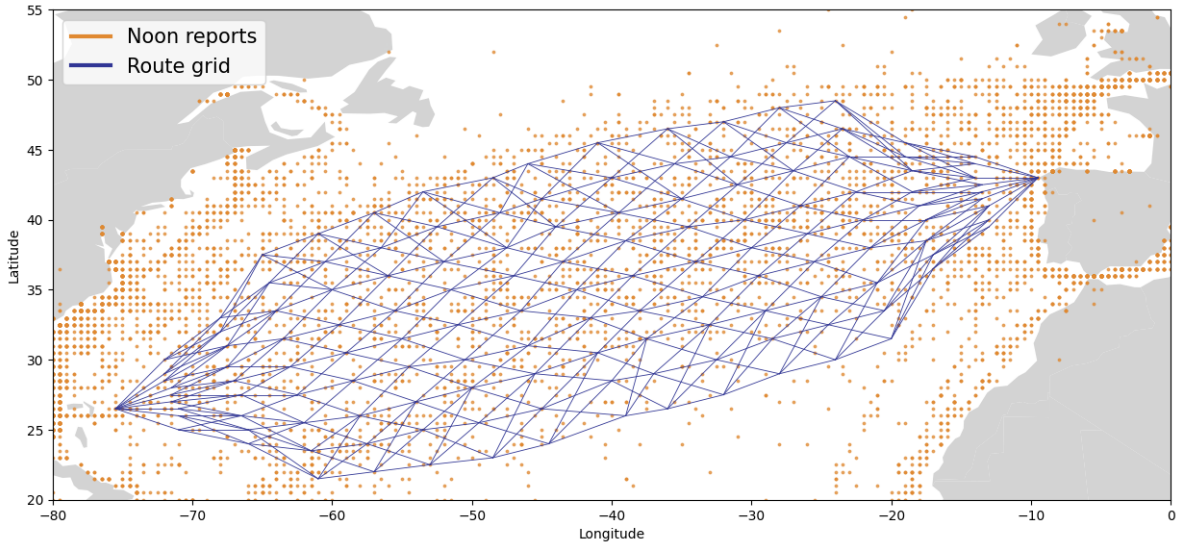

Figure 5.2: Visualization of the route grid network on top of all historic I-class and L-class vessel positioning data from noon reports over the Atlantic

The optimization algorithm's objective is to find the optimal route for any given month. To reiterate, we define the optimal route as the cost-minimizing route while not sailing over a coordinate that has observed a mean hourly wave height exceeding six meters during the day of passing. To define the route's cost, we split the cost function into two parts: the cost of fuel consumption and the cost of time. We calculate the fuel

consumption cost by multiplying the estimated consumption of the route by the bunker cost. Due to high energy demand, the current bunker cost is at $800 per metric ton (Ship & Bunker, 2022). We add the consumption cost together with the estimated time cost that is derived from the length and speed of the route and multiplied by the day rate of the ship at $25 000 per day. Note that only whole days are counted, meaning that if a trip takes 11 days and 12 hours, the day rate of 12 days will be used to calculate the cost. There are months where the model cannot identify an optimal route without breaking the defined weather constraint. For these routes, an additional day rate is added to the cost to simulate the need to wait a day for milder weather conditions. The change in ETA will likely be insignificant to the customer of the shipment, however we assume that due to vessel scarcity, the vessel must be replaced in the market, hence the extra day rate.

Due to computational limitations, pre-calculations and data pre-fetching are needed to reduce the time needed to find the optimal route. The data is calculated on an edge level, meaning that we connect all of our parameters and calculations to the edges rather than the nodes. Our pre-calculations consist of obtaining the angle and length of the edge as these remain constant between routes and months. The angle of the edge is calculated in relation to the north direction. While the length of the edge is the distance, d, calculated by using the haversine formula designed to find the length between two points on a sphere (Kettle, 2017). This calculation is displayed in Equation 5.2, where r refers to the radius of Earth, ϕ_1 and ϕ_2 is the latitude of the two points, and λ_1 and λ_2 represent the longitude of the two points, respectively.

$$d = 2r \sin^{-1} \sqrt{sin^2(\frac{\phi_2 - \phi_1}{2}) + cos(\phi_1)cos(\phi_2)sin^2(\frac{\lambda_2 - \lambda_1}{2})} \tag{5.2}$$

Furthermore, we attach the weather data to each edge. We use the weather data for the ending node of the edge as it represents the weather the ship is sailing into.

Central to our pathfinding algorithm is our machine learning model, which we use to estimate a vessel's fuel consumption when sailing an edge. For the inputs of the machine learning model, we need the following parameters from the edge: Wave height (*swh*), wave period (*mwp*), wave direction (*mwd*), wind eastward direction (*u10*), wind northward direction (*v10*), and the vessel's bearing (*bearing*). In addition, the speed (*speed_made_good*), *draft*, and *age* must be determined prior to our fuel consumption

calculations. For the ship's speed, the current climate and bunker prices dictate the ship to sail at a speed of 13.5 knots, referred to as "super eco" speed. Alternatively, a vessel's speed could be predicted based on the impact of experienced weather effects during the voyage, which would provide more nuanced speed estimations. However, the high computational effort associated with this, prohibits the inclusion of speed predictions in our approach. We therefore assume that a constant speed provides a sufficient approximation of the ship's speed during open sea sailing. This also holds true for draft predictions; we assume that the vessels will be loaded to the a draft of 10.5 meters based on recommendations from G2 Ocean. A vessel's age is also needed as a parameter in the model as a vessel experience increasing hull roughness and decreasing propulsion system efficiency as it ages. As a result, newer ships will be more energy efficient and able to consume less bunker fuel for the same speeds. We have set the age of the two classes I and L to 20 (built around 2000) and 10 (built around 2012), respectively.

From the noon reports and data from the speed and distance calculations, we found that the Atlantic crossing takes approximately 14 days. Based on this information, we can assume that there is about a day between each of our 15 stages, and can thereby map known weather observations to each node. This means that a node in the first stage will contain all corresponding hourly weather conditions of the first day of the month, a node in the second stage will contain all corresponding hourly weather conditions of the second day of the month, etc. This approach enables us to extend the model's functionality to incorporate weather forecasts on a later stage if we wish to do so. Switching to forward-looking weather data rather than historical weather data is simply a matter of changing the weather data at the nodes.

5.2.2 Methodology workflow

In accordance with the aforementioned methodology, our optimization approach can be summarized in four steps: firstly, we start with the generation of the edges that will make up each possible route between Spain and Houston. Secondly, we pre-calculate the angle and the distance of each edge and fetch all of the weather parameters attached to the edges. Thirdly, we calculate the fuel consumption for each edge using our best-performing machine learning model. Fourthly, we generate our input files for the optimization model. Lastly, to conduct our analysis, we run our route optimization simulation for all 192

months from 2006 to 2021 and find the optimal route for each month. By running the simulation for such an extended period, we decrease the impact outliers have on our results, and are therefore more likely to capture the overall trend. This process is visualized in figure 5.3.

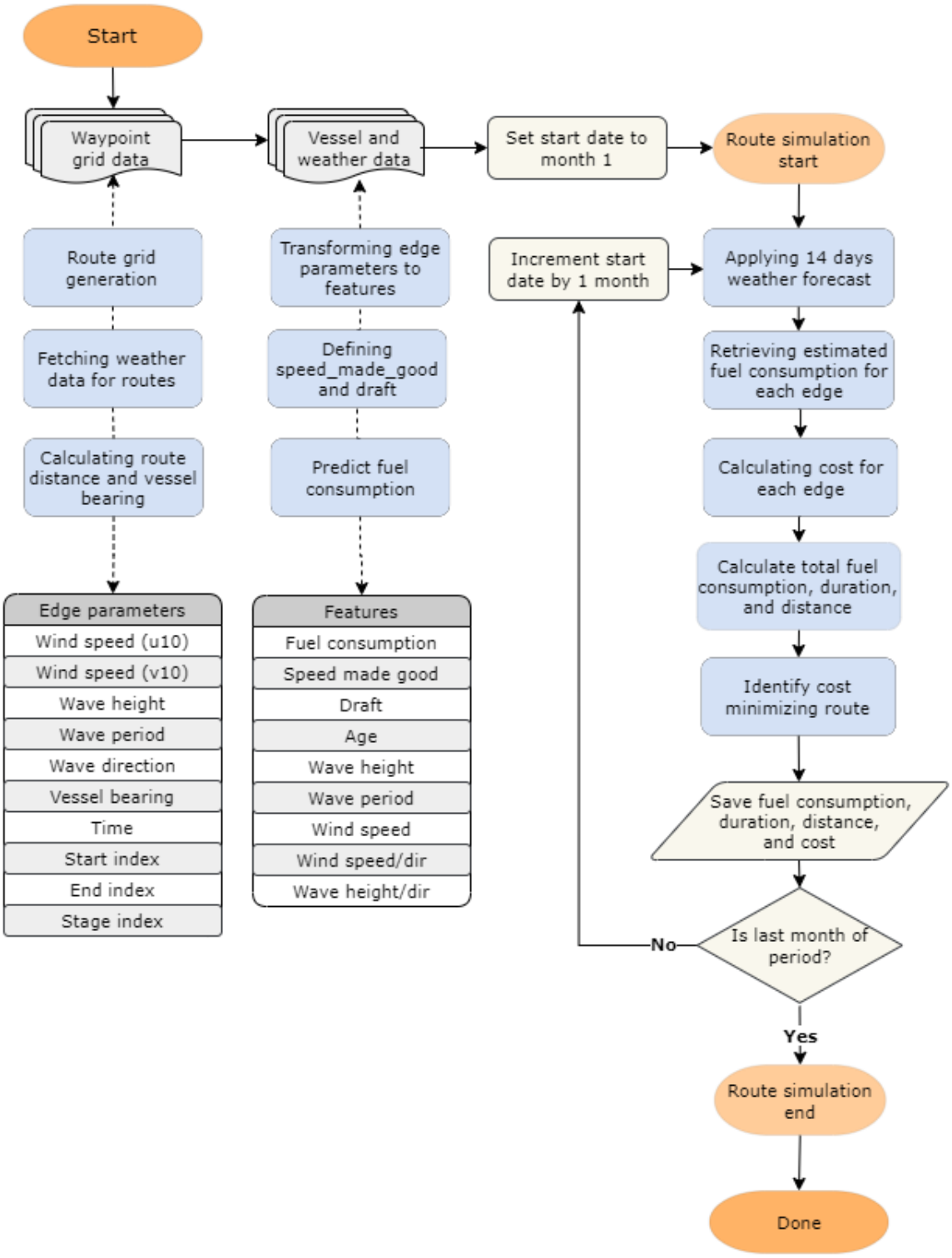

Figure 5.3: Visualization of our weather routing simulation pipeline

6 Results and Discussion

6.1 Vessel fuel estimations

We have produced two models for each machine learning algorithm by fitting the training data to the two vessel designs, *Class I* and *Class L*, independently. The notion is that each vessel design possesses different vessel-specific properties; a combination of the two classes might reduce a model's capability to identify the appropriate relationships between input values and fuel consumption. We will dedicate separate discussions for both models as their results might draw different conclusions.

6.1.1 Hyperparameters

Regarding the I-class model, the KNN-model yields a k value of 13 after tuning the model for 15 randomly drawn k numbers of neighbors. For the NNET-model, we opt for a grid search, which finds that the best performing combination of parameters is 10 in the hidden layer (size) and 0.3 as the regularization parameter (decay). Both ensemble models, RF and ET, use a random search parameter of 7, corresponding to the maximum number of input features. This yields a parameter value of 7 randomly selected predictors (mtry) for the RF-model and 5 for the ET-model. The ET-model contains an additional parameter describing the number of random cuts (numRandomCuts) during the evaluation period, which is found to be 7 for the best tuned model.

An identical approach is used during the evaluation period for the L-class model, however, the optimal parameter selection deviates to some extent from the I-class model. The KNN-model finds that a k value of 11 produces the best tune for this model. The L-class NNET-model yields quite a similar tune as the I-class, with 9 units in the hidden layer (size) and 0.3 as the regularization parameter (decay). The optimal RF-model and ET-model detect a value of randomly selected predictors (mtry) of 4 and 7, respectively. The L-class ET-model identifies a random cut parameter (numRandomCuts) of 4 to produce the best tune for the model.

We should emphasize that the selection of hyperparameters for neither the I-class nor L-class contains an exhaustive list of possible hyperparameters, but rather a comprised

subset of possible hyperparameters assumed most relevant to achieving the optimal model performance. Furthermore, the predictive performance of the models during the evaluation period is measured using MAE, as we find this to be the most unambiguous measure of error.

6.1.2 Feature importance

We implement a feature importance technique to help improve understanding of each feature's relationship with the dependent variable. This technique ranks all features based on a score, where a high score indicates high importance on the dependent variable and vice versa. There are mainly two groups of functions to evaluate the feature importance: model-specific metrics or model-independent metrics (Kuhn, 2019). We choose a model-based approach as this is more closely tied to the model's performance, and is able to incorporate the correlation structure between predictions into the feature importance calculations. Not all machine learning models possess the correct properties to determine each feature's predictive power. However, the RF-model is suited for this purpose, thus we use this model for our feature importance calculations. Figure 6.1 displays the identified feature importance when training the model for L-class vessels.

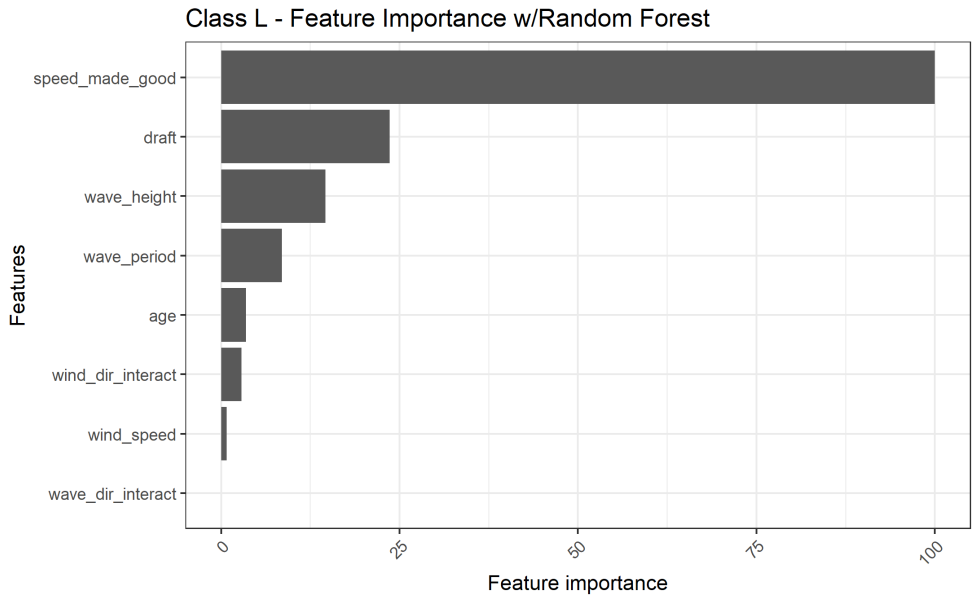

Figure 6.1: Identified feature importance for L-class vessel designs using Random Forest

A potential drawback of utilizing model-specific metrics is the risk of minor irregularities or variations between models, which might be influenced by patterns identified in a

specific model. This is to some extent found in this case as Figure 6.2 displays the feature importance of I-class vessel designs. We find that the main findings stay consistent between models, but there are some minor discrepancies between the features of lesser importance. This might result from the vast disparity in sample size, however the differences are so small that we consider them to be negligible.

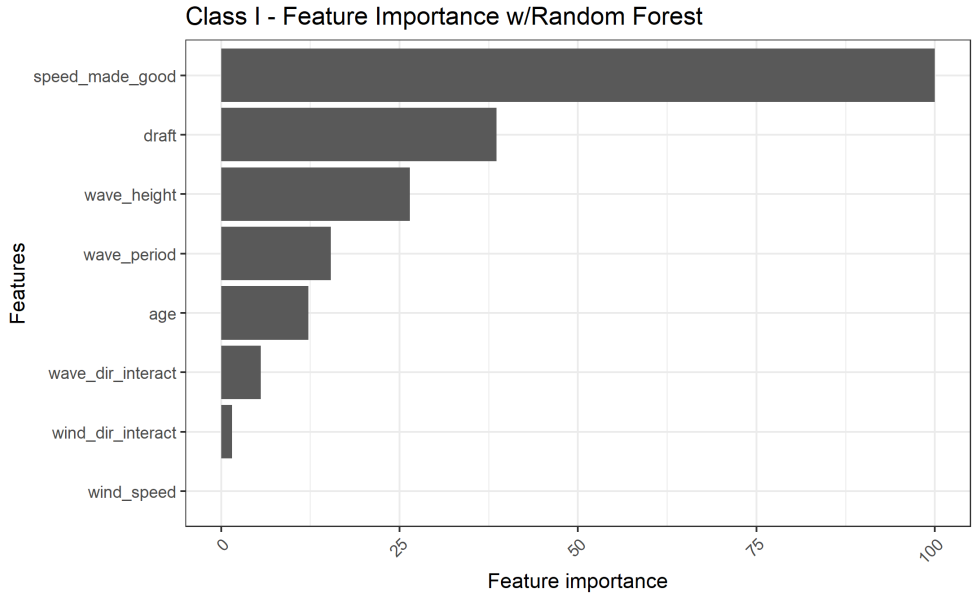

Figure 6.2: Identified feature importance for I-class vessel designs using Random Forest

It is evident that the feature *speed_made_good* seems to be of especially high importance, which is in accordance with the fuel-speed theory discussed in section 3. Next, we find that *draft* is the second most important feature. A deeper draft will expose a vessel's hull to increased resistance, thus more power is required to maintain the same speed. An increased power output results in higher fuel consumption, and it is therefore expected that a higher value of draft corresponds to higher fuel consumption. Furthermore, the *wave_height* and corresponding *wave_period* appears to be of significant importance. A higher wave height and period increase the residual resistance and causes pitch and heave motions of the vessel, consequently, we expect a significant correlation between wave effects and fuel consumption. The last feature to be of significant importance is *age*. Age corresponds to the inevitable deterioration of the vessel and its engine; it is likely that a higher age correlates with higher fuel output. The remaining directional weather features are found to be of lesser importance. The amount of influence the features exert on the dependent variable is based on the current selection of features. This means that

the relative importance can change dramatically given a different selection of features. It is possible that the exclusion of the wave and wind magnitude could have increased the relative importance of the directional weather features. However, they are included in our dataset as we believe these are necessary to achieve the highest prediction performance.

6.1.3 Model selection

To identify the best performing machine learning model for each class of vessel, each machine learning algorithm is fitted to an identical training set separated by vessel class. The predictive performance is measured using the out-of-sample test set. Table 6.1 and 6.2 displays the sMAPE, RMSE, MAE, and R^2 score for the I-class and L-class models, respectively.

Additionally, we have included a basic linear regression model (LM) to serve as a benchmark for our more complex machine learning models. As discussed in Section 3, a non-linear relationship exists between vessel fuel consumption and speed. The features *fo_ at_ sea* (dependent variable), *speed_ made_ good*, as well as *draft* should therefore ideally be log-transformed to be consistent with "cubic law". This would unfortunately diminish its capability as a benchmark as we cannot compare the residual standard errors of dependent variables on different scales. For this reason, we opted not to log-transform the features of the linear regression model. However, to provide context we have included the performance metrics of the corresponding log-transformed linear regression model in Table A2.1 in the Appendix.

Table 6.1: [Class I] Performance metrics of models

Models	sMAPE (%)	RMSE	MAE	RSquared
Extra Trees	8.155	3.209	1.639	0.886
Random Forest	12.770	3.974	2.612	0.839
KNN	22.314	6.787	5.005	0.478
NNET	22.178	6.852	4.993	0.462
LM	24.403	7.459	5.582	0.381

Table 6.2: [Class L] Performance metrics of models

Models	sMAPE (%)	RMSE	MAE	RSquared
Extra Trees	6.269	2.566	1.219	0.894
Random Forest	10.007	3.106	1.977	0.855
KNN	17.809	5.148	3.713	0.568
NNET	18.154	5.295	3.809	0.540
LM	20.393	5.800	4.223	0.443

It is apparent that regardless of which performance metrics are used, the two ensemble models, Extra Trees (ET) and Random Forest (RF), perform significantly better than the remaining models. Both models have a R^2 score in the range of 0.8 to 0.9, which means they explain close to 90% of the observed variance. The L-class ET model has a MAE value of 1.219, while the I-class has a MAE value of 1.639. In terms of fuel consumption, this indicates an average prediction error of 1.219 and 1.639 metric tons of fuel per day. This is far superior to the linear benchmark model producing an average error of 4.223 for the L-class and 5.582 metric tons for the I-class. We can safely conclude that the ET and RF models rank number 1 and 2, respectively, in regards to predictive performance. However, due to some unreliable dependencies of the ET models that occasionally produce computational difficulties, we have opted to use the RF models as our best-performing models instead. Consequently, this will, unfortunately, reduce the accuracy of our fuel predictions to some extent.

6.1.4 Prediction performance

As our fuel prediction approach involves estimating the daily cumulative fuel consumption for a specific vessel design, the predictive performance of our models should be evaluated based on error minimization of fuel predictions for an I-class and a L-class vessel rather than a generic vessel. For this purpose, we have arbitrarily selected one vessel from each class, namely Star Istind and Star Lysefjord to represent I-class and L-class, respectively. The predictive performance is displayed in Figure 6.3 and Figure 6.4.

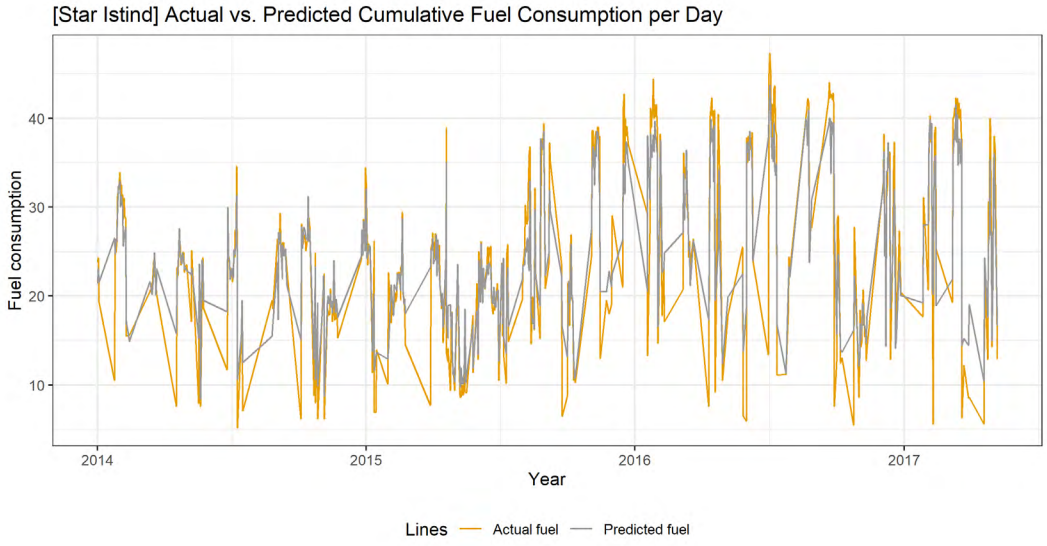

Figure 6.3: Daily cumulative fuel consumption observations plotted against daily cumulative fuel consumption predictions using the Random Forest model. The sample size includes all fuel consumption reports of the vessel Star Istind

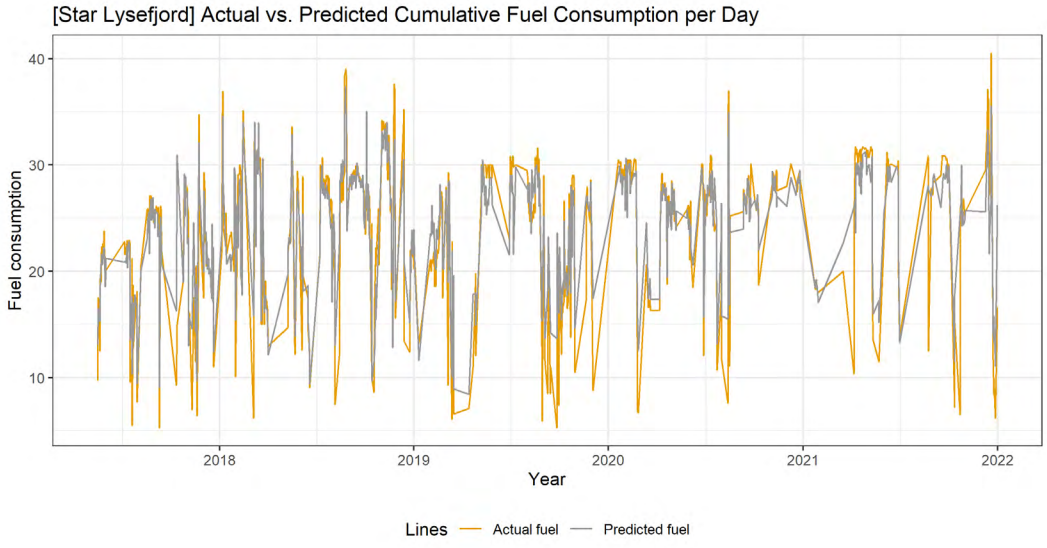

Figure 6.4: Daily cumulative fuel consumption observations plotted against daily cumulative fuel consumption predictions using the Random Forest model. The sample size includes all fuel consumption reports of the vessel Star Lysefjord

We find that the daily cumulative predictions do not correspond perfectly with the actual daily cumulative fuel consumption. The models are able to capture the general trend and variance in the observed data, however the models struggle to explain the most extreme fuel observations. Further, we want to examine the performance over an extended period. We opt for a 14-day period as it is the approximate duration of a voyage. The performance

of the I-class model can be observed in Figure 6.5. We find that the model, though not perfectly accurate on a daily level, performs sufficiently well over an extended period. We calculate a MAE score of 8.975, representing an average prediction error of 8.975 metric tons of fuel over a 14-day period. We considered this to be an insignificant amount, and conclude that our models are relatively unbiased and function as ample vessel fuel predictors.

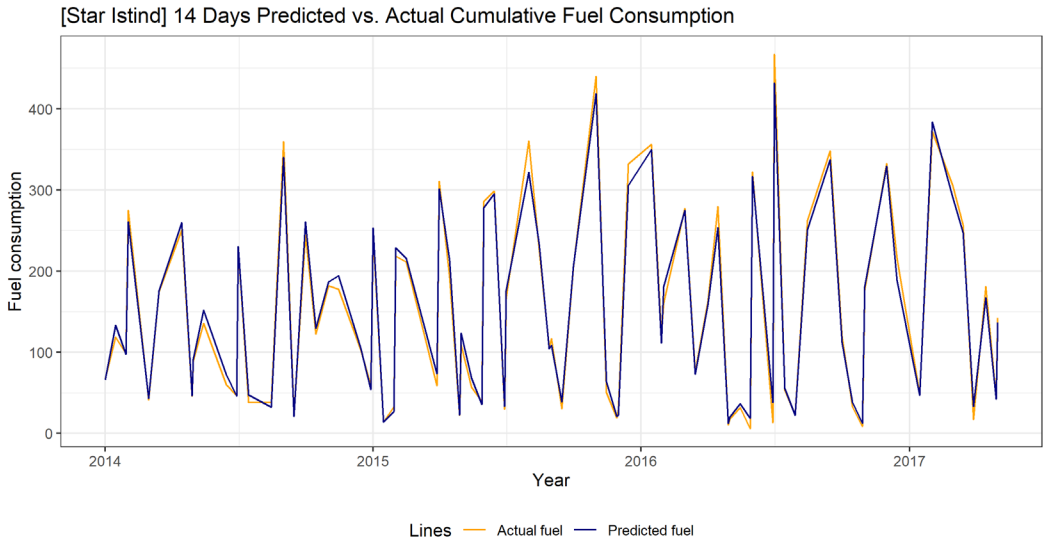

Figure 6.5: The difference between actual cumulative fuel consumption and predicted cumulative fuel consumption over a period of 14 days. The data includes all fuel consumption reports of the vessel Star Istind

6.2 Ship weather routing

This segment will highlight the results from our weather routing model. The results are summarized descriptive data, consisting of the optimal routes for each of the 192 months in our simulation. Our preliminary assumption is that the seasonal weather conditions will have a substantial impact on optimal route selection. As displayed in Figure 2.2 in Section 2, we find significantly more frequent extreme weather observations in the winter months compared to the summer months. Consequently, we will provide separate discussions for the summer and winter months to examine if the seasonality differences affects the expected voyage cost. To clarify, by summer months we refer to the months of April, May, June, July, August, and September, while winter months refer to the remaining months, namely January, February, March, October, November, and December.

Our primary objective is to identify the incremental cost associated with the weather-sensitive cargo risk. This is found by analyzing the differences in comparative routes between vessels with and without weather constraints. To generalize our results, the findings are presented as the relative incremental cost and not as absolute cost differences. Furthermore, we would like to emphasize that we have produced separate results for the I-class and L-class vessels. As these models produce relatively similar results, we will focus our attention on the results from the L-class. This is due to the superior sample size of the L-class model, which ensures reliability of results and improved predictions performance, as shown in Table 6.1 and Table 6.2. The results of the I-class is displayed in Table A5.1 in the Appendix.

6.2.1 Route selection

The figures 6.6 and 6.7 presents an overview of the general differences in route selection based on seasonality. These contain the optimal route for each month between 2006 and 2021 generated by our optimization algorithm. Two examples of these routes are included in the Appendix (Figure A4.1 and Figure A4.2). The dotted line in the figures 6.6 and 6.7 represents the closest approximation to the great circle line. Given no weather constraints and zero cost benefits by choosing a route with milder weather conditions, all routes should follow this line. Additionally, the figures include the dispersion of coordinates with historical wave height observations exceeding 6 meters to showcase how the seasonality

effect influences the routing choices. The visualization includes the total number of wave height observations above 6 meters from the period 2006 to 2021. The total number ranges from 0 to 8000 observations for each coordinate, where a highly saturated coordinate indicates a high number of observations.

There are apparent differences between the summer and the winter route selections and nuances within the figures. The most striking observation is the differences in extreme weather conditions. The weather in Figure 6.6 showcases a high variation in terms of the number of extreme weather observations, including coordinates with more than 8000 observations, all the way down to coordinates at the lower end of the scale at less than 1000. There is also a clear distinction in weather between the northern and southern regions of the Atlantic. The same distinction is not as evident during the summer months, as seen in Figure 6.7. It is difficult to distinguish coordinates that experience more extreme weather than others, as almost all coordinates have less than 1000 adverse weather observations. Consequently, we observe that the winter routes in Figure 6.6 span a much greater number of selected routes compared to the summer routes in Figure 6.7. While the summer routes are primarily grouped around the great circle line, the winter routes can be scattered well south of the shortest path, opting to take a significant distance penalty to avoid weather constraints and optimize cost savings.

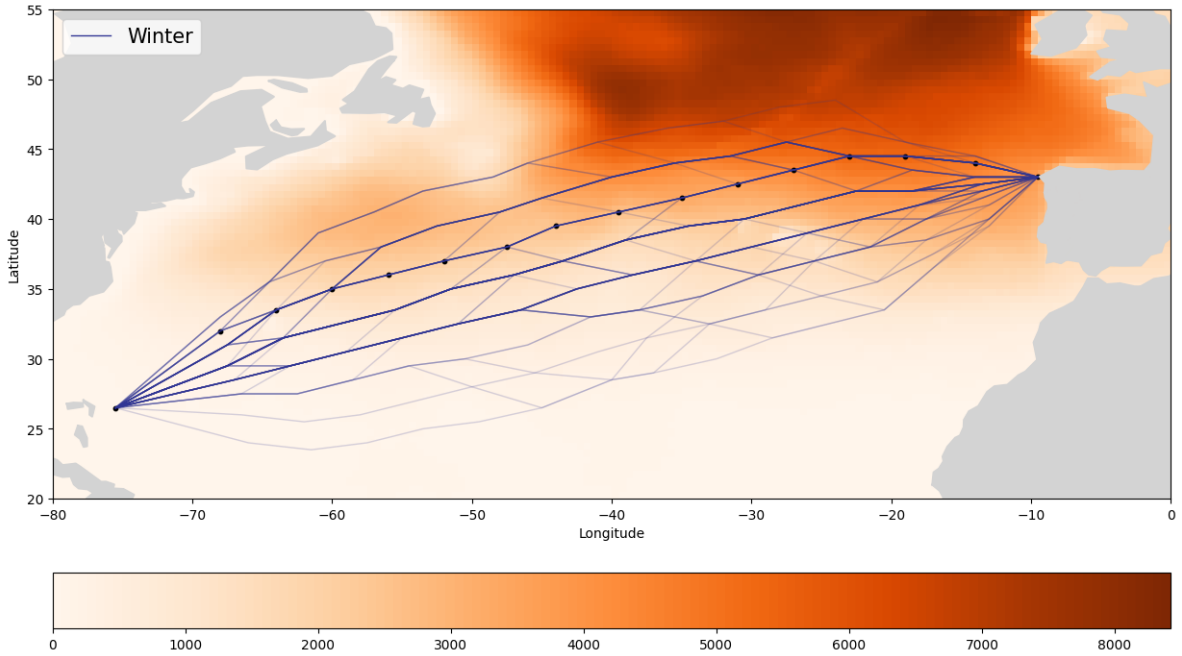

Figure 6.6: Routes generated by simulating for winter months with heatmap of extreme weather observations (wave height $>$ 6m). The dotted line is the shortest path possible in our route grid

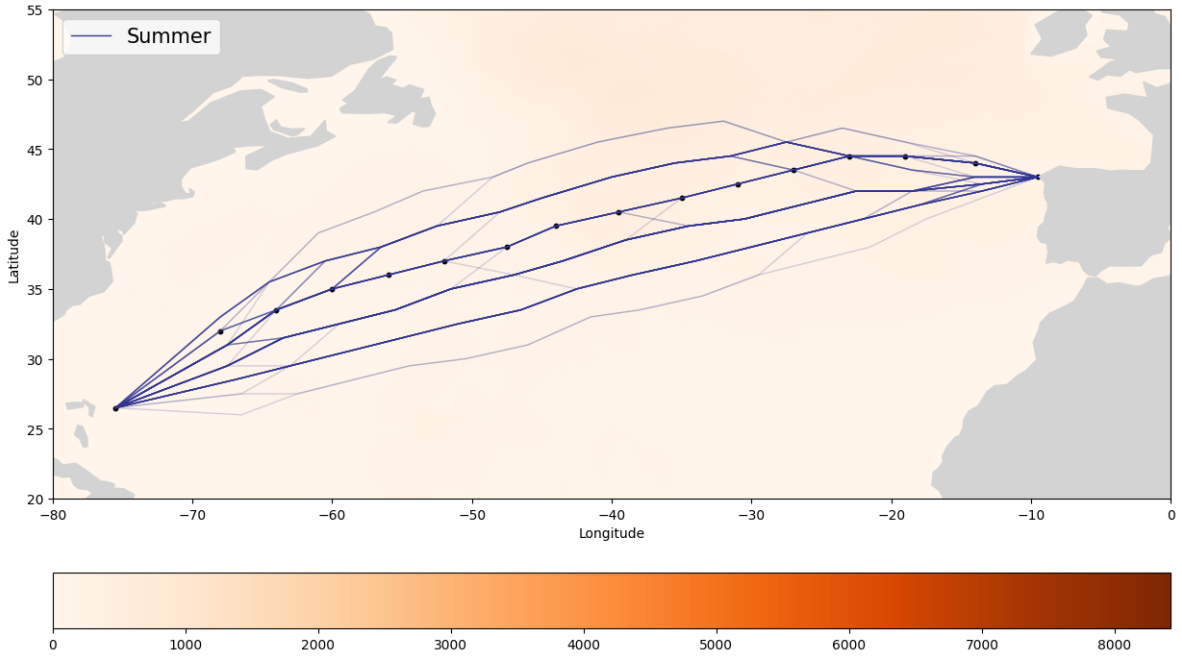

Figure 6.7: Routes generated by simulating for summer months with heatmap of extreme weather observations (wave height $>$ 6m). The dotted line is the shortest path possible in our route grid

The clear difference in route selection between Figure 6.6 and in Figure 6.7, indicates that there is an obvious increased risk associated with taking on a weather-sensitive shipment

in the winter compared to the summer season. Although, we do find a surprisingly high variance in route selection within the summer months. This variance can imply that there is still some cost-benefit from diverging from the shortest path and sailing into calmer seas. Another interesting remark is that we still see many winter routes covering the same area as the summer routes, which implies that there are still many winter routes that are not constrained by adverse weather conditions. The clear trend, however, is that if the ship needs an alternative route when subjected to adverse weather conditions, the ship opts to take a path further south. This can be an instructive hedge technique if an operator wants to reduce the chance of running into rough weather.

6.2.2 Cost estimation

Now that we have explored the dissimilarities in seasonal route selection, we would like to examine if this translates to differences in cost estimates. We investigate the incremental differences in cost and relevant parameters for vessels subjected to weather constraints and vessels that are not. We find the distance, time, fuel, and cost from the generated routes of vessels subjected to weather constraints and then subtract the results from the corresponding routes of vessels not subjected to weather constraints. This implies that vessels subjected to weather constraints, but do not deviate from the optimal path due to insufficient waves, will have no additional cost. We look at the results in the following table 6.3.

Table 6.3: [Class L] Results from Comparative Analysis of Vessels Transporting Weather Sensitive deck cargo and Regular Cargo

	Change in Parameters (%)			Incremental Cost (%)				
Month	Distance (km)	Time (hrs)	Fuel (mt)	Mean	Max.	Min.	Var.	Std.
January	0.94	0.94	0.99	1.74	9.30	0.00	7.93	2.82
February	1.54	1.54	1.47	2.26	10.80	0.00	15.22	3.90
March	1.41	1.41	2.37	1.49	13.80	0.00	12.11	3.48
April	0.07	0.07	0.03	0.01	0.20	0.00	0.00	0.05
May	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
June	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
July	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
August	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
September	0.05	0.05	0.11	0.05	0.70	0.00	0.03	0.18
October	0.96	0.96	0.89	0.43	4.40	0.00	1.22	1.10
November	0.57	0.57	0.47	0.21	1.60	0.00	0.22	0.46
December	1.88	1.88	2.00	2.56	10.70	0.00	14.25	3.78

Similarly to the results from Figure 6.6 and Figure 6.7, we identify an indisputable difference in incremental cost between routes generated in the winter months compared to the summer months, displayed in Table 6.3. We find that the months between December and March have the highest associated risk, where certain routes can be upwards of 13.80% more costly when carrying weather-sensitive cargo. On the other hand, the same months have a mean incremental costs closer to the minimum than the maximum value. This bias implies that there are only a small number of routes that deviate a lot compared to those that do not, meaning there are still good chances of normal sailing during these months. Note that all months have routes that do not deviate from the optimal path, as all months have a minimum incremental cost of zero. Summer months have a maximum incremental cost of 0.70%, meaning that they are either barely affected or not affected by weather constraints. The mean incremental cost tells us that the cost of transporting weather-sensitive cargo often does not differ substantially from regular cargo. Nevertheless, the high variance suggests that the transportation of weather-sensitive cargo is still significantly exposed to the risk of higher costs.

Based on the results from Table 6.3, we believe that there is potentially a big difference in choosing which month to carry weather-sensitive cargo. For instance, waiting a month between March and April results in a significant risk reduction in increasing costs or cargo loss. The summer months are generally safe to carry weather-sensitive cargo, as they rarely encounter weather constraints. Suppose there is a need to transport weather-sensitive cargo during the winter; our findings suggests the sailing to have a fairly low chance of substantially deviating from the optimal path due to weather constraints. However, the operators still have to be mindful of the worst-case scenario imposing an additional cost of 13.80%.

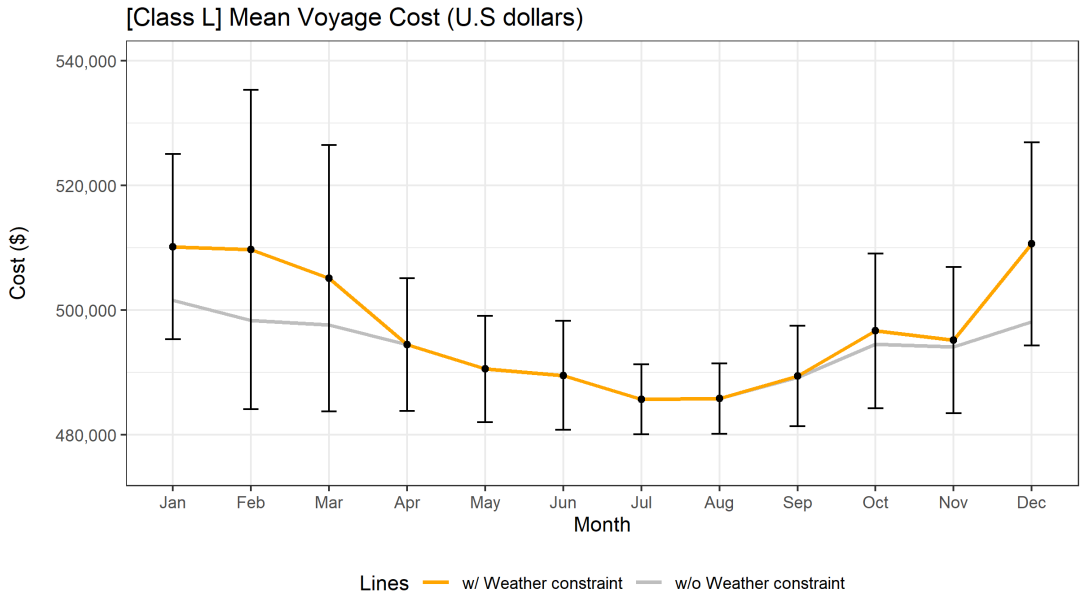

Figure 6.8: Cost development throughout the year compared to non-weather constrained routes

Figure 6.8 visualizes the differences in absolute voyage cost between vessels with and without weather constraints. We see that in the summer months, from April to September, the mean of the weather-constrained routes is close to or equal to the mean of the non-constrained routes. The cost variance in these routes is also minuscule compared to the winter routes. As the months turn to winter, the mean of the weather-constrained routes deviates progressively from the mean of the non-constrained routes. In addition, we see that the cost estimate gets increasingly more uncertain in the winter months, peaking in February, where we can expect a cost ranging between $485 000 and $535 000.

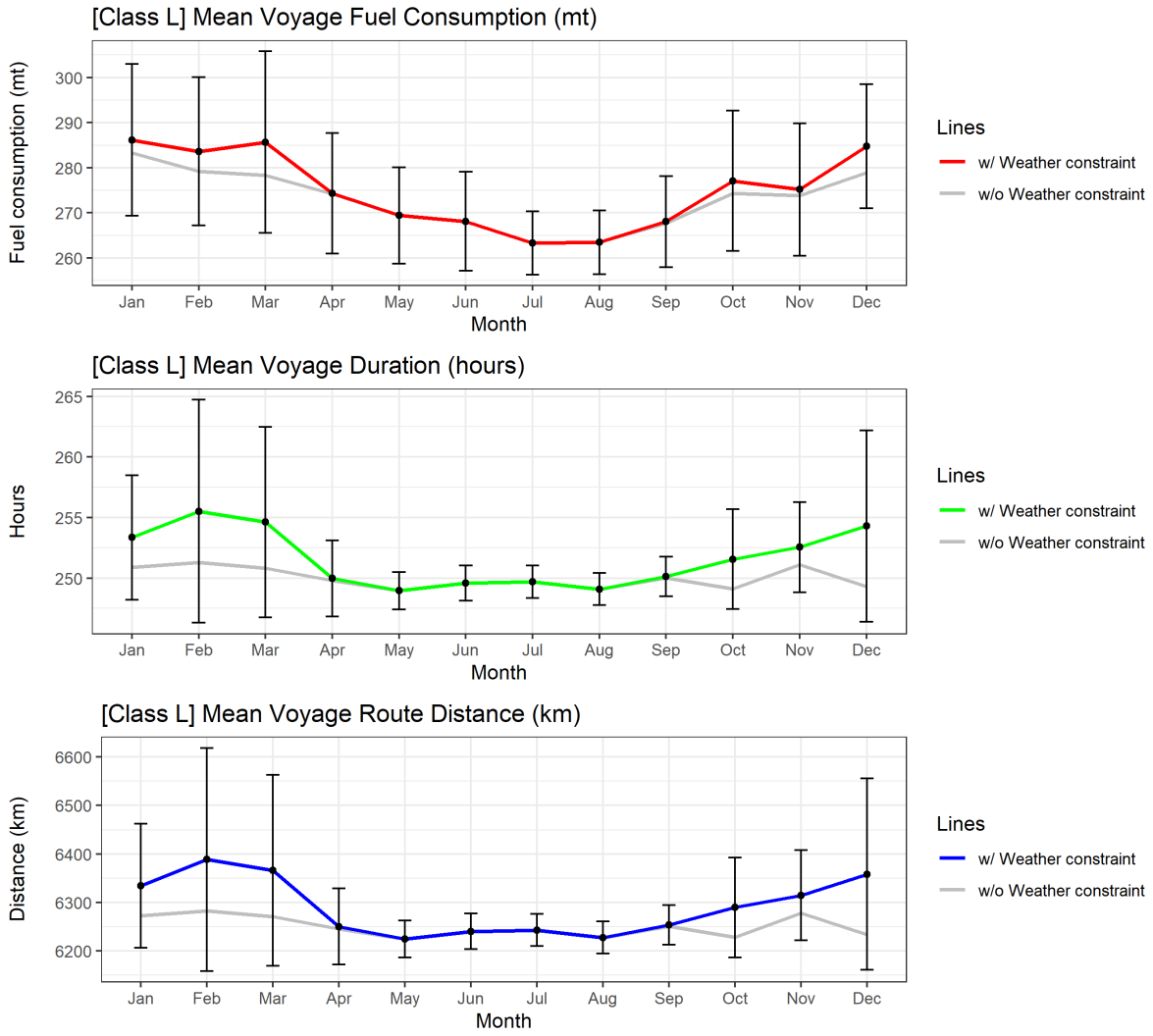

Figure 6.9: Class L voyage statistics displaying fuel consumption (mt), duration (hrs), and distance (km) compared to non-weather constrained routes

The same pattern is also true for the differences in relevant voyage parameters between the constrained and non-constrained routes, presented in Figure 6.9. This figure showcases the monthly variations in total voyage duration, distance, and fuel consumption. We find that the difference in duration and distance is greater than the fuel consumption. This is likely because fuel consumption decreases as the deviating route selects calmer seas, reducing energy demands. In comparison, the duration and distance parameters only depend on the distance between the nodes included in the route. Note that the relationship between duration and distance is constant as our model does not support varying speeds. The variance of the fuel consumption is also more consistent throughout the months than the variance of duration and distance, indicating that weather conditions play a significant role in the consumption.

6.2.3 Sensitivity analysis

Further, we will examine how the value of the weather constraint impacts the expected incremental cost. This study assumes that there is no uncertainty in the weather observations. In a forward-looking system, however, a buffer on wave heights should be utilized due to the potential uncertainty of the predictions. The advantage of knowing the cost of such a buffer will be significant as it sets a price of being too careful in regards to adverse weather. To examine this, we have simulated route selection for the 192 months, constraining the simulations by a limit on wave heights ranging from 4 to 8 meters with half meter increments. Figure 6.10 and Table 6.4 showcase the generated plot and table of the additional cost and variance for each wave constraint.

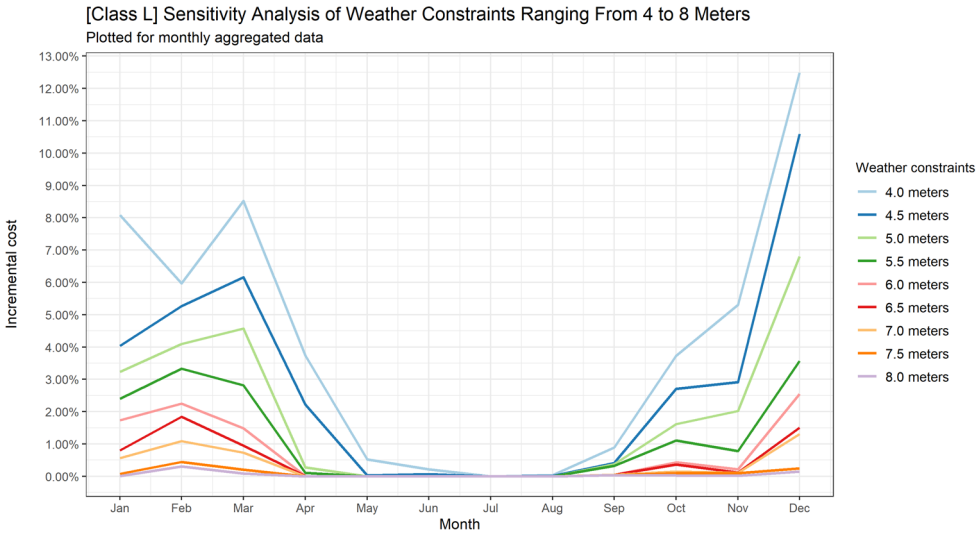

Figure 6.10: Sensitivity analysis showcasing the expected incremental cost for a weather constraint ranging from 4 to 8 meters

From the Figure 6.10, we observe primarily the same pattern for the summer and winter months identified in our prior results. An additional cost is found on almost every level of wave height during the winter months. The mean additional cost is found to increase by 1-2 percentage points for almost every level of wave height. We can therefore conclude that the winter months are relatively sensitive to wave constraints. However, the interesting fact is that the summer months do not see a significant increase in additional costs. There is almost no change between April and September until we set a wave constraint of 4.5 meters. May to September is barely affected even with a wave constraint of as low as 4 meters. These results further demonstrate the difference between carrying weather-sensitive cargo

in the summer and winter when trying to minimize the risk of additional costs.

Table 6.4: [Class L] Sensitivity analysis for wave constraint ranging from 4-8 meters

	Summer Incremental Cost (%)					Winter Incremental Cost (%)				
Constraint	Mean	Max.	Min.	Var.	Std.	Mean	Max.	Min.	Var.	Std.
4.0	0.90	12.09	0.00	5.19	2.28	7.35	23.57	0.00	38.83	6.23
4.5	0.46	8.12	0.00	2.34	1.53	5.28	21.66	0.00	34.47	5.87
5.0	0.11	5.10	0.00	0.31	0.55	3.72	18.97	0.00	26.13	5.11
5.5	0.07	5.10	0.00	0.28	0.53	2.33	15.91	0.00	18.73	4.33
6.0	0.01	0.73	0.00	0.01	0.08	1.45	13.76	0.00	8.77	2.96
6.5	0.01	0.73	0.00	0.01	0.08	0.93	11.41	0.00	5.59	2.36
7.0	0.01	0.73	0.00	0.01	0.07	0.66	9.53	0.00	3.92	1.98
7.5	0.01	0.54	0.00	0.00	0.06	0.19	3.42	0.00	0.30	0.55
8.0	0.01	0.54	0.00	0.00	0.06	0.10	1.92	0.00	0.11	0.33

Looking at Table 6.4, we observe considerably higher additional costs for the winter months for every level of wave constraint. To exemplify the differences, we will look at a scenario where one would like to consider the extra costs associated with having a 1 meter wave height buffer from 6 meters, which implicitly mean sailing with a wave constraint of 5 meters. From Table 6.4, we find an expected 0.11% increase in costs on average if we choose to transport weather-sensitive cargo during the summer months compared to a 3.72% increase during the winter months. This incremental cost can be expected to vary by 0.55% in the summer compared to the 5.11% in the comparable winter months. However, the vast span in variance suggests that we actually risk an additional cost of 5.10% in the summer compared to 18.97% in the winter. When looking at the difference between using 5 meter as the constraint compared to 6 meters, it is evident that one exposes the sailing to an ampler risk range of divergences using a buffer. Nonetheless, the increase is always much higher in the winter compared to the summer months.

7 Limitations and Further research

We acknowledge that our study has certain weaknesses that can affect the accuracy of our findings. First, we rely on the use of noon reports during the training process of our machine learning algorithms. As discussed in Section 4, the noon reports are inherently uncertain and are considered an inferior alternative to the automated data logging & monitoring systems (Aldous et al., 2015). However, it is likely that we manage to reduce this uncertainty by including third-party weather data and performing a rigorous pre-processing. This is supported by the high explanatory power of our machine learning models. Furthermore, it should be emphasized that the results of our fuel consumption estimates can only be considered applicable to the respective vessel designs used during the training process and not as an objective prediction model for all vessel types.

Second, as the noon reports only register the consumption since the previous report, the weather observations should preferably be matched at the time and location of the midpoint between the preceding report and the point of reporting. This could arguably provide a more accurate representation of the observed conditions since the previous report. Instead, we opted to predict the fuel consumption based on the weather observations at the reporting point. This was done to correspond with the approach used in the pathfinding algorithm, which, due to its structure, is limited to use either the starting or ending nodes as reference. Additionally, due to the low spatial resolution of our third-party weather data, it is possible that we would be unable to match the appropriate weather observation to the corresponding midpoint. Improper matching would increase the uncertainty of our dataset as the route may reference weather points that are not actually on the route.

Lastly, our pathfinding algorithm identifies the most optimal route from the generated grid, which makes the accuracy of our findings highly dependent on the grid resolution (Wang et al., 2017). As the grid resolution is the direct cause of the computational effort required, we cannot construct a sufficiently high grid resolution to cover all potential routes.

For further research into the transportation of weather-sensitive cargo, the optimization algorithm should be expanded to include additional factors that affect the cargo load and thus affect the limitations of sailing. When assessing routes, factors like the length

of waves and wave duration relative to the ship structure and specifications should be considered. Additionally, the specific details about the wave's effect on the given cargo should be considered as some cargo and cargo supporting structures require different considerations than others.

Due to simplifications, the optimization model is limited by setting a predetermined time at each stage to fetch each node's weather data. Optimally the model should fetch weather data for a node live as the solver checks the nodes in the route. The solver could then support cases where an edge will have different data associated with it, depending on the preceding edges. In this way, the model can consider more accurate and realistic data when making the routing decision. Live weather data fetching at runtime in the solver would also enable the possibility of performing live consumption predictions on each edge and significantly increase the method's applicability. Lastly, we would like to include actual predictions of speed on a route rather than let this be a constant like it is now. The speed variation would likely impact fuel consumption, yielding more accurate cost results.

8 Conclusion

In this thesis, we have sought to determine the incremental cost associated with weather-sensitive cargo risk. We have examined how adverse weather affects the routing choices of vessels transporting weather-sensitive components compared to vessels without this type of cargo. We propose a ship weather routing model comprised of one of the prevailing pathfinding algorithms combined with a sophisticated ensemble machine learning model. Our machine learning model produces highly accurate fuel consumption estimates, which provide context to the vessel's cost of sailing along a given route. This approach enables multi-objective optimization, allowing the model to identify the optimal cost-minimizing route while accounting for the associated environmental constraints of weather-sensitive cargo.

As a part of our collaboration with G2 Ocean, we wanted to explore routes relevant to their operations. Therefore, our thesis's scope only covers the route from Spain to Houston. This limits the relevant area of this study to the northern part of the Atlantic. Additionally, G2 Ocean has supplied us with noon reports from two of their vessel designs, namely I-class and L-class. These have, in combination with third-party meteorologic data, been our primary data sources. The data span the period from 2009 to 2021, providing a sufficient sample size for our analysis.

Our findings indicate that the vessels transporting weather-sensitive deck cargo have a significantly higher tendency to deviate from the objectively considered optimal route relative to vessels without this type of cargo. The results further identify that the seasonal variations in weather conditions heavily influence the route selection. The winter months typically experience substantially rougher weather conditions relative to the summer months. This results in the winter months yielding incremental costs upwards of 13.8%. Contrarily, sailing during the summer months only yield a maximum incremental cost of 0.70%, indicating that the vessels largely follow the optimal route of vessels with cargo impervious to adverse weather during these months. Therefore, we conclude that there is a significantly higher risk of increased costs associated with the transport of weather-sensitive deck cargo from Spain to Houston, however the degree of impact is seasonal determined.

References

Abebe, M., Shin, Y., Noh, Y., Lee, S., and Lee, I. (2020). Machine learning approaches for ship speed prediction towards energy efficient shipping. *Applied Sciences*, 10(7).

Adland, R., Cariou, P., Jia, H., and Wolff, F.-C. (2018). The energy efficiency effects of periodic ship hull cleaning. *Journal of Cleaner Production*, 178:1–13.

Adland, R., Cariou, P., and Wolff, F.-C. (2020). Optimal ship speed and the cubic law revisited: Empirical evidence from an oil tanker fleet. *Transportation Research Part E: Logistics and Transportation Review*, 140:101972.

Aldous, L., Smith, T., and Bucknall, R. (2013). Noon report data uncertainty.

Aldous, L., Smith, T., Bucknall, R., and Thompson, P. (2015). Uncertainty analysis in ship performance monitoring. *Ocean Engineering*, 110:29–38. Energy Efficient Ship Design and Operations.

Aniruddha, B. (2020). Feature scaling for machine learning: Understanding the difference between normalization vs. standardization. https://www.analyticsvidhya.com/blog/2020/04/feature-scaling-machine-learning-normalization-standardization/. Accessed: 2022-05-01.

Anish, W. (2021). What is noon report on ships and how is it prepared? https://www.marineinsight.com/guidelines/what-is-noon-report-on-ships/. Accessed: 2022-04-21.

Arribas, F. P. (2007). Some methods to obtain the added resistance of a ship advancing in waves. *Ocean Engineering*, 34(7):946–955.

Beheshti, N. (2022). Random forest regression. https://towardsdatascience.com/random-forest-regression-5f605132d19d. Accessed: 2022-05-03.

Bellman, R. (1954). The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13:281–305.

Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.

Bommert, A., Sun, X., Bischl, B., Rahnenführer, J., and Lang, M. (2020). Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics & Data Analysis*, 143:106839.

Carlton, J. (2018). *Marine propellers and propulsion*. Butterworth-Heinemann.

Chen, H. (1978). *A dynamic program for minimum cost ship routing under uncertainty*. PhD thesis, Massachusetts Institute of Technology.

Christopher, A. (2021). K-nearest neighbor. https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4. Accessed: 2022-05-03.

Chu, P. C., Miller, S. E., and Hansen, J. A. (2015). Fuel-saving ship route using the navy's ensemble meteorological and oceanic forecasts. *The Journal of Defense Modeling and Simulation*, 12:41–56.

Clarke, B., Fokoue, E., Zhang, H. H., et al. (2009). *Principles and theory for data mining and machine learning*. Springer.

Clarksons Research (2022). World fleet register. https://www.clarksons.net/n/#/portal. Accessed: 2022-05-30.

Diesel, M. and Turbo, S. (2004). Basic principles of ship propulsion. *MAN Diesel & Turbo publication*.

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

G2 Ocean (2022a). About g2 ocean. https://www.g2ocean.com/about-us/. Accessed: 2022-05-01.

G2 Ocean (2022b). Project cargo services. https://www.g2ocean.com/project/project-cargo/. Accessed: 2022-05-01.

Gkerekos, C., Lazakis, I., and Theotokatos, G. (2019). Machine learning models for

predicting ship main engine fuel oil consumption: A comparative study. *Ocean Engineering*, 188:106–282.

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

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., J., M. S., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2018). Era5 hourly data on single levels from 1979 to present. copernicus climate change service (c3s) climate data store (cds). https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form. (Accessed on: 2022-04-23), 10.24381/cds.adbb2d47.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

James, R. W. (1957). *Application of wave forecasts to marine navigation*. New York University.

Jie, J. and Miao, C. (2021). Analysis and selection of shipping route in ocean. In *Journal of Physics: Conference Series*, volume 2029, page 012151. IOP Publishing.

Karney, C. F. F. (2013). Algorithms for geodesics. *Journal of Geodesy*, 87:43–55.

Kettle, S. (2017). Distance on a sphere: The haversine formula. https://community.esri.com/t5/coordinate-reference-systems-blog/distance-on-a-sphere-the-haversine-formula/ba-p/902128. Accessed: 2022-05-18.

Kim, S. and Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3):669–679.

Kossin, J. P. (2008). Is the north atlantic hurricane season getting longer? *Geophysical Research Letters*, 35(23).

Kuhn, M. (2019). the caret package. https://topepo.github.io/caret/index.html. Accessed: 2022-05-14.

Lin, Y.-H., Fang, M.-C., and Yeung, R. W. (2013). The optimization of ship weather-

routing algorithm based on the composite influence of multi-dynamic elements. *Applied Ocean Research*, 43:184–194.

Liu, J. H. (2021). Shipment of project and equipment cargoes. https://www.skuld.com/topics/cargo/project-cargo/shipment-of-project-and-equipment-cargoes/. Accessed: 2022-05-01.

Magnussen, A. K. (2017). Rational calculation of sea margin. Master's thesis, NTNU.

Mannarini, G., Coppini, G., Oddo, P., and Pinardi, N. (2013). A prototype of ship routing decision support system for an operational oceanographic service. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, 7(1).

Meng, Q., Du, Y., and Wang, Y. (2016). Shipping log data based container ship fuel efficiency modeling. *Transportation Research Part B: Methodological*, 83:207–229.

Menon, A. (2020). What is draft or draught of a ship? https://www.marineinsight.com/naval-architecture/vessel-draft-vessel-draught-ship/. Accessed: 2022-05-25.

Nilsson, J. and Nilsson, M. (2021). Estimating weather margin seasonality in shipping using machine learning. Master's thesis, Norwegian School of Economics.

Padhy, C. P., Sen, D., and Bhaskaran, P. K. (2008). Application of wave model for weather routing of ships in the north indian ocean. *Natural Hazards*, 44(3):373–385.

Prashant, G. (2017). Cross-validation in machine learning. https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f. Accessed: 2022-05-02.

Project Cargo Journal (2019). G2 ocean: "project cargo will become a bigger part of our income". https://www.projectcargojournal.com/shipping/2019/12/04/g2-ocean-project-cargo-will-become-a-bigger-part-of-our-income/. Accessed: 2022-05-16.

Project Cargo Journal (2021). G2 ocean sees southeast asia, africa rising to project cargo prominence. https://www.g2ocean.com/g2-ocean-sees-southeast-asia-africa-rising-to-project-cargo-prominence/. Accessed: 2022-05-01.

Ripley, B. and Venables, W. (2022). Feed-forward neural networks and multinomial

log-linear models. https://cran.r-project.org/web/packages/nnet/nnet.pdf. Accessed: 2022-05-03.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.

Russel, S. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Pearson (2010).

Shao, W., Zhou, P., and Thong, S. K. (2012). Development of a novel forward dynamic programming method for weather routing. *Journal of marine science and technology*, 17(2):239–251.

Ship & Bunker (2022). World bunker prices. https://shipandbunker.com/prices. Accessed: 2022-05-18.

Stopford, M. (2009). *Maritime economics*. Routledge, 3 edition.

Svozil, D., Kvasnicka, V., and Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and intelligent laboratory systems*, 39(1):43–62.

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

Tunnicliffe, D. (2021). Extra trees, please. https://towardsdatascience.com/extra-trees-please-cec916e24827. Accessed: 2022-05-03.

UNCTAD (2021). Review of maritime transport 2021. In *United Nations Conference on Trade and Development, Geneva, Switzerland*.

Wang, C. and Chrétienne, P. (1993). Routage des voiliers et programmation dynamique. *RAIRO-Operations Research*, 27(1):61–76.

Wang, H., Mao, W., and Eriksson, L. (2017). Benchmark study of five optimization algorithms for weather routing. In *International Conference on Offshore Mechanics and Arctic Engineering*, volume 57748, page V07BT06A023. American Society of Mechanical Engineers.

Wang, K., Yan, X., Yuan, Y., and Li, F. (2016). Real-time optimization of ship energy

efficiency based on the prediction technology of working condition. *Transportation Research Part D: Transport and Environment*, 46:81–93.

Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82.

Zheng, J., Zhang, H., Yin, L., Liang, Y., Wang, B., Li, Z., Song, X., and Zhang, Y. (2019). A voyage with minimal fuel consumption for cruise ships. *Journal of Cleaner Production*, 215:144–153.

Zis, T., North, R. J., Angeloudis, P., Ochieng, W. Y., and Harrison Bell, M. G. (2014). Evaluation of cold ironing and speed reduction policies to reduce ship emissions near and at ports. *Maritime Economics & Logistics*, 16(4):371–398.

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

Appendix

A1 Descriptive statistics

Table A1.1: [Class L] Descriptive statistics of the L-class dataset

n = 14 307								
Feature	n_missing	mean	sd	p0	p25	p50	p75	p100
fuel_consumption	0	23.44	7.78	5.10	18.30	24.10	29.40	59.20
speed_made_good	0	13.29	1.67	5.28	12.30	13.45	14.46	18.38
wave_period	0	7.77	2.17	1.67	6.37	7.84	9.23	15.84
wave_height	0	1.89	1.02	0.03	1.21	1.76	2.40	10.34
draft	0	9.71	1.87	5.90	8.02	9.80	11.41	12.77
wind_speed	0	12.61	6.17	0.10	8.00	12.25	16.50	43.46
wind_dir_interact	0	-94.33	772.50	-3287.17	-602.93	0.00	431.72	3360.46
wave_dir_interact	0	-8.72	107.34	-791.21	-74.40	0.00	58.34	527.49
age	0	4.23	2.46	0.00	2.00	4.00	6.00	9.00

Table A1.2: [Class I] Descriptive statistics of the I-class dataset

n = 5 729								
Feature	n_missing	mean	sd	p0	p25	p50	p75	p100
fuel_consumption	0	25.95	9.45	5.10	19.80	26.40	32.00	47.30
speed_made_good	0	13.60	1.67	5.42	12.64	13.65	14.75	18.91
wave_period	0	7.66	2.17	2.05	6.23	7.68	9.14	17.69
wave_height	0	1.99	1.13	0.09	1.20	1.78	2.59	10.44
draft	0	9.60	1.76	0.00	8.34	9.83	11.16	12.37
wind_speed	0	13.18	6.62	0.05	8.28	12.62	17.48	49.12
wind_dir_interact	0	-117.33	803.00	-3218.53	-638.21	0.00	399.67	3526.58
wave_dir_interact	0	-7.77	118.62	-632.74	-79.54	0.00	62.23	666.35
age	0	14.68	3.61	9.00	12.00	14.00	18.00	22.00

A2 Speed performance metrics

Table A2.1: [Class L] Performance metrics of speed prediction model

Models	sMAPE (%)	RMSE	MAE	RSquared
LM (log-transformed)	7.284	0.312	0.212	0.398

A3 Fuel consumption - prediction error distribution

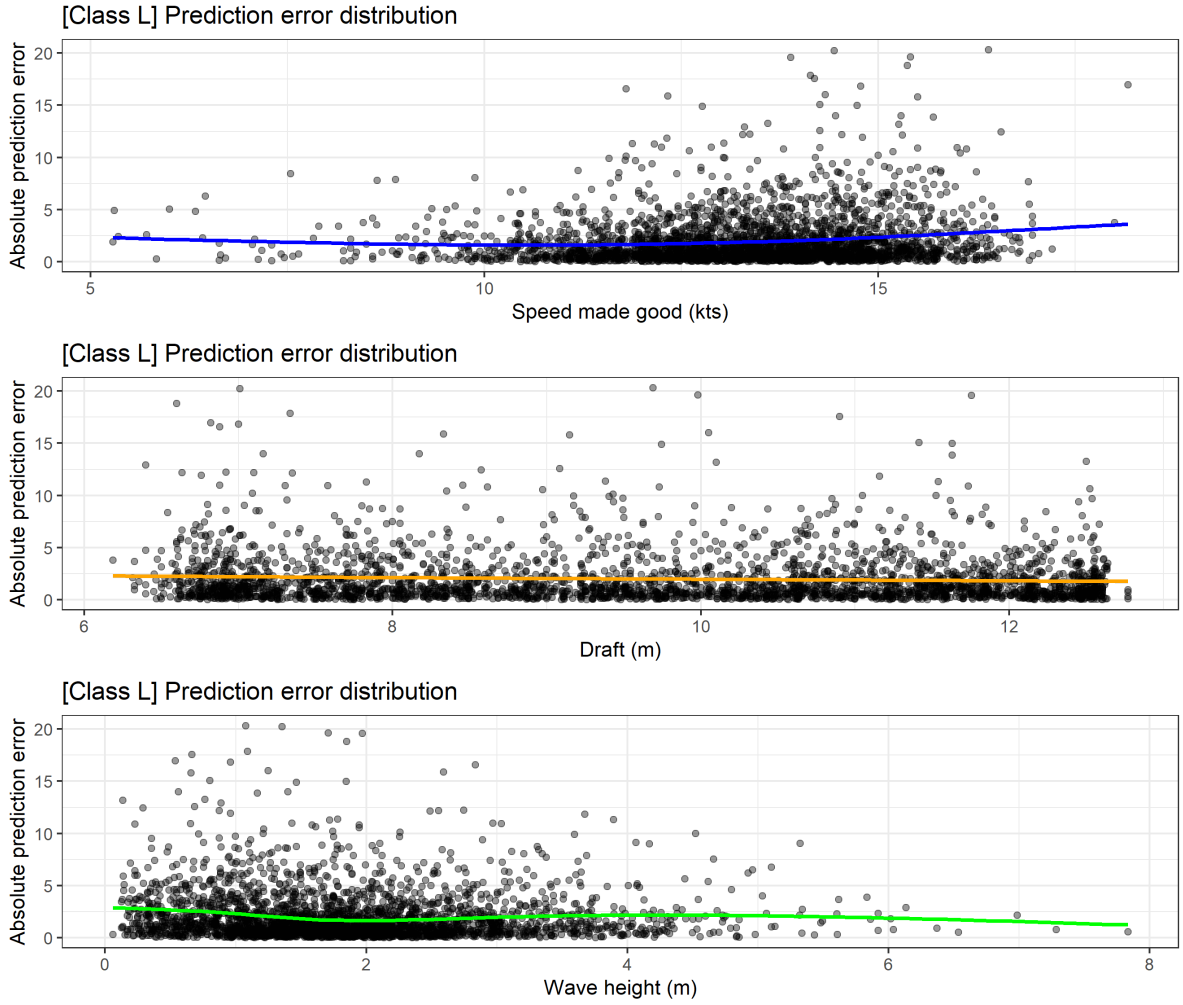

Figure A3.1: Absolute prediction error on fuel consumption sorted by *speed_made_good* (kts), *draft* (m), and *wave_height* (m)

A4 Route selection

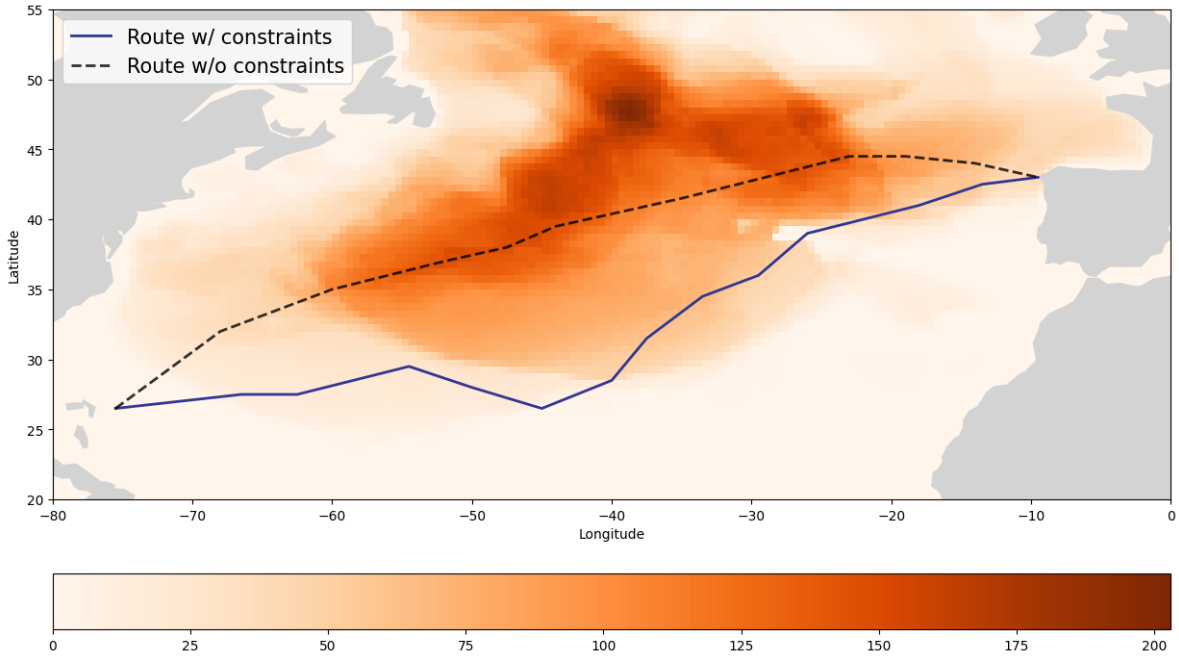

Figure A4.1: Example of optimal route in February 2010.

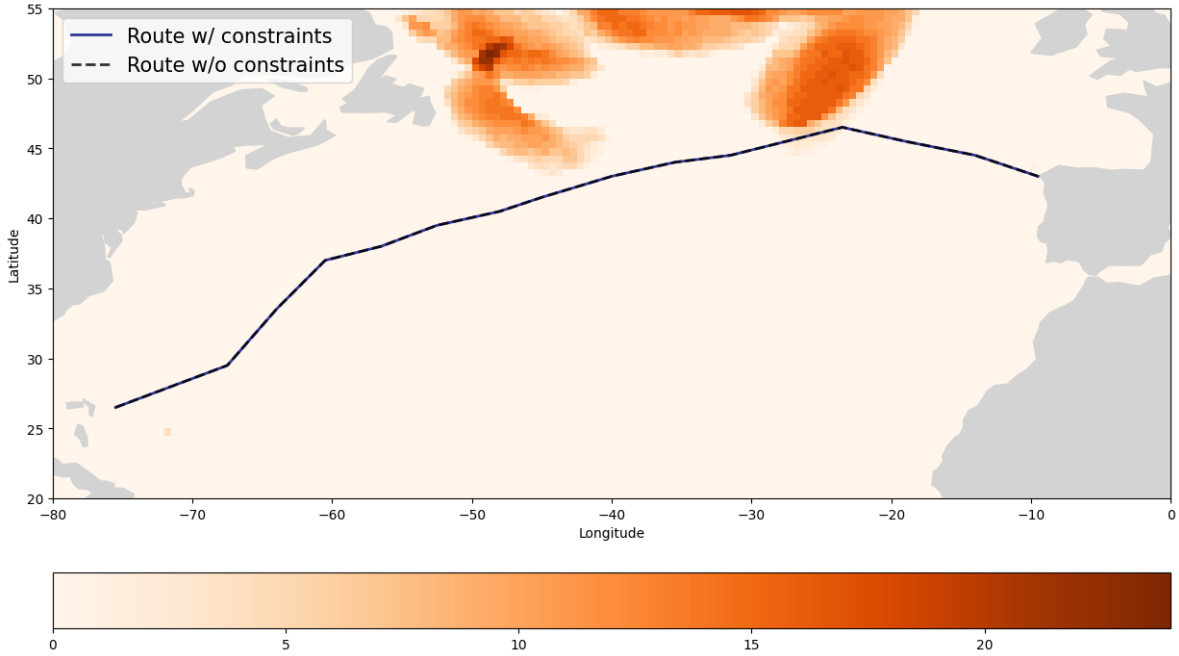

Figure A4.2: Example of optimal route in September 2015

A5 Cost estimations

Table A5.1: [Class I] Results from Comparative Analysis of Vessels Transporting Weather Sensitive deck cargo and Regular Cargo

	Change in Parameters (%)			Incremental Cost (%)				
Month	Distance (km)	Time (hrs)	Fuel (mt)	Mean	Max.	Min.	Var.	Std.
January	0.82	0.82	1.64	2.04	8.90	0.00	7.05	2.66
February	2.03	2.03	1.86	2.48	12.90	0.00	18.58	4.31
March	1.99	1.99	2.66	1.62	14.50	0.00	12.98	3.60
April	0.13	0.13	0.24	0.11	0.90	0.00	0.08	0.29
May	0.00	0.00	0.01	0.01	0.10	0.00	0.00	0.03
June	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
July	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
August	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
September	0.13	0.13	0.26	0.11	0.70	0.00	0.05	0.23
October	0.98	0.98	0.88	0.41	2.70	0.00	0.53	0.73
November	0.73	0.73	0.82	0.37	1.90	0.00	0.42	0.65
December	1.84	1.84	2.49	2.76	11.10	0.00	12.67	3.56